



Protist Diversity and Metabolic Strategy in Freshwater Lakes Are Shaped by Trophic State and Watershed Land Use on a Continental Scale

 Rebecca E. Garner,^{a,f}
 Susanne A. Kraemer,^{a,b,c,f}
 Vera E. Onana,^{a,f}
 Yannick Huot,^{d,f}
 Irene Gregory-Eaves,^{e,f}
 David A. Walsh^{a,f}

^aDepartment of Biology, Concordia University, Montreal, Quebec, Canada

^bGenome Centre, Department of Microbiology & Immunology, McGill University, Montreal, Quebec, Canada

^cDepartment of Civil Engineering, McGill University, Montreal, Quebec, Canada

^dDépartement de géomatique appliquée, Université de Sherbrooke, Sherbrooke, Quebec, Canada

^eDepartment of Biology, McGill University, Montreal, Quebec, Canada

^fGroupe de recherche interuniversitaire en limnologie, Montreal, Quebec, Canada

ABSTRACT Protists play key roles in aquatic food webs as primary producers, predators, nutrient recyclers, and symbionts. However, a comprehensive view of protist diversity in freshwaters has been challenged by the immense environmental heterogeneity among lakes worldwide. We assessed protist diversity in the surface waters of 366 freshwater lakes across a north temperate to subarctic range covering nearly 8.4 million km² of Canada. Sampled lakes represented broad gradients in size, trophic state, and watershed land use. Hypereutrophic lakes contained the least diverse and most distinct protist communities relative to nutrient-poor lakes. Greater taxonomic variation among eutrophic lakes was mainly a product of heterotroph and mixotroph diversity, whereas phototroph assemblages were more similar under high-nutrient conditions. Overall, local physicochemical factors, particularly ion and nutrient concentrations, elicited the strongest responses in community structure, far outweighing the effects of geographic gradients. Despite their contrasting distribution patterns, obligate phototroph and heterotroph turnover was predicted by an overlapping set of environmental factors, while the metabolic plasticity of mixotrophs may have made them less predictable. Notably, protist diversity was associated with variation in watershed soil pH and agricultural crop coverage, pointing to human impact on the land-water interface that has not been previously identified in studies on smaller scales. Our study exposes the importance of both within-lake and external watershed characteristics in explaining protist diversity and biogeography, critical information for further developing an understanding of how freshwater lakes and their watersheds are impacted by anthropogenic stressors.

IMPORTANCE Freshwater lakes are experiencing rapid changes under accelerated anthropogenic stress and a warming climate. Microorganisms underpin aquatic food webs, yet little is known about how freshwater microbial communities are responding to human impact. Here, we assessed the diversity of protists and their myriad ecological roles in lakes varying in size across watersheds experiencing a range of land use pressures by leveraging data from a continental-scale survey of Canadian lakes. We found evidence of human impact on protist assemblages through an association with lake trophic state and extending to agricultural activity and soil characteristics in the surrounding watershed. Furthermore, trophic state appeared to explain the distributions of phototrophic and heterotrophic protists in contrasting ways. Our findings highlight the vulnerability of lake ecosystems to increased land use and the importance of assessing terrestrial interfaces to elucidate freshwater ecosystem dynamics.

Editor Ashley Shade, Michigan State University

Copyright © 2022 Garner et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Rebecca E. Garner, rebecca.garner@mail.concordia.ca, or David A. Walsh, david.walsh@concordia.ca.

The authors declare no conflict of interest.

Received 1 April 2022

Accepted 13 May 2022

Published 22 June 2022

KEYWORDS microbial eukaryotes, trophic state, plankton, phototrophy, heterotrophy, mixotrophy, human impact

Protists have evolved a vast morphological and ecological diversity (1). In aquatic ecosystems, protists play key roles in the transfer of energy and nutrients by converting sunlight into chemical energy, remineralizing organic matter, controlling microbial biomass, feeding higher trophic levels, and maintaining symbioses, some of which recruit prokaryotic metabolisms (2–4).

Elucidating protist diversity in lakes is relevant for clarifying microbial distributions across a wide array of environmental conditions and for tracking the health of critical freshwaters. Covering <1% of Earth's surface (5), lakes contribute disproportionately to the global carbon cycle (6–8) and hold essential water resources (9). Lakes display a rich environmental heterogeneity, generated by integrating fluxes of materials and energy from their catchments and airsheds (10, 11). Freshwater lakes are hotspots of biodiversity, collectively containing higher levels of eukaryote richness and endemism (12) and protist community turnover (13) than the marine and terrestrial realms. In the Anthropocene, lakes are increasingly altered by eutrophication (14), warming temperatures (15), deoxygenation (16), salinization (17), and myriad other persistent and emergent stressors (18). There is accumulating evidence that anthropogenic modifications to lake habitats affect protist assemblages (19–21), in turn influencing ecosystem dynamics.

Efforts to study protists from a variety of biomes have unearthed a vast diversity and begun to map their distributions on a global scale (22–24). Such investigations have provided insights into protist environmental preferences and community assembly processes (25–28), food web dynamics (29–33), symbioses (34–36), viruses (37), functional traits (38), and bioindicator values (39, 40). Large-scale lake surveys have shown that protist assemblages are shaped by broad biogeographic patterns (41–43) but are also influenced by local environmental factors and interactions with bacteria (44). Phytoplankton surveys recapitulate these observations while exclusively investigating photosynthetic taxa including *Cyanobacteria* (45–47). However, there is no clear understanding yet of the distributions of different trophic life history strategies and the environmental drivers underlying their diversity. This knowledge gap is particularly glaring, since heterotrophy is likely the most abundant trophic mode (48). Meanwhile, mixotrophs have dramatically altered our view of plankton food webs by combining primary production and prey consumption (49), sometimes surpassing obligate heterotrophs as the leading grazers of bacteria in lakes (50, 51).

In this study, 18S rRNA gene amplicon sequencing was used to investigate the distributions of protists in the surface waters of 366 freshwater lakes across a north temperate to subarctic continental range. We hypothesized that (i) protist diversity at the local scale and community turnover decrease under high-nutrient conditions; (ii) trophic groups respond to different environmental factors, and specifically, phototrophs are more sensitive than heterotrophs to bottom-up resource availability; and (iii) because lakes integrate their catchments, protist diversity in lakes should reflect watershed conditions. This project was conducted within the LakePulse survey, which sampled hundreds of lakes of different sizes in watersheds across a gradient of human impact with the primary aim to assess lake health through a multidisciplinary lens (52). The current study fills a gap in the mapping of microbial biogeography through this first standardized assessment of protist diversity across Canada, which stewards the greatest abundance of lakes worldwide (53). Our study draws attention to the diversity of protists and the ecological patterns that emerge in a broad collection of newly explored habitats being reshaped by increasing human impact.

RESULTS

Sampled lakes and watersheds display high environmental heterogeneity.

Protist assemblages were surveyed in the euphotic zones of 366 freshwater to oligosaline lakes in 12 ecozones across Canada (43 to 68°N, 53 to 141°W) (Fig. 1). Watersheds

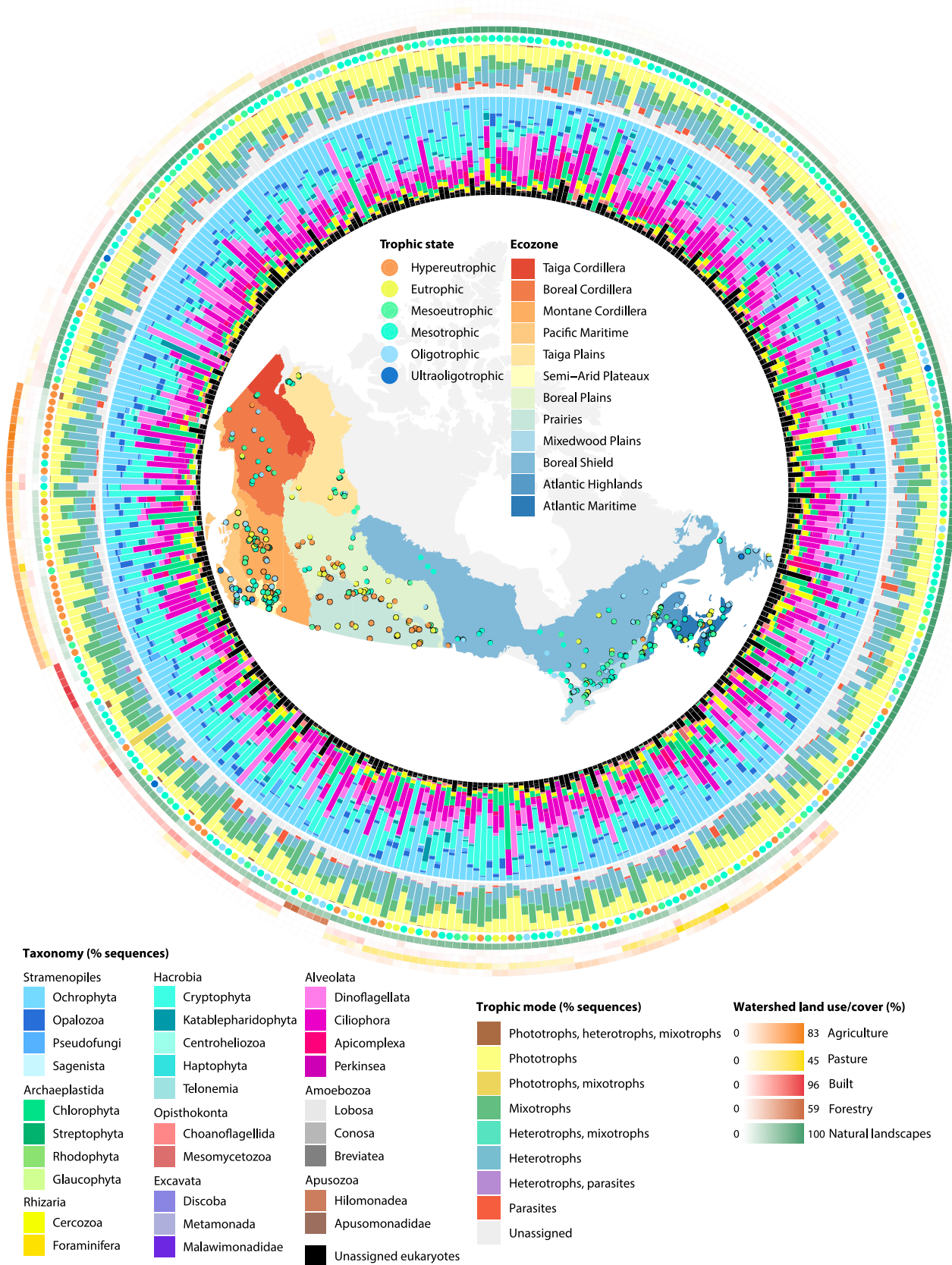


FIG 1 Diversity and distributions of protists across 366 Canadian lakes. Lake trophic states and ecozones are shown on the map of sampling sites as coloured circles and polygons, respectively. The relative sequence abundance of protist taxonomic divisions in each lake is represented in the inner (Continued on next page)

ranged widely in area (0.3 to 9,332.3 km²) and were characterized by a variety of human population densities (0 to 3,785 people/km²) and land use, including different proportions of crop agriculture (0 to 81%) and built development (0 to 93%). Lakes had a wide range of surface areas (0.05 to 99.66 km²) and maximum depths (1 to >150 m) and were either vertically mixed (~40% of lakes) or thermally stratified (~60%) at the time of sampling. Physicochemical conditions differed substantially across lakes, as represented by a broad pH gradient (5.6 to 10.2) and ultraoligotrophic to hypereutrophic states evaluated by total phosphorus (TP) concentrations (2 to 2,484 µg/L) (Fig. 1).

Sampled lakes and watersheds reflected regional variation in environmental conditions, including anthropogenic gradients (see Fig. S1 in the supplemental material). Lakes in northern Canada (Taiga Cordillera, Boreal Cordillera, and Taiga Plains ecozones) were subject to the coldest climates and lowest intensities and proportions of land use within their watersheds. Lakes in western Canada (Montane Cordillera, Pacific Maritime, and Semi-Arid Plateaux) were the deepest on average and located in watersheds with the largest proportions of harvested forests. In central Canada (Boreal Plains and Prairies), a region dominated by plains and prairies with extensive agriculture, lakes were generally shallow, exposed to winds, productive, and high in pH, carbon, ions, and nutrients. Lakes in eastern Canada (Mixedwood Plains, Boreal Shield, Atlantic Highlands, and Atlantic Maritime) had the warmest surface waters and generally the most built-up landscapes, a feature shared with watersheds in the Pacific Maritime. Overall, nutrient-, ion-, and carbon-rich waters, and high-pH lake conditions were most often observed in agricultural watersheds with alkaline soils (Fig. S2).

Lakes support taxonomically and functionally diverse protist assemblages.

Eukaryotic diversity was assessed through the sequencing of 18S rRNA gene fragments amplified from DNA collected in 0.22- to 100-µm surface water particles. A total of 15,848 amplicon sequence variants (ASVs) were inferred in 17,749,930 sequences across 366 lake samples. A final data set of 13,046 putative protist ASVs encompassing 14,622,273 sequences was retained after ASVs assigned to animals, fungi, and plants were removed (Table S1). The rarefaction of pooled samples showed that the sampling of new ASVs plateaued toward 2,000,000 sequences, signaling that the global sequencing effort had exhaustively captured the protist diversity targeted by the primer pair (Fig. S3). The most abundant taxonomic groups were Ochrophyta (24% of all sequences), Cryptophyta (18%), Ciliophora (15%), and Dinoflagellata (11%) (Fig. 1). Lineages with the highest ASV richness were Ochrophyta (22% of all ASVs), Dinoflagellata (12%), and Chlorophyta (11%).

We analyzed ASV incidence to assess the contribution of individual assemblages to total landscape diversity. New genotypes accumulated at a high rate in the first ~100 randomly ordered assemblages, followed by a gradual deceleration (Fig. 2A). The majority of ASVs were restricted to one or a few lakes (Fig. 2B). A smaller number of ASVs were distributed widely, including one ASV assigned to *Cryptomonas curvata* (Cryptophyta) that was ubiquitous yet highly variable in relative abundance (0.0031 to 64%) across all the lakes sampled.

Trophic functions were assigned to ASVs representing 85% of the total sequence space by leveraging natural history descriptions summarized from the literature (Fig. 1; Table S1). Of the functionally annotated ASVs, most were classified as obligate phototrophs (32% of all sequences). The most abundant phototrophs were classified as Chrysophyceae (Ochrophyta; 16% of all sequences), Bacillariophyta (Ochrophyta; 3%), and Chlorophyceae (Chlorophyta; 5%). Bacterivory (20% of all sequences) and mixotrophy (20%) were the next most abundant trophic modes. The most abundant mixotrophs were classified as Cryptophyceae (Cryptophyta; 17% of all sequences) and Dinophyceae (Dinoflagellata; 2%). Cytotrophy (i.e., feeding on other protists), parasitism, commensalism,

FIG 1 Legend (Continued)

tract of bar plots. The middle tract of bar plots shows the relative sequence abundance of trophic modes. The heat map on the outer edge illustrates the proportions of land use and land cover associated with the watershed of each lake, whose trophic state is represented in an adjacent coloured circle. Watershed land use proportions are hierarchically clustered to highlight the relationship between agriculture and trophic state.

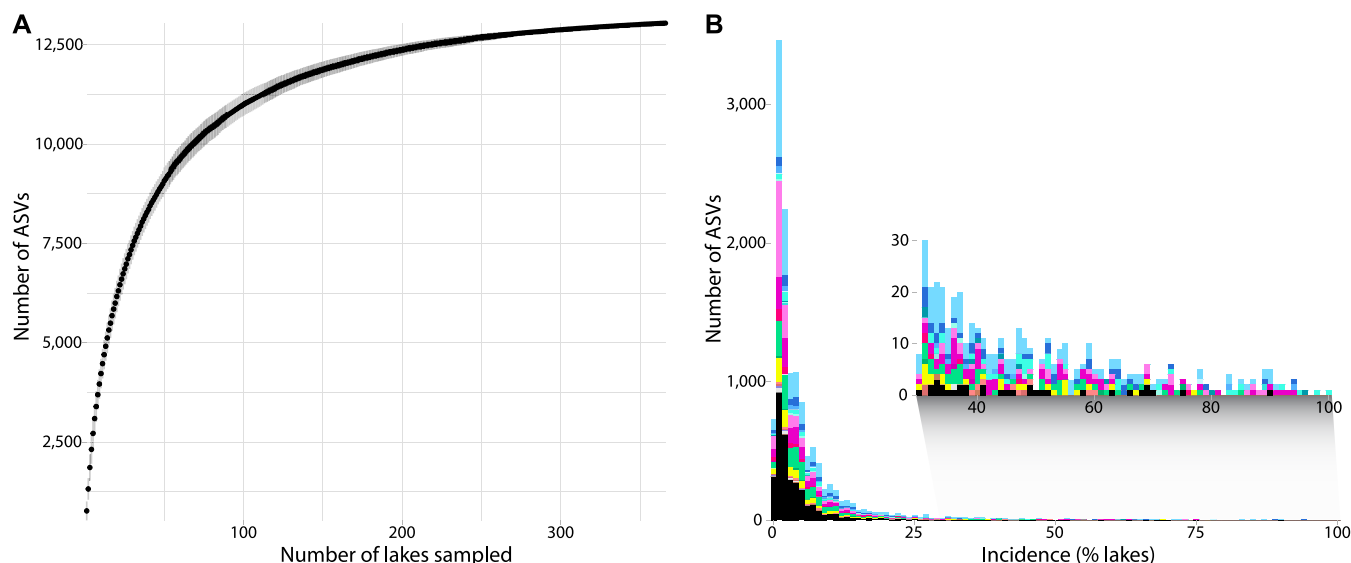


FIG 2 Accumulation and incidence of genotypes across lakes. (A) Accumulation curve of genotypes in a random ordering of lakes. Vertical bars are standard deviations. The accumulation of new genotypes was rapid in the first ~100 lakes, followed by a gradual deceleration. (B) Incidence of genotypes across lakes. Most taxa are distributed across one or a few lakes, whereas a few taxa (magnified in the inset) are widely distributed. ASV taxonomic classifications are coloured according to the taxonomic divisions in Fig. 1.

saprotrophy, and osmotrophy were detected to a lesser extent. Parasites (3% of all sequences) were most abundant in the Coccidiomorpha (Apicomplexa; 2%) and Oomycota (Pseudofungi; 1%). Heterotrophs—which broadly encompassed bacterivores, cytrotrophs, saprotrophs, and osmotrophs—accounted for 27% of all sequences. Heterotrophy was most abundant in the ciliates Spirotrichea (9% of all sequences), Oligohymenophorea (2%), and Litostomatea (2%) and in other lineages, including Bicoecia (Opalozoa; 3%) and Katablepharidaceae (Katablepharidophyta; 2%).

Genotypic similarity to known protist 18S rRNA gene diversity was assessed through the global alignment of ASVs to V7 region fragments in the Protist Ribosomal Reference database (PR²) (Fig. S4). Of the 12,511 ASVs that met the threshold global alignment length, the majority showed high sequence similarity with references in the database. Most ASVs (58%) were $\geq 96\%$ identical to references. ASVs with 100% similarity to reference sequences occupied 56% of the total sequence space, while ASVs with $\geq 96\%$ similarity occupied 91% of sequences. The most novel genotypes included the 738 ASVs with $< 90\%$ sequence similarity to PR² references and were either not assigned to a supergroup (415 ASVs) or primarily classified as Opisthokonta (141 ASVs), Alveolata (88 ASVs), or Stramenopiles (30 ASVs).

Local diversity at each lake was estimated by richness (67 to 1,275 ASVs) (Fig. S5), the Shannon index (0.17 to 5.67) (Fig. 3A), Pielou's evenness index (0.04 to 0.83), and Faith's phylogenetic diversity index (8.15 to 76.04). The Shannon index was negatively correlated with magnesium ($r = -0.33$), total nitrogen (TN; $r = -0.32$), dissolved inorganic carbon (DIC; $r = -0.31$), and potassium ($r = -0.30$) concentrations (all correlations had P values of < 0.001). Analyses of variance (ANOVAs) followed by *post hoc* Tukey's tests comparing the association of trophic state with local diversity showed that mean richness, Shannon diversity, and evenness were significantly lower in hypereutrophic lakes than in eutrophic, mesoeutrophic, mesotrophic, or oligotrophic lakes (all $P \leq 0.002$) (Fig. 3B).

Protist assemblages vary regionally and across lake trophic states. Next, we looked at how communities varied in taxonomic and phylogenetic composition among lakes. A principal component analysis (PCA) of ASV assemblages showed a clear pattern of taxonomic variation by lake trophic state and ecozone along the first dimension (Fig. 4A). Assemblages in the typically nutrient- and ion-rich lakes of the Prairies and Boreal Plains were distinguished from assemblages in the lower-nutrient lakes of the Boreal Shield and other eastern regions. Cryptophyte diversity (*Cryptomonas*, *Geminigera*, *Plagioselmis*, and

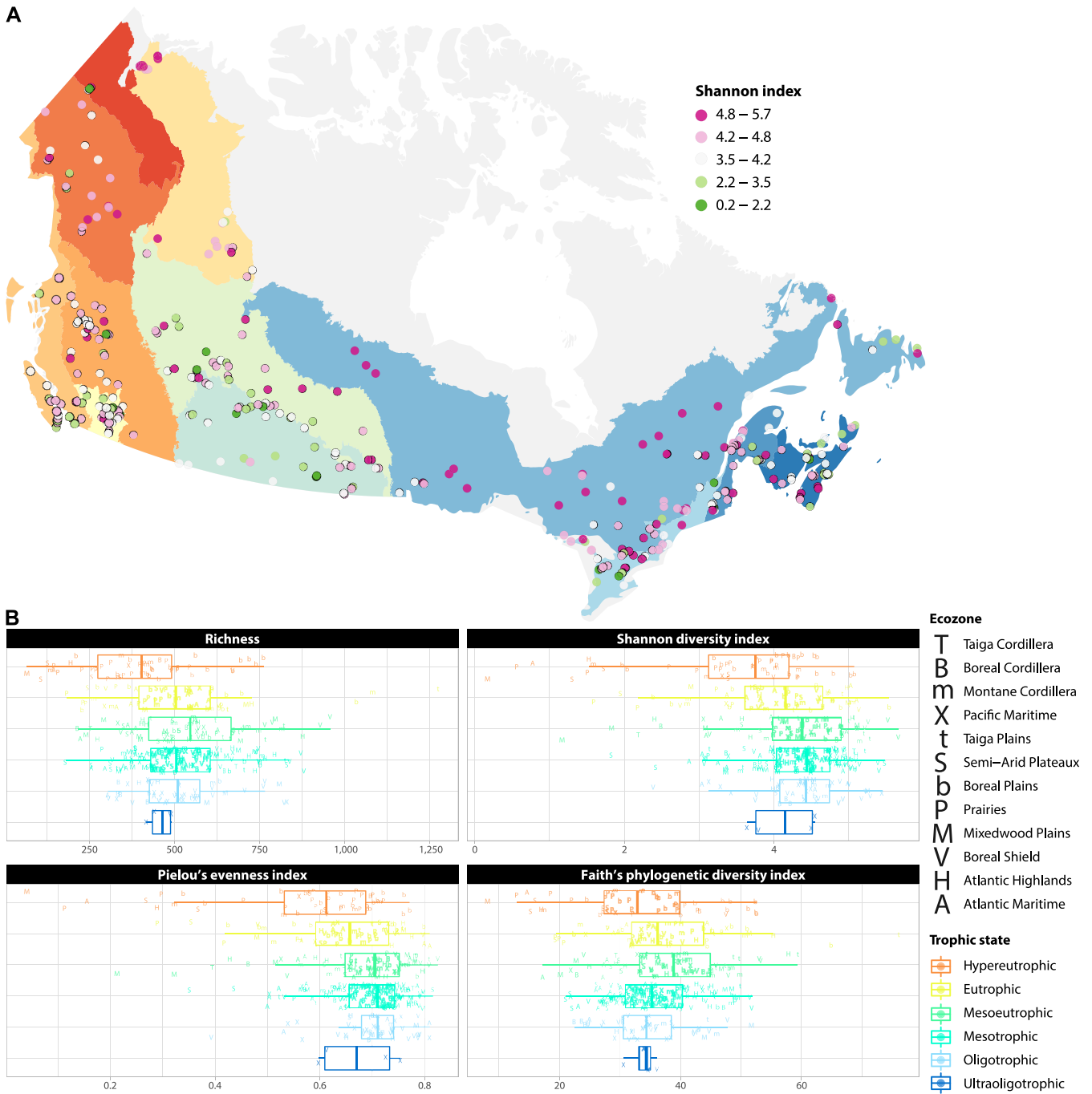


FIG 3 Local protist diversity of each lake. (A) Rarefied Shannon diversity index calculated for each protist assemblage across the Canadian landscape. Ecozones are identified in the key of Fig. 1. (B) Local diversity metrics categorized by lake trophic state. TP concentration thresholds and sample sizes of lakes categorized under each trophic state are as follows: ultraoligotrophic (TP concentration, <4 μg/L; n = 4), oligotrophic (4 to 10 μg/L; n = 48), mesotrophic (10 to 20 μg/L; n = 132), mesoeutrophic (20 to 35 μg/L; n = 71), eutrophic (35 to 100 μg/L; n = 61), and hypereutrophic (>100 μg/L; n = 50).

Komma species) contributed the most strongly to the variation among assemblages (Fig. 4A). A principal coordinate analysis (PCoA) of generalized UniFrac distances between assemblages showed patterns of phylogenetic variation that were highly congruent with the taxonomic variation observed in the PCA, as evaluated by an RV coefficient correlating the first two dimensions of each ordination (RV = 67%, P = 1.3 × 10⁻⁹⁸) (Fig. 4B).

Taxonomically distinct assemblages as quantified by local contributions to β-diversity (LCBD) were mostly localized in the Prairies and Boreal Plains, with a few high-LCBD

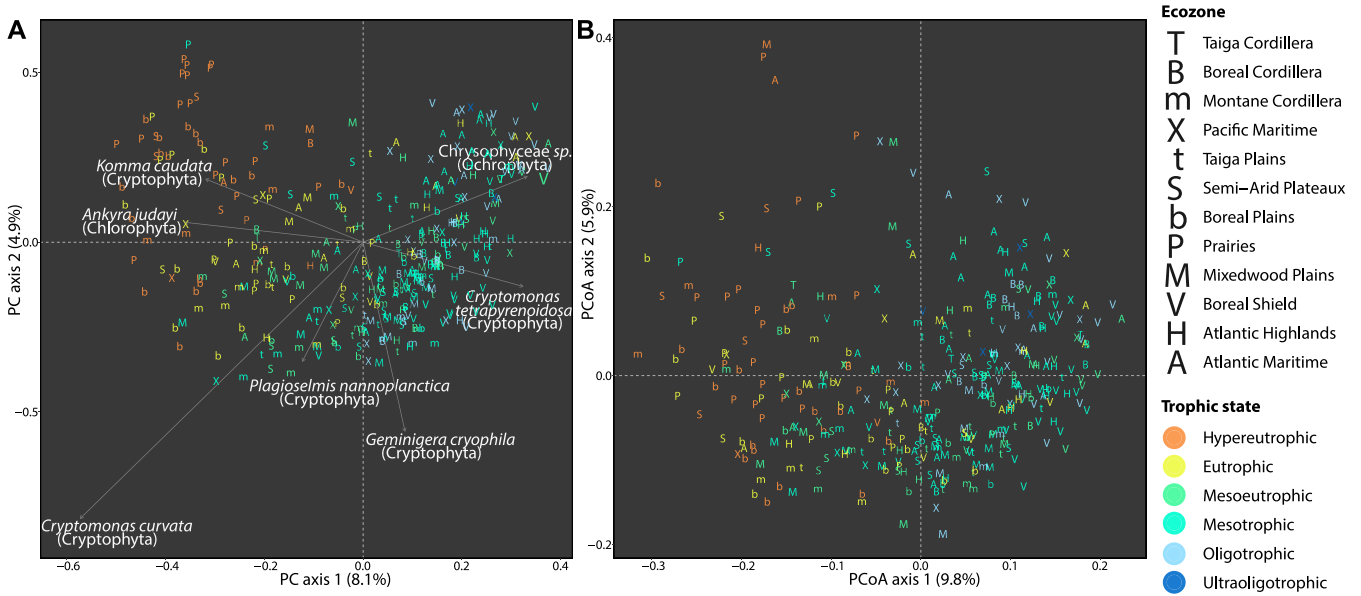


FIG 4 Taxonomic and phylogenetic variation of protist assemblages among lakes. (A) PCA of the taxonomic variation among protist assemblages at the level of individual ASVs. The relative contributions and taxonomic assignments are shown for the top 7 ASVs (an arbitrary cutoff selected for illustrative clarity) contributing to the variation explained by the first two principal component dimensions. (B) PCoA of the phylogenetic variation among protist assemblages.

assemblages scattered across other regions (Fig. 5). Beta-diversity partitioning showed that the taxonomic dissimilarities between assemblages were primarily generated through ASV turnover (75% of total variance) and to a lesser extent through differences in richness (25%). Highly significant ($P < 0.001$) positive correlations were detected between LCBD and a complement of physicochemical variables, including potassium

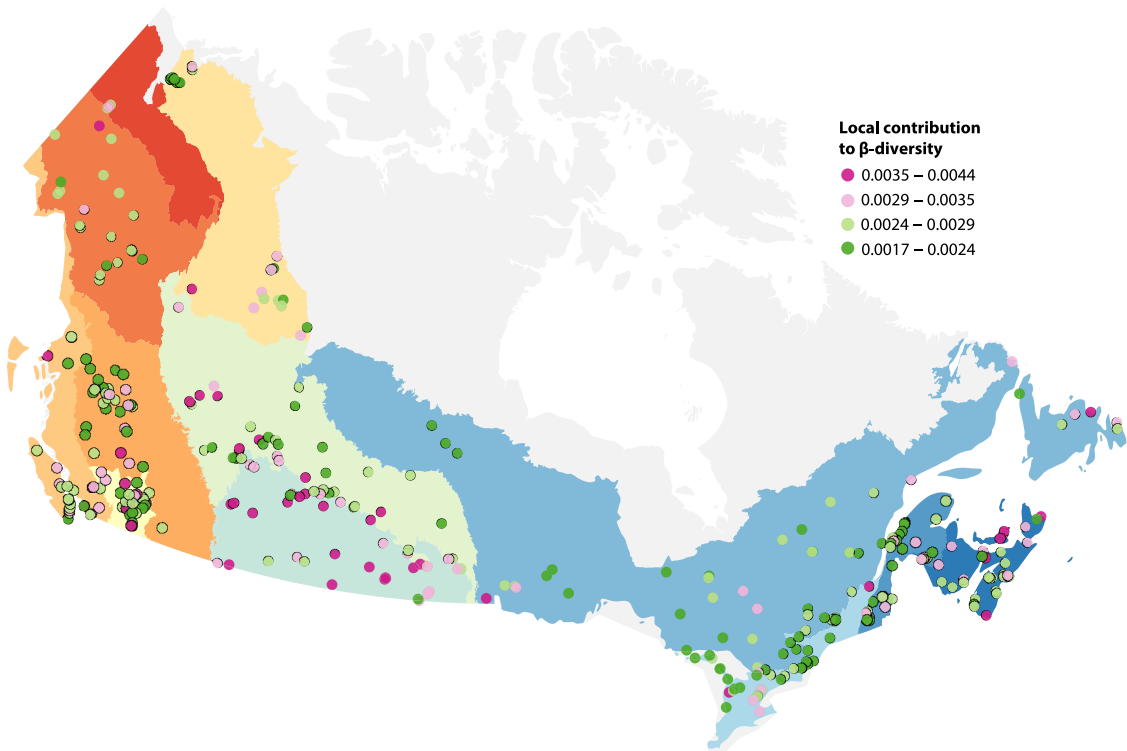


FIG 5 Local contribution to β -diversity (LCBD) of protist assemblages across the Canadian landscape. LCBD describes the taxonomic uniqueness of a given assemblage, i.e., how much the taxonomic composition differs from the rest of the communities in the landscape. Ecozones are identified in the map legend of Fig. 1.

TABLE 1 Percent deviance explained by generalized dissimilarity models (GDMs) fitting the responses of protist assemblages to environmental gradients^a

| Explanatory variables | Deviance explained (%) | | | | |
|-----------------------|------------------------|-----------|---------------------------|--------------|------------|
| | All protists | | Trophic mode ^b | | |
| | Taxonomy | Phylogeny | Phototrophs | Heterotrophs | Mixotrophs |
| Physicochemistry | 38 | 33 | 35 | 42 | 18 |
| Watershed | 15 | 16 | 20 | 19 | NS |
| Morphometry | 6 | 8 | 5 | NS | NS |
| Weather | NS | NS | NS | 3 | NS |
| Geography | NS | NS | NS | 2 | NS |

^aGDMs were constructed using various community response data and categories of environmental explanatory variables. NS, model was not statistically significant ($P \geq 0.05$).

^bAnalysis was performed only on taxonomic composition response data.

($r = 0.53$), TN ($r = 0.50$), DIC ($r = 0.45$), TP ($r = 0.44$), dissolved organic carbon (DOC; $r = 0.42$), magnesium ($r = 0.40$), sodium ($r = 0.38$), sulfate ($r = 0.37$), chlorophyll *a* ($r = 0.33$), lake colour ($r = 0.22$), and pH ($r = 0.22$). LCBD further correlated ($P < 0.001$) with watershed crop cover ($r = 0.35$) and soil pH ($r = 0.27$) and correlated negatively with soil nitrogen ($r = -0.25$). LCBD was negatively correlated with local diversity, estimated as ASV richness ($r = -0.44$) and the Shannon ($r = -0.54$) and Faith's phylogenetic diversity ($r = -0.34$) indices (all $P < 0.001$).

Physicochemical and watershed conditions predict community turnover. Following the previous observations of high community variability across an environmentally heterogeneous set of lakes, we evaluated the drivers of taxonomic and phylogenetic turnover based on five categories of environmental variables: (i) geography (i.e., latitude, longitude, and altitude), (ii) weather, (iii) lake morphometry, (iv) physicochemistry, and (v) watershed characteristics, including land use and edaphic properties.

We employed generalized dissimilarity models (GDMs) to detect nonlinear trends between community turnover and environmental gradients. Physicochemical factors explained the highest deviance in GDMs modeling taxonomic or phylogenetic turnover. In order of decreasing strength, DIC, TP, chlorophyll *a*, magnesium, pH, potassium, lake colour, surface temperature, and DOC were statistically significant predictors of taxonomic turnover, whereas chlorophyll *a*, magnesium, TN, pH, calcium, and colour were significant predictors of phylogenetic turnover. After physicochemistry, watershed characteristics were the most important predictors of community turnover, with both taxonomic and phylogenetic diversity responding most strongly to soil pH and then proportion of cropland cover. The volumetric fraction of soil coarse fragments and natural landscape coverage were additional predictors of taxonomic and phylogenetic turnover, respectively. Lake morphometry, comprising maximum depth, watershed slope, and shoreline circularity, had relatively weak effects on turnover. Weather and geography did not generate statistically significant GDMs. Model deviances are summarized in Table 1, and the partial effects of individual variables are summarized in Fig. S6 and Table S2.

Trophic strategies exhibit contrasting distributions. To examine the distributions of different trophic strategies, we investigated the taxonomic variation of phototroph, heterotroph, and mixotroph communities in separate PCAs (Fig. 6). Reflecting the taxonomic variation emerging at the whole-community level, assemblages of each trophic mode were distinguished by lake trophic state and ecozone along the first dimension. However, phototrophs displayed taxonomic variation patterns contrasting those of heterotrophs and mixotrophs. To compare the taxonomic turnover of phototrophs, heterotrophs, or mixotrophs among lakes of the same trophic state, we measured the distances of assemblages to the trophic state median within the two-dimensional principal coordinate space. The mean distance of phototroph assemblages to the trophic state median (i.e., turnover) was significantly lower in hypereutrophic lakes than in eutrophic, mesoeutrophic, mesotrophic, or oligotrophic lakes (all $P < 0.001$) (Fig. S7). In contrast, the mean distance of heterotroph assemblages was significantly higher in

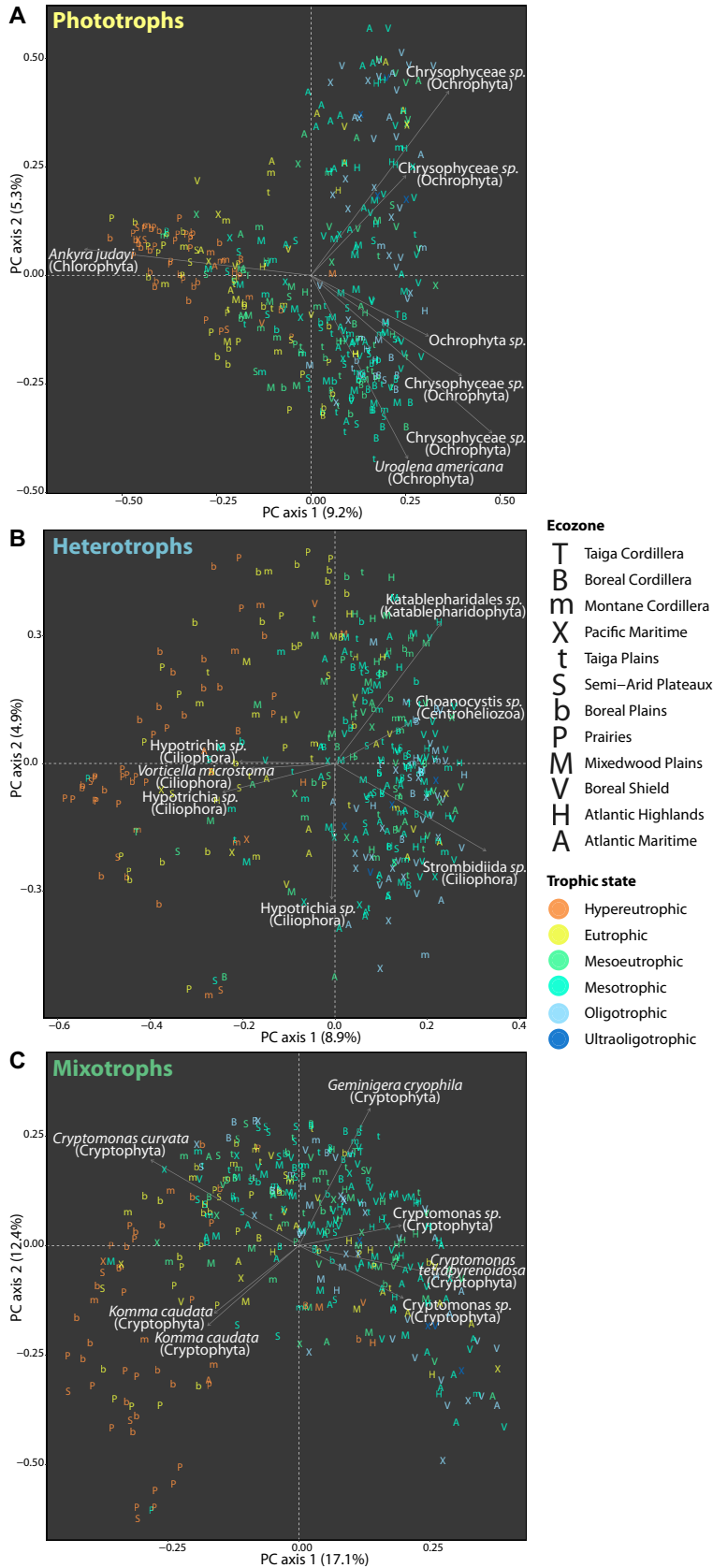


FIG 6 PCA of the taxonomic variation among (A) phototroph, (B) heterotroph, and (C) mixotroph assemblages at the level of individual ASVs. The relative contributions and taxonomic assignments (Continued on next page)

eutrophic lakes than in mesotrophic or oligotrophic lakes (all $P \leq 0.007$) (Fig. S7). Compared with either phototroph or heterotroph assemblages, mixotroph assemblages were highly dispersed within trophic state groups (Fig. S7).

Partitioning protist assemblages by trophic function further allowed us to examine the responses of different groups to environmental conditions. GDMs showed that phototrophs, heterotrophs, and mixotrophs each responded most strongly to physicochemical gradients. Chlorophyll *a*, DIC, potassium, and pH were predictors common to the three trophic modes, while calcium, colour, and chloride were additional predictors of phototroph turnover, and TP, surface temperature, sulfate, TN, colour, and chloride were additional predictors of heterotroph turnover. Chlorophyll *a* followed by potassium were the top predictors of phototroph turnover, while DIC followed by chlorophyll *a* were the top predictors of heterotroph and mixotroph turnover. Watershed characteristics explained an important amount of deviance in phototroph and heterotroph turnover: phototroph turnover was explained by soil pH, crop coverage, soil organic carbon density, and built development, whereas heterotroph turnover was additionally explained by soil coarse fragments. Only phototroph turnover was predicted by lake morphometry variables (watershed slope, lake maximum depth, and circularity). Heterotroph turnover was weakly explained by weather, specifically air temperature and wind speed, and geography, specifically, altitude and distances between lakes. Mixotroph turnover was the least predictable from environmental conditions, as physicochemical factors explained less deviance than for obligate phototroph or heterotroph turnover and GDMs modeled on geography, weather, lake morphometry, and watershed characteristics were not statistically significant. Model deviances are summarized in Table 1, and the partial effects of individual variables are summarized in Fig. S8 and Table S2.

DISCUSSION

Protist diversity unveiled at a continental scale. Establishing a comprehensive perspective of freshwater protist diversity is challenging given the substantial environmental heterogeneity of the millions of lakes distributed globally. Our study fills a sizable gap by mapping protist distributions across hundreds of lakes spanning an area covering nearly 8.4 million km² in the largest study of its kind to employ a standardized sampling scheme (52). Protist diversity was determined in the sunlit surface waters of 366 freshwater lakes varying in size and degree of human impact on the watershed. The sampled biomass bridged cell diameters across 4 orders of magnitude (0.22 to 100 μm), allowing us to recover pico- to microscale organisms, many of which are not resolved under standard light microscopy and which constitute an expansive microbial diversity encompassing the major eukaryotic lineages and trophic strategies. Our global rarefaction analysis showed that the sequencing effort provided a reasonable estimate of the genotypic variation captured by the primer pair. The rapid accumulation of genotypes across sites indicated that sampling hundreds of lakes was required to assess landscape diversity. The integration of our work with the collection of recent large-scale surveys is leading to a synoptic view of protist ecology (23, 24, 28, 42, 43, 54–57).

Surface water assemblages contain high proportions of phototrophs and heterotrophs. Lacustrine protist diversity was dominated by ochrophytes, which accounted for both the highest sequence abundance and ASV richness. Cryptophytes, ciliates, dinoflagellates, and chlorophytes were also highly represented, reflecting taxonomic profiles typical of freshwater biomes on other continents and identified with primer pairs targeting other gene regions (58). The main taxa contributing to the dissimilarity between assemblages were mixotrophic cryptophytes, whose distributions as major bacterivores may be dependent on the occurrence of specific prey (59).

Phototrophic taxa accounted for the greatest richness and relative abundance of ASVs,

FIG 6 Legend (Continued)

are shown for the top 7 ASVs contributing to the variation explained by the first two PC dimensions. Ecozone affiliations and trophic state classifications of lakes are represented by letter symbols and colour, respectively.

perhaps not unexpectedly given that assemblages were sampled from the euphotic zone. Notably, the dominance of phototrophs within the protist fraction appears to be a distinct feature of freshwater photic zones in contrast with the prevalence of heterotrophs in the sunlit ocean and surface soils (13). We showed that heterotrophs and mixotrophs were numerically important groups after phototrophs, which, along with a smaller collection of parasites, illustrates that lake surface waters harbour a broad array of microbial functions linking multiple trophic levels.

Protist communities respond to local environmental conditions. The key environmental drivers of protist diversity in Canadian lakes differed from those observed in previous large-scale surveys. Assemblages showed the strongest responses to physicochemical factors, including nutrient, major ion, and chlorophyll *a* concentrations, pH, and lake colour. The environmental drivers in lakes differed from those in marine and soil ecosystems, where protist diversity on a large scale is generally predicted by temperature (60) and annual precipitation (28), respectively. The differences in environmental filtering between biomes likely represent fundamental differences in the types of habitats and degrees of ecosystem connectivity. Compared with the more spatially continuous and expansive ocean and soil macroenvironments, lakes are fragmented across the landscape, and neighbouring lakes can exhibit widely contrasting physicochemical attributes (10, 61). Lake heterogeneity is amplified by temporal variability and punctuated perturbations. Higher community turnover has previously been measured among lakes than within marine or soil ecosystems (12, 13), which we reassert is linked to physicochemical heterogeneity.

We found that the influence of local environmental conditions far outweighed the effects of geographic variation across the continental extent. Among LakePulse sites, systems with the highest trophic states (typically, Prairies and Boreal Plains lakes) were located at intermediate longitudes, latitudes, and altitudes. A study of Scandinavian boreal lakes spanning a longitudinally aligned and narrower trophic state gradient (oligotrophic to mesoeutrophic) reported that geography explained more protist community variation than water chemistry (42), complementing our assessment that regional physicochemical heterogeneity is a major determinant of protist diversity. Other surveys identified biogeographic patterns of protist diversity structured by the isolation and dispersal limitation of mountain lake communities (41, 43). Geographic barriers (e.g., the Rocky Mountains) did not appear to generate strong compositional divisions in our set of protist assemblages but were identified as having an important influence on the distributions across LakePulse sites of crustacean zooplankton (62), a group with greater dispersal limitation due to their larger body sizes. Instead, we found the greatest taxonomic and phylogenetic divisions between regions distinguished by differences in lake trophic state and other local environmental conditions.

Partitioning protist trophic diversity allowed us to examine how different components of freshwater food webs respond to the environment. We observed contrasting distribution patterns for each trophic mode. Phototroph assemblages among hypereutrophic lakes exhibited significantly lower taxonomic turnover (as evaluated by mean distance to the trophic state median) than lakes at lower nutrient states, whereas heterotroph assemblages did not follow this trend but turned over significantly more rapidly among eutrophic lakes than mesotrophic or oligotrophic lakes. However, the turnover within each trophic mode was predicted by a mostly overlapping suite of physicochemical factors, including chlorophyll *a*, DIC, ions, pH, and colour, although rank order of importance varied among groups. Nutrient (TP and TN) concentrations and surface temperature were exclusive predictors of heterotroph turnover. All of the environmental predictors of mixotroph turnover were common to both phototrophs and heterotrophs, with no predictors unique to mixotrophs. Changes in low levels of chlorophyll *a* were associated with the most rapid turnover in phototroph composition, which is to be expected given that chlorophyll *a* is linked to phytoplankton biomass and phototroph diversity shifts along a lake productivity gradient. The next strongest predictor of phototroph turnover was potassium, which is not a limiting resource (63) but displayed extreme regional variation, peaking in

the Prairies, followed by the Boreal Plains. Heterotroph and mixotroph turnover was primarily predicted by DIC concentrations. Overall, the environmental drivers of heterotroph diversity mostly reflect the bottom-up controls on primary producers traversing multiple trophic levels but are rendered more complex by the added effects of nutrient and temperature factors. Bacterial prey and top-down controls, encompassing predation and parasitism (not measured in this study), likely also determine trophic functional diversity.

Mixotroph distributions were the least aligned with trophic state and the least predictable from environmental conditions, which should be expected for organisms with the metabolic versatility to occupy variable niche spaces. The balance between primary production and prey consumption is dependent on a mixotroph's phenotypic plasticity (64). While mixotrophy is competitively advantageous over obligate phototrophy or heterotrophy under low-nutrient conditions (65), primarily phototrophic mixotrophs prevail in oligotrophic lakes and are replaced by primarily heterotrophic mixotrophs as trophic state increases (66). Another variable driving mixotroph diversity in aquatic ecosystems is light availability (67, 68), which is modulated by lake colour, a predictor identified in this study. Here, too, bacterial prey and zooplankton predators with a preference for nutritious mixotrophs likely also exert controls (69).

Hypereutrophic lakes are taxonomically distinct. Hypereutrophic lakes, located mostly in agricultural watersheds, contained protist assemblages with the lowest diversity and highest taxonomic distinctness (i.e., LCBD) relative to other lakes in the landscape. Specifically, Shannon diversity was inversely related to ion and nutrient concentrations, while taxonomic distinctness tracked with ion- and nutrient-rich conditions. Hypereutrophic conditions potentially filter protist communities by creating relatively extreme conditions (e.g., light attenuation), tolerated by a small number of taxa assembling into uneven communities distinct from those in lakes at lower-nutrient states. Phosphorus is often the limiting nutrient of phytoplankton in numerous freshwater systems (70, 71), yet TP was a predictor of turnover exclusive to heterotrophs. Phosphorus was found to be an important predictor of littoral protist diversity in European lakes (44) and of long-term microeukaryote community turnover evidenced from paleolimnological trends (19). Interactions between protists and *Cyanobacteria* likely also play a role in determining protist assemblages in hypereutrophic lakes, especially as *Cyanobacteria* of the genus *Microcystis* were found to be associated with high-nutrient conditions across LakePulse sites (72).

Ordinations of separate trophic modes showed that the high community variation among eutrophic lakes was generated by heterotroph diversity, whereas obligate phototroph assemblages were the least varied under hypereutrophic conditions. A positive relationship between compositional heterogeneity and trophic state runs counter to the expectation of reduced community variability (e.g., for phytoplankton [73]) that is predicted to follow the leveling of abiotic conditions among lakes induced by land use and eutrophication. Biotic homogenization appears to have trended with long-term climate warming and eutrophication in *Cyanobacteria* (74) and protist (20) assemblages reconstructed from sediment core chronologies. Following our observation of increased compositional heterogeneity among lakes as a function of trophic state, we posit that protist communities in productive lakes are less stable over time, including over the same season, as observed in bacterial time series (75). The temporal fluxes in abiotic conditions prompting succession may be induced by allochthonous inputs or nutrient resuspension from sediments accrued at higher rates in regions of extensive lake use and high populations densities (76). Furthermore, given that many Prairies and Boreal Plains lakes are shallow and exposed to winds, temporary stratification followed by destratification is not uncommon (77). We speculate that taxon replacement linked to land use and eutrophication may force a re-evaluation of human impact on biodiversity. In particular, anthropogenic pressures may not inherently decrease diversity but instead increase turnover, possibly at the expense of rare or specialist taxa disappearing from the landscape pool (78).

Watersheds influence lacustrine protist diversity. While the importance of physicochemical factors for lacustrine protist diversity has been described (42, 44) and elaborated upon in this study, the influence of the watershed, in particular soil properties

and land use, which are often but not entirely correlated with lake physicochemical attributes, until now has not been documented on a continental scale. We found that the taxonomic distinctness of local assemblages (i.e., LCBD) corresponded strongly with the proportion of crop agriculture in the watershed, while surface soil pH was an important predictor of community turnover. Because of their concave topographies and position in the landscape, lakes are recipients of major allochthonous subsidies, with global effects (e.g., carbon storage) that are disproportionate to the spatial extent of lakes (6). The influence of terrestrial catchments on within-lake community dynamics is compelling evidence for why lakes cannot be studied apart from their watersheds.

The filtering of lacustrine protist diversity by watershed soil chemistry points to the influence of external factors on lake conditions. Soil buffering capacity, determined by soil texture, organic matter content, and mineral composition, is a main abiotic control on lake water pH (79). Furthermore, soil properties control the mobility of nutrients and their eventual input into lakes. Specifically, soil pH determines the availability and chemical forms of nutrients, and particle size governs the movement of groundwater carrying released nutrients (80). Soil properties also determine the composition and activity of soil microbiomes, which perform biogeochemical transformations modulating the availability of nutrients (81) and seed potential colonists from soils to lakes. Altered precipitation regimes and warming temperatures associated with climate change are expected to increase soil erosion (82) as well as terrestrial nutrient exports (83).

Given the continual increase in land surface transformed by agricultural production and urbanization (84), accelerated watershed land use conversions are widespread. Changes in soils associated with human activities range from increased nutrient loading and acidification by nitrogenous fertilizers in agriculture (85) to shifts in carbon storage precipitated by changes in land management practices or climate (86). Moreover, land use and climate change interact to increase the frequency and magnitude of nutrient and carbon pulses to water bodies (87). In the interest of securing a healthy future for critical freshwaters, we suggest that current soil chemistry heterogeneity across the landscape can inform predictions about the potential consequences of anthropogenic watershed alterations for lakes. In particular, work can be done to understand the microbial diversity and food web dynamics driven by various soil states in future land use scenarios. Overall, the ability to predict lacustrine protist diversity from watershed conditions, as demonstrated in this study, highlights an expanded potential for monitoring lake ecosystems using remote sensing products (88).

Conclusion. This is the first study to examine the taxonomic and trophic functional variation in protist diversity across the expansive and lake-rich Canadian landscape. We showed that on this continental scale, lakes displayed broad environmental heterogeneity, including substantial variation in local physicochemical conditions driving taxonomic and phylogenetic community turnover. Watershed soil pH and crop agriculture additionally predicted community turnover and exceptional local-scale diversity. Hypereutrophic lakes were found to contain less diverse and more distinct assemblages than lower-nutrient lakes, primarily as a product of their variable heterotroph and mixotroph compositions. In contrast, phototroph assemblages were more similar among hypereutrophic lakes. While phototrophy was the prevailing nutritional strategy in lake euphotic zones, heterotrophy was nearly as numerically important; each of these trophic modes was highly predictable from physicochemical and other environmental factors. Our survey and findings serve as a valuable resource for mapping species distributions and provide a basis for future research into the increasing anthropogenic impact on lake microbiomes.

MATERIALS AND METHODS

Lake selection and sampling. Hundreds of lakes were sampled between July and early September in 2017 to 2019 by the Natural Sciences and Engineering Research Council of Canada (NSERC) Canadian Lake Pulse Network (52). Sampling was timed to coincide with the summertime period of water column thermal stratification, where relevant. Lakes were sampled across 12 terrestrial ecozones, regions defined by landform, geology, and vegetation (89). Lake selection was stratified across lake surface area and watershed land use impact categories to capture natural and human-mediated lake heterogeneity. Only

natural lakes with a maximum depth of at least 1 m and within 1 km from a road were considered. Freshwater to oligosaline lakes (identified as having conductivity of <8 mS/cm and total major ion concentrations of <4,000 mg/L) were selected for this analysis.

Water was collected using an integrated tube sampler from the euphotic zone over a depth of up to 2 m below the surface at the deepest point in the lake (90). The site of maximum lake depth was located by depth sounding with the aid of bathymetric maps where available. The depth of the euphotic zone was estimated as twice the Secchi disk depth. All water sampling equipment was acid washed and rinsed three times with lake water before use. Carboys were stored in icepack-chilled coolers until water filtration later in the day. Water was prefiltered through 100- μ m nylon mesh and vacuum filtered on 47-mm-diameter 0.22- μ m Durapore membranes through a glass funnel at a maximum pressure of 8 inHg. Up to 500 mL of water was filtered until the filter was nearly clogged. Filters were stored in sterile cryovials at -80°C .

18S rRNA gene amplification and sequencing. DNA was extracted using the DNeasy PowerWater kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions with the addition of two optional steps: after bead beating and centrifugation, 1 μ L RNase A was added to samples, followed by 30 min incubation at 37°C . DNA was quantified using the Qubit dsDNA BR assay (Invitrogen, Carlsbad, CA, USA). A \sim 265-bp fragment of the 18S rRNA gene V7 region was amplified with the primers 960F (5'-GGCTAATTTGACTCAACRCG-3') (91) and NSR1438 (5'-GGGCATCACAGACCTGTAT-3') (92) to broadly target microeukaryotes (93). Each PCR contained a total 25- μ L mixture of 14.25 μ L Milli-Q water, 5 μ L 5 \times High-Fidelity buffer, 1.25 μ L of each 10 μ M primer, 0.5 μ L 10 mM deoxynucleoside triphosphates (dNTPs), 0.25 μ L dimethyl sulfoxide (DMSO), 0.5 μ L Phusion DNA polymerase (Thermo Fisher Scientific, Waltham, MA, USA), and 2 ng DNA template. PCR conditions consisted of an initial denaturation at 98°C for 1 min, 30 cycles of 98°C for 10 s, 60°C melting for 30 s, and 72°C for 20 s, and a final extension at 72°C for 5 min. PCR products were loaded with Orange G dye on an ethidium bromide-stained 2% agarose gel and electrophoresed at 40 V for 100 min. DNA bands aligned at the target fragment length against a 100-bp DNA ladder were excised with razor blades and gel extracted with the QIAquick gel extraction kit (Qiagen, Hilden, Germany), modified by final elution into Milli-Q water. PCR products were submitted to Genome Quebec for library barcoding and sequencing of 250-bp paired-end reads in three sequencing runs on an Illumina MiSeq platform.

ASV inference and annotation. Primer sequences were removed in Cutadapt v. 3.1 (94). Trimmed reads were processed into ASVs through DADA2 v. 1.16 (95). Samples were pooled for ASV inference using otherwise default parameters. Taxonomy was assigned with naive Bayesian classification trained on PR² v. 4.12.0 (96). Potentially spurious ASVs were removed by visually inspecting a *de novo* alignment performed in MAFFT (97). To retain only putative protist ASVs, ASVs assigned to Metazoa, Fungi, and Embryophyceae were removed. Taxa were assigned to trophic functional groups as either photoautotrophs, heterotrophs (bacterivores, cytotrophs, saprotrophs, or osmotrophs), mixotrophs, or parasites according to lineage-specific feeding habits summarized by Adl et al. (98).

Sequence similarities with known diversity. A database of 18S rRNA gene references restricted to the V7 region was constructed by applying the 960F/NSR1438 primer pair to PR² v. 4.13.0 in Cutadapt. ASV top hits were queried against the PR² database in BLAST v. 2.6.0+ (99). Sequence identities were reported for ASVs that were globally aligned to references over a length \geq 220 nucleotides.

Environmental data collection. Lake trophic states were assigned based on TP concentration thresholds estimated for Canadian freshwater systems: ultraoligotrophic (TP concentration, <4 $\mu\text{g/L}$), oligotrophic (4 to 10 $\mu\text{g/L}$), mesotrophic (10 to 20 $\mu\text{g/L}$), mesoeutrophic (20 to 35 $\mu\text{g/L}$), eutrophic (35 to 100 $\mu\text{g/L}$), and hypereutrophic (>100 $\mu\text{g/L}$) (100). Meteorological conditions recorded over 7 days leading up to sampling and ice disappearance day data were accessed from ERA5-Land hourly reanalysis (101). Data on watershed slope and lake volume, discharge, and hydraulic residence time were accessed from HydroLAKES v. 1.0 (5). Watershed surface soil properties were accessed from SoilGrids250m (102). Land cover information was compiled as described by Huot et al. (52). Maps were constructed in R with the NAD 83 coordinate reference system and using the coordinates of Canada from the package maps (103) and ecozone shapefiles sourced from the Canada Council of Ecological Areas (89).

Environmental data were categorized into thematic groups of variables. Latitude, longitude, and altitude were categorized as geography variables. Ice disappearance day and meteorological variables (air temperature, precipitations, and net solar radiation) were categorized as weather variables. Lake surface area, circularity, volume, maximum depth, discharge, residence time, watershed slope within 100 m of the shoreline, watershed area, and lake-to-watershed area ratio were categorized as lake morphometry variables. Watershed land use (crop agriculture, pasture, built development, and clear-cut forestry) and natural land cover fractions, human population density, and mean surface soil properties (bulk density of the fine earth fraction, cation exchange capacity, nitrogen, pH, organic carbon density, organic carbon content in the fine earth fraction, volumetric fraction of coarse fragments, clay, sand, and silt) were categorized as watershed variables. Surface water temperature, calcium, magnesium, potassium, sodium, chloride, sulfate, TP, TN, DIC, DOC, and chlorophyll *a* concentrations, pH, and lake colour were categorized as lake physicochemical variables. Missing physicochemical data were replaced with ecozone median values. Highly collinear variables, evaluated by a Pearson's correlation *r* value of \geq 0.7, within all categories except physicochemistry were removed.

Diversity analyses. ASVs were aligned in the SILVA Incremental Aligner v. 1.7.2 (104) against the SILVA 138.1 SSU Ref NR 99 database (August 27, 2020 release) (105). A maximum-likelihood phylogeny was constructed in FastTree v. 2.1.11 using the Generalized Time-Reversible model of nucleotide evolution (106). Phylogenetic dissimilarities between ASV assemblages were calculated as generalized UniFrac distances, which are sensitive to compositional changes in lineages of intermediate abundance (107). Generalized UniFrac distances ($\alpha = 0.5$) were computed using the GUniFrac package.

To deal with uneven total sequence abundance across samples, sequence count composition was scaled to relative abundance. Rarefaction analysis was conducted on the total data set by measuring ASV richness in assemblages randomly subsampled at each 1,000-sequence step. Taxon accumulation was estimated in a random ordering of lakes using 100 permutations in the package *vegan* (108). Local diversity indices (richness, Pielou's evenness, Shannon diversity, and Faith's phylogenetic diversity) were calculated from rarefied community data (i.e., randomly subsampled to the lowest sample abundance equaling 10,069 sequences) in the R packages *vegan* (108) and *picante* (109). PCAs were computed on Hellinger-transformed community data in *vegan*.

LCBD and β -diversity partitioning analyses were performed on β -diversity estimated using 100 permutations from Hellinger-transformed community data in the package *adespatial* (110). RV coefficients were computed from the first two principal components or coordinates in the package *FactoMineR* (111). Nonlinear relationships between β -diversity and untransformed environmental gradients were modeled in GDMs (112, 113) in the package *gdm* (114). To create GDM site-pair tables, pairwise dissimilarities between sites were weighted proportionally to the total number of sequences associated with each sample. Variable selection for GDMs was performed using backward elimination with 100 permutations per step. To assess the dispersion in taxonomic composition among lakes of the same trophic state, the mean Bray-Curtis distances of assemblages to trophic-state medians (i.e., centroids) calculated across the first two principal coordinates were compared using ANOVAs, followed by *post hoc* Tukey's tests.

Data wrangling and statistical analysis were performed in R v. 4.0.2 (115).

Data availability. Sequence data have been deposited in the European Nucleotide Archive under study accession number [PRJEB42538](https://www.ebi.ac.uk/PRJEB42538) (www.ebi.ac.uk). Scripts are accessible from https://github.com/rebeccagamer/lakepulse_protists.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

FIG S1, PDF file, 0.8 MB.

FIG S2, PDF file, 0.8 MB.

FIG S3, PDF file, 0.4 MB.

FIG S4, PDF file, 0.4 MB.

FIG S5, PDF file, 1.4 MB.

FIG S6, PDF file, 7.6 MB.

FIG S7, PDF file, 1.5 MB.

FIG S8, PDF file, 10.1 MB.

TABLE S1, CSV file, 3.9 MB.

TABLE S2, DOCX file, 0.02 MB.

ACKNOWLEDGMENTS

This research was funded by the NSERC Canadian Lake Pulse Network (Strategic Network Grant NETGP-479720) and Canada Research Chairs held by D.A.W., I.G.-E., and Y.H. R.E.G. and V.E.O. were supported by the NSERC CREATE ÉcoLac training program in lake and fluvial ecology. R.E.G. also acknowledges support from a *Fonds de recherche du Québec—Nature et technologies* doctoral research scholarship and the Stephen Bronfman Scholarship in Environmental Studies.

We thank the LakePulse sampling crews and the many landowners, municipal and park employees, lake associations, and First Nations who welcomed and facilitated the sampling effort. We also thank the researchers in LakePulse who contributed to the data set, Zofia Taranu for helpful discussions on statistical methods, and Paul MacKeigan, Cindy Paquette, Marieke Beaulieu, and Katherine Griffiths for their collaboration on data quality assessment and curation.

REFERENCES

1. Simpson AGB, Eglit Y. 2016. Protist diversification, p 344–360. In Kliman RM (ed), *Encyclopedia of evolutionary biology*. Academic Press, Waltham, MA.
2. Gadd GM, Raven JA. 2010. Geomicrobiology of eukaryotic microorganisms. *Geomicrobiol J* 27:491–519. <https://doi.org/10.1080/01490451003703006>.
3. Worden AZ, Follows MJ, Giovannoni SJ, Wilken S, Zimmerman AE, Keeling PJ. 2015. Rethinking the marine carbon cycle: factoring in the multifarious lifestyles of microbes. *Science* 347:1257594. <https://doi.org/10.1126/science.1257594>.
4. Caron DA, Alexander H, Allen AE, Archibald JM, Armbrust EV, Bachy C, Bell CJ, Bharti A, Dyhrman ST, Guida SM, Heidelberg KB, Kaye JZ, Metzner J, Smith SR, Worden AZ. 2017. Probing the evolution, ecology and physiology of marine protists using transcriptomics. *Nat Rev Microbiol* 15: 6–20. <https://doi.org/10.1038/nrmicro.2016.160>.
5. Messenger ML, Lehner B, Grill G, Nedeva I, Schmitt O. 2016. Estimating the volume and age of water stored in global lakes using a geo-statistical approach. *Nat Commun* 7:13603. <https://doi.org/10.1038/ncomms13603>.
6. Tranvik LJ, Downing JA, Cotner JB, Loiselle SA, Striegl RG, Ballatore TJ, Dillon P, Finlay K, Fortino K, Knoll LB, Kortelainen PL, Kutser T, Larsen S, Laurion I, Leech DM, Leigh MS, McKnight DM, Melack JM, Overholt E, Porter JA, Prairie Y, Renwick WH, Roland F, Sherman BS, Schindler DW, Sobek S, Tremblay A, Vanni MJ, Verschoor AM, Von Wachenfeldt E,

- Weyhenmeyer GA. 2009. Lakes and reservoirs as regulators of carbon cycling and climate. *Limnol Oceanogr* 54:2298–2314. https://doi.org/10.4319/lo.2009.54.6_part_2.2298.
7. Beaulieu JJ, DeSontro T, Downing JA. 2019. Eutrophication will increase methane emissions from lakes and impoundments during the 21st century. *Nat Commun* 10:1375. <https://doi.org/10.1038/s41467-019-09100-5>.
 8. Anderson NJ, Heathcote AJ, Engstrom DR, Globocarb Data Contributors. 2020. Anthropogenic alteration of nutrient supply increases the global freshwater carbon sink. *Sci Adv* 6:eaa2145. <https://doi.org/10.1126/sciadv.aaw2145>.
 9. Heino J, Alahuhta J, Bini LM, Cai Y, Heiskanen A, Hellsten S, Kortelainen P, Kotamäki N, Tolonen KT, Vihervaara P, Vilmi A, Angeler DG. 2021. Lakes in the era of global change: moving beyond single-lake thinking in maintaining biodiversity and ecosystem services. *Biol Rev Camb Philos Soc* 96:89–106. <https://doi.org/10.1111/brv.12647>.
 10. Kratz TK, MacIntyre S, Webster KE. 2005. Causes and consequences of spatial heterogeneity in lakes, p 329–347. *In* Lovett GM, Jones C, Turner MG, Weathers KC (ed), *Ecosystem function in heterogeneous landscapes*. Springer, New York, NY.
 11. Williamson CE, Dodds W, Kratz TK, Palmer MA. 2008. Lakes and streams as sentinels of environmental change in terrestrial and atmospheric processes. *Front Ecol Environ* 6:247–254. <https://doi.org/10.1890/070140>.
 12. Strayer DL, Dudgeon D. 2010. Freshwater biodiversity conservation: recent progress and future challenges. *J N Am Benthol Soc* 29:344–358. <https://doi.org/10.1899/08-171.1>.
 13. Singer D, Seppey CVW, Lentendu G, Dunthorn M, Bass D, Belbahri L, Blandenier Q, Debroas D, de Groot GA, de Vargas C, Domaizon I, Duckert C, Izaguirre I, Koenig I, Mataloni G, Schiaffino MR, Mitchell EAD, Geisen S, Lara E. 2021. Protist taxonomic and functional diversity in soil, freshwater and marine ecosystems. *Environ Int* 146:106262. <https://doi.org/10.1016/j.envint.2020.106262>.
 14. Smith VH, Schindler DW. 2009. Eutrophication science: where do we go from here? *Trends Ecol Evol* 24:201–207. <https://doi.org/10.1016/j.tree.2008.11.009>.
 15. O'Reilly CM, Sharma S, Gray DK, Hampton SE, Read JS, Rowley RJ, Schneider P, Lenters JD, MacIntyre PB, Kraemer BM, Weyhenmeyer GA, Straile D, Dong B, Adrian R, Allan MG, Anneville O, Arvola L, Austin J, Bailey JL, Baron JS, Brookes JD, Eyto E, Dokulil MT, Hamilton DP, Havens K, Hetherington AL, Higgins SN, Hook S, Izmeš'eva LR, Joehnk KD, Kangur K, Kasprzak P, Kumagai M, Kuusisto E, Leshkevich G, Livingstone DM, MacIntyre S, May L, Melack JM, Mueller-Navarra DC, Naumenko M, Noges P, Noges T, North RP, Plisnier P, Rigosi A, Rimmer A, Rogora M, Rudstam LG, Rusak JA, Salmaso N, Samal NR, Schindler DE, Schladow SG, Schmid M, Schmidt SR, Silow E, Soylu ME, et al. 2015. Rapid and highly variable warming of lake surface waters around the globe. *Geophys Res Lett* 42:10773–10781. <https://doi.org/10.1002/2015GL066235>.
 16. Jane SF, Hansen GJA, Kraemer BM, Leavitt PR, Mincer JL, North RL, Pilla RM, Stetler JT, Williamson CE, Woolway RI, Arvola L, Chandra S, DeGasperi CL, Diemer L, Dunalska J, Erina O, Flaím G, Grossart H-P, Hambright KD, Hein C, Hejzlar J, Janus LL, Jenny J-P, Jones JR, Knoll LB, Leoni B, Mackay E, Matsuzaki S-IS, McBride C, Müller-Navarra DC, Pierson AM, Pierson D, Rogora M, Rusak JA, Sadro S, Saulnier-Talbot E, Schmid M, Sommaruga R, Thiery W, Verburg P, Weathers KC, Weyhenmeyer GA, Yokota K, Rose KC. 2021. Widespread deoxygenation of temperate lakes. *Nature* 594:66–70. <https://doi.org/10.1038/s41586-021-03550-y>.
 17. Dugan HA, Skaff NK, Doubek JP, Bartlett SL, Burke SM, Krivak-Tetley FE, Summers JC, Hanson PC, Weathers KC. 2020. Lakes at risk of chloride contamination. *Environ Sci Technol* 54:6639–6650. <https://doi.org/10.1021/acs.est.9b07718>.
 18. Reid AJ, Carlson AK, Creed IF, Elision EJ, Gell PA, Johnson PTJ, Kidd KA, MacCormack TJ, Olden JD, Ormerod SJ, Smol JP, Taylor WW, Tockner K, Vermaire JC, Dudgeon D, Cooke SJ. 2019. Emerging threats and persistent conservation challenges for freshwater biodiversity. *Biol Rev Camb Philos Soc* 94:849–873. <https://doi.org/10.1111/brv.12480>.
 19. Capo E, Debroas D, Arnaud F, Perga ME, Chardon C, Domaizon I. 2017. Tracking a century of changes in microbial eukaryotic diversity in lakes driven by nutrient enrichment and climate warming. *Environ Microbiol* 19:2873–2892. <https://doi.org/10.1111/1462-2920.13815>.
 20. Keck F, Millet L, Debroas D, Etienne D, Galop D, Rius D, Domaizon I. 2020. Assessing the response of micro-eukaryotic diversity to the Great Acceleration using lake sedimentary DNA. *Nat Commun* 11:3831. <https://doi.org/10.1038/s41467-020-17682-8>.
 21. Salmaso N, Tolotti M. 2021. Phytoplankton and anthropogenic changes in pelagic environments. *Hydrobiologia* 848:251–284. <https://doi.org/10.1007/s10750-020-04323-w>.
 22. Massana R, Castresana J, Balagué V, Guillou L, Romari K, Groisillier A, Valentin K, Pedrós-Alió C. 2004. Phylogenetic and ecological analysis of novel marine stramenopiles. *Appl Environ Microbiol* 70:3528–3534. <https://doi.org/10.1128/AEM.70.6.3528-3534.2004>.
 23. de Vargas C, Audic S, Henry N, Decelle J, Mahe F, Logares R, Lara E, Berney C, Le Bescot N, Probert I, Carmichael M, Poulain J, Romac S, Colin S, Aury J-M, Bittner L, Chaffron S, Dunthorn M, Engelen S, Flegontova O, Guidi L, Horak A, Jaillon O, Lima-Mendez G, Luke J, Malviya S, Morard R, Mulot M, Scalco E, Siano R, Vincent F, Zingone A, Dimier C, Picheral M, Searson S, Kandels-Lewis S, Acinas SG, Bork P, Bowler C, Gorsky G, Grimsley N, Hingamp P, Iudicone D, Not F, Ogata H, Pesant S, Raes J, Sieracki ME, Speich S, Stemann L, Tara Oceans Coordinators, et al. 2015. Eukaryotic plankton diversity in the sunlit ocean. *Science* 348:1261605. <https://doi.org/10.1126/science.1261605>.
 24. Mahé F, de Vargas C, Bass D, Czech L, Stamatakis A, Lara E, Singer D, Mayor J, Bunge J, Sernaker S, Siemensemeyer T, Trautmann I, Romac S, Berney C, Kozlov A, Mitchell EAD, Seppey CVW, Egge E, Lentendu G, Wirth R, Trueba G, Dunthorn M. 2017. Parasites dominate hyperdiverse soil protist communities in Neotropical rainforests. *Nat Ecol Evol* 1:e0091. <https://doi.org/10.1038/s41559-017-0091>.
 25. Giner CR, Balagué V, Krabberød AK, Ferrera I, Reñé A, Garcés E, Gasol JM, Logares R, Massana R. 2019. Quantifying long-term recurrence in planktonic microbial eukaryotes. *Mol Ecol* 28:923–935. <https://doi.org/10.1111/mec.14929>.
 26. Logares R, Deuschmann IM, Junger PC, Giner CR, Krabberød AK, Schmidt TSB, Rubinat-Ripoll L, Mestre M, Salazar G, Ruiz-González C, Sebastián M, de Vargas C, Acinas SG, Duarte CM, Gasol JM, Massana R. 2020. Disentangling the mechanisms shaping the surface ocean microbiota. *Microbiome* 8:55. <https://doi.org/10.1186/s40168-020-00827-8>.
 27. Vass M, Székely AJ, Lindström ES, Langenheder S. 2020. Using null models to compare bacterial and microeukaryotic metacommunity assembly under shifting environmental conditions. *Sci Rep* 10:2455. <https://doi.org/10.1038/s41598-020-59182-1>.
 28. Oliverio AM, Geisen S, Delgado-Baquerizo M, Maestre FT, Turner BL, Fierer N. 2020. The global-scale distributions of soil protists and their contributions to belowground systems. *Sci Adv* 6:eaa8787. <https://doi.org/10.1126/sciadv.aax8787>.
 29. Azam F, Fenichel T, Field J, Gray J, Meyer-Reil L, Thingstad F. 1983. The ecological role of water-column microbes in the sea. *Mar Ecol Prog Ser* 10:257–263. <https://doi.org/10.3354/meps010257>.
 30. Sherr EB, Sherr BF. 2002. Significance of predation by protists in aquatic microbial food webs. *Antonie Van Leeuwenhoek* 81:293–308. <https://doi.org/10.1023/a:1020591307260>.
 31. Lima-Mendez G, Faust K, Henry N, Decelle J, Colin S, Carcillo F, Chaffron S, Ignacio-Espinosa JC, Roux S, Vincent F, Bittner L, Darzi Y, Wang J, Audic S, Berline L, Bontempi G, Cabello AM, Coppola L, Cornejo-Castillo FM, d'Ovidio F, De Meester L, Ferrera I, Garet-Delmas M-J, Guidi L, Lara E, Pesant S, Royo-Llonch M, Salazar G, Sánchez P, Sebastian M, Souffreau C, Dimier C, Picheral M, Searson S, Kandels-Lewis S, Gorsky G, Not F, Ogata H, Speich S, Stemann L, Weissenbach J, Wincker P, Acinas SG, Sunagawa S, Bork P, Sullivan MB, Karsenti E, Bowler C, de Vargas C, Raes J, Tara Oceans Coordinators. 2015. Determinants of community structure in the global plankton interactome. *Science* 348:1262073. <https://doi.org/10.1126/science.1262073>.
 32. Bjorbækmo MFM, Evenstad A, Røsaeg LL, Krabberød AK, Logares R. 2020. The planktonic protist interactome: where do we stand after a century of research? *ISME J* 14:544–559. <https://doi.org/10.1038/s41396-019-0542-5>.
 33. Piwosz K, Mukherjee I, Salcher MM, Grujić V, Šimek K. 2021. CARD-FISH in the sequencing era: opening a new universe of protistan ecology. *Front Microbiol* 12:640066. <https://doi.org/10.3389/fmicb.2021.640066>.
 34. Gast RJ, Sanders RW, Caron DA. 2009. Ecological strategies of protists and their symbiotic relationships with prokaryotic microbes. *Trends Microbiol* 17:563–569. <https://doi.org/10.1016/j.tim.2009.09.001>.
 35. Monteil CL, Vallenet D, Menguy N, Benzerara K, Barbe V, Fouteau S, Craud C, Floriani M, Viollier E, Adryanczyk G, Leonhardt N, Faivre D, Pignol D, López-García P, Weld RJ, Lefevre CT. 2019. Ectosymbiotic bacteria at the origin of magnetoreception in a marine protist. *Nat Microbiol* 4:1088–1095. <https://doi.org/10.1038/s41564-019-0432-7>.
 36. Husnik F, Tashyreva D, Boscaro V, George EE, Lukeš J, Keeling PJ. 2021. Bacterial and archaeal symbioses with protists. *Curr Biol* 31:R862–R877. <https://doi.org/10.1016/j.cub.2021.05.049>.

37. Coy SR, Gann ER, Pound HL, Short SM, Wilhelm SW. 2018. Viruses of eukaryotic algae: diversity, methods for detection, and future directions. *Viruses* 10:487. <https://doi.org/10.3390/v10090487>.
38. Ramond P, Sourisseau M, Simon N, Romac S, Schmitt S, Rigaut-Jalabert F, Henry N, de Vargas C, Siano R. 2019. Coupling between taxonomic and functional diversity in protistan coastal communities. *Environ Microbiol* 21:730–749. <https://doi.org/10.1111/1462-2920.14537>.
39. Payne RJ. 2013. Seven reasons why protists make useful bioindicators. *Acta Protozool* 52:105–113.
40. Sagova-Mareckova M, Boenigk J, Bouchez A, Cermakova K, Chonova T, Cordier T, Eisendle U, Elerseck T, Fazi S, Fleituch T, Frühe L, Gajdosova M, Graupner N, Haegerbaeumer A, Kelly A-M, Kopecky J, Leese F, Nöges P, Orlic S, Panksep K, Pawlowski J, Petrussek A, Piggott JJ, Rusch JC, Salis R, Schenk J, Simek K, Stovicek A, Strand DA, Vasquez MI, Vrålstad T, Zlatkovic S, Zupancic M, Stoeck T. 2021. Expanding ecological assessment by integrating microorganisms into routine freshwater biomonitoring. *Water Res* 191:116767. <https://doi.org/10.1016/j.watres.2020.116767>.
41. Filker S, Sommaruga R, Vila I, Stoeck T. 2016. Microbial eukaryote plankton communities of high-mountain lakes from three continents exhibit strong biogeographic patterns. *Mol Ecol* 25:2286–2301. <https://doi.org/10.1111/mec.13633>.
42. Khomich M, Kausered H, Logares R, Rasconi S, Andersen T. 2016. Planktonic protistan communities in lakes along a large-scale environmental gradient. *FEMS Microbiol Ecol* 93:fw231. <https://doi.org/10.1093/femsec/fw231>.
43. Boenigk J, Wodniok S, Bock C, Beisser D, Hempel C, Grossmann L, Lange A, Jensen M. 2018. Geographic distance and mountain ranges structure freshwater protist communities on a European scale. *Metabarcoding Metagenomics* 2:e21519. <https://doi.org/10.3897/mbmg.2.21519>.
44. Bock C, Jensen M, Forster D, Marks S, Nuy J, Psenner R, Beisser D, Boenigk J. 2020. Factors shaping community patterns of protists and bacteria on a European scale. *Environ Microbiol* 22:2243–2260. <https://doi.org/10.1111/1462-2920.14992>.
45. Ptasnik R, Lepistö L, Willén E, Brettum P, Andersen T, Rekolainen S, Lyche Solheim A, Carvalho L. 2008. Quantitative responses of lake phytoplankton to eutrophication in Northern Europe. *Aquat Ecol* 42:227–236. <https://doi.org/10.1007/s10452-008-9181-z>.
46. Stomp M, Huisman J, Mittelbach GG, Litchman E, Klausmeier CA. 2011. Large-scale biodiversity patterns in freshwater phytoplankton. *Ecology* 92:2096–2107. <https://doi.org/10.1890/10-1023.1>.
47. Maileht K, Nöges T, Nöges P, Ott I, Mischke U, Carvalho L, Dudley B. 2013. Water colour, phosphorus and alkalinity are the major determinants of the dominant phytoplankton species in European lakes. *Hydrobiologia* 704:115–126. <https://doi.org/10.1007/s10750-012-1348-x>.
48. Keeling PJ, del Campo J. 2017. Marine protists are not just big bacteria. *Curr Biol* 27:R541–R549. <https://doi.org/10.1016/j.cub.2017.03.075>.
49. Stoecker DK, Hansen PJ, Caron DA, Mitra A. 2017. Mixotrophy in the marine plankton. *Annu Rev Mar Sci* 9:311–335. <https://doi.org/10.1146/annurev-marine-010816-060617>.
50. Berninger U-G, Caron DA, Sanders RW. 1992. Mixotrophic algae in three ice-covered lakes of the Pocono Mountains, U.S.A. *Freshwater Biol* 28:263–272. <https://doi.org/10.1111/j.1365-2427.1992.tb00583.x>.
51. Hitchman RB, Jones HLJ. 2000. The role of mixotrophic protists in the population dynamics of the microbial food web in a small artificial pond. *Freshw Biol* 43:231–241. <https://doi.org/10.1046/j.1365-2427.2000.00541.x>.
52. Huot Y, Brown CA, Potvin G, Antoniadis D, Baulch HM, Beisner BE, Bélanger S, Brazeau S, Cabana H, Cardille JA, del Giorgio PA, Gregory-Eaves I, Fortin M-J, Lang AS, Laurion I, Maranger R, Prairie YT, Rusak JA, Segura PA, Siron R, Smol JP, Vinebrooke RD, Walsh DA. 2019. The NSERC Canadian Lake Pulse Network: a national assessment of lake health providing science for water management in a changing climate. *Sci Total Environ* 695:133668. <https://doi.org/10.1016/j.scitotenv.2019.133668>.
53. Minns CK, Moore JE, Shuter BJ, Mandrak NE. 2008. A preliminary national analysis of some key characteristics of Canadian lakes. *Can J Fish Aquat Sci* 65:1763–1778. <https://doi.org/10.1139/F08-110>.
54. Giner CR, Pernice MC, Balagué V, Duarte CM, Gasol JM, Logares R, Massana R. 2020. Marked changes in diversity and relative activity of picoeukaryotes with depth in the world ocean. *ISME J* 14:437–449. <https://doi.org/10.1038/s41396-019-0506-9>.
55. Ortiz-Álvarez R, Triadó-Margarit X, Camarero L, Casamayor EO, Catalan J. 2018. High planktonic diversity in mountain lakes contains similar contributions of autotrophic, heterotrophic and parasitic eukaryotic life forms. *Sci Rep* 8:4457. <https://doi.org/10.1038/s41598-018-22835-3>.
56. Massana R, Gobet A, Audic S, Bass D, Bittner L, Boutte C, Chambouvet A, Christen R, Claverie J-M, Decelle J, Dolan JR, Dunthorn M, Edvardsen B, Forn I, Forster D, Guillou L, Jaillon O, Kooistra WHCF, Logares R, Mahé F, Not F, Ogata H, Pawlowski J, Pernice MC, Probert I, Romac S, Richards T, Santini S, Shalchian-Tabrizi K, Siano R, Simon N, Stoeck T, Vaultot D, Zingone A, de Vargas C. 2015. Marine protist diversity in European coastal waters and sediments as revealed by high-throughput sequencing. *Environ Microbiol* 17:4035–4049. <https://doi.org/10.1111/1462-2920.12955>.
57. Parfrey LW, Walters WA, Lauber CL, Clemente JC, Berg-Lyons D, Teiling C, Kodira C, Mohiuddin M, Brunelle J, Driscoll M, Fierer N, Gilbert JA, Knight R. 2014. Communities of microbial eukaryotes in the mammalian gut within the context of environmental eukaryotic diversity. *Front Microbiol* 5:298. <https://doi.org/10.3389/fmicb.2014.00298>.
58. Debros D, Domaizon I, Humbert J-FF, Jardillier L, Lepère C, Oudart A, Taïb N, Lepère C, Oudart A, Taïb N. 2017. Overview of freshwater microbial eukaryotes diversity: a first analysis of publicly available metabarcoding data. *FEMS Microbiol Ecol* 93:fx023. <https://doi.org/10.1093/femsec/fix023>.
59. Grujčić V, Nuy JK, Salcher MM, Shabarova T, Kasalický V, Boenigk J, Jensen M, Simek K. 2018. Cryptophyta as major bacterivores in freshwater summer plankton. *ISME J* 12:1668–1681. <https://doi.org/10.1038/s41396-018-0057-5>.
60. Ibarbalz FM, Henry N, Brandão MC, Martini S, Busseni G, Byrne H, Coelho LP, Endo H, Gasol JM, Gregory AC, Mahé F, Rigonato J, Royo-Llonch M, Salazar G, Sanz-Sáez I, Scalco E, Soviadan D, Zayed AA, Zingone A, Labadie K, Ferland J, Marec C, Kandels S, Picheral M, Dimier C, Poulain J, Pisarev S, Carmichael M, Pesant S, Acinas SG, Babin M, Bork P, Boss E, Bowler C, Cochrane G, de Vargas C, Follows M, Gorsky G, Grimsley N, Guidi L, Hingamp P, Iudicone D, Jaillon O, Karp-Boss L, Karsenti E, Not F, Ogata H, Poulton N, Raes J, Sardet C, Speich S, Stemmann L, Tara Oceans Coordinators, et al. 2019. Global trends in marine plankton diversity across kingdoms of life. *Cell* 179:1084–1097. <https://doi.org/10.1016/j.cell.2019.10.008>.
61. Martin SL, Soranno PA. 2006. Lake landscape position: relationships to hydrologic connectivity and landscape features. *Limnol Oceanogr* 51:801–814. <https://doi.org/10.4319/lo.2006.51.2.0801>.
62. Paquette C, Gregory-Eaves I, Beisner BE, Heino J. 2021. Multi-scale biodiversity analyses identify the importance of continental watersheds in shaping lake zooplankton biogeography. *J Biogeogr* 48:2298–2311. <https://doi.org/10.1111/jbi.14153>.
63. Jaworski GHM, Talling JF, Heaney SI. 2003. Potassium dependence and phytoplankton ecology: an experimental study. *Freshw Biol* 48:833–840. <https://doi.org/10.1046/j.1365-2427.2003.01051.x>.
64. Wilken S, Choi CJ, Worden AZ. 2020. Contrasting mixotrophic lifestyles reveal different ecological niches in two closely related marine protists. *J Phycol* 56:52–67. <https://doi.org/10.1111/jpy.12920>.
65. Katechakis A, Stibor H. 2006. The mixotroph *Ochromonas tuberculata* may invade and suppress specialist phago- and phototroph plankton communities depending on nutrient conditions. *Oecologia* 148:692–701. <https://doi.org/10.1007/s00442-006-0413-4>.
66. Saad JF, Unrein F, Tribelli PM, López N, Izaguirre I. 2016. Influence of lake trophic conditions on the dominant mixotrophic algal assemblages. *J Plankton Res* 38:818–829. <https://doi.org/10.1093/plankt/fbw029>.
67. Wilken S, Soares M, Urrutia-Cordero P, Ratcovich J, Ekvall MK, Van Donk E, Hansson L-A. 2018. Primary producers or consumers? Increasing phytoplankton bacterivory along a gradient of lake warming and browning. *Limnol Oceanogr* 63:S142–S155. <https://doi.org/10.1002/lno.10728>.
68. Fischer R, Giebel H-A, Hillebrand H, Ptasnik R. 2017. Importance of mixotrophic bacterivory can be predicted by light and loss rates. *Oikos* 126:713–722. <https://doi.org/10.1111/oik.03539>.
69. Hansson TH, Grossart H, Giorgio PA, St-Gelais NF, Beisner BE. 2019. Environmental drivers of mixotrophs in boreal lakes. *Limnol Oceanogr* 64:1688–1705. <https://doi.org/10.1002/lno.11144>.
70. Schindler DW, Armstrong FAJ, Holmgren SK, Brunskill GJ. 1971. Eutrophication of Lake 227, Experimental Lakes Area, northwestern Ontario, by addition of phosphate and nitrate. *J Fish Res Bd Can* 28:1763–1782. <https://doi.org/10.1139/f71-261>.
71. Liang Z, Soranno PA, Wagner T. 2020. The role of phosphorus and nitrogen on chlorophyll *a*: evidence from hundreds of lakes. *Water Res* 185:116236. <https://doi.org/10.1016/j.watres.2020.116236>.
72. MacKeigan PW, Garner RE, Monchamp M-É, Walsh DA, Onana VE, Kraemer SA, Pick FR, Beisner BE, Agbeti MD, da Costa NB, Shapiro BJ, Gregory-Eaves I. 2022. Comparing microscopy and DNA metabarcoding

- techniques for identifying cyanobacteria assemblages across hundreds of lakes. *Harmful Algae* 113:102187. <https://doi.org/10.1016/j.hal.2022.102187>.
73. Maloufi S, Catherine A, Mouillot D, Louvard C, Couté A, Bernard C, Troussellier M. 2016. Environmental heterogeneity among lakes promotes hyper β -diversity across phytoplankton communities. *Freshw Biol* 61:633–645. <https://doi.org/10.1111/fwb.12731>.
 74. Monchamp M-E, Spaak P, Domaizon I, Dubois N, Bouffard D, Pomati F. 2018. Homogenization of lake cyanobacterial communities over a century of climate change and eutrophication. *Nat Ecol Evol* 2:317–324. <https://doi.org/10.1038/s41559-017-0407-0>.
 75. Aguilar P, Sommaruga R. 2020. The balance between deterministic and stochastic processes in structuring lake bacterioplankton community over time. *Mol Ecol* 29:3117–3130. <https://doi.org/10.1111/mec.15538>.
 76. Baud A, Jenny J-P, Francus P, Gregory-Eaves I. 2021. Global acceleration of lake sediment accumulation rates associated with recent human population growth and land-use changes. *J Paleolimnol* 66:453–467. <https://doi.org/10.1007/s10933-021-00217-6>.
 77. Mitchell P, Prepas EE (ed). 1990. Atlas of Alberta lakes. University of Alberta Press, Edmonton, Canada.
 78. Dornelas M, Gotelli NJ, McGill B, Shimadzu H, Moyes F, Sievers C, Magurran AE. 2014. Assemblage time series reveal biodiversity change but not systematic loss. *Science* 344:296–299. <https://doi.org/10.1126/science.1248484>.
 79. Dillon PJ, Jeffries DS, Snyder W, Reid R, Yan ND, Evans D, Moss J, Scheider WA. 1978. Acidic precipitation in south-central Ontario: recent observations. *J Fish Res Bd Can* 35:809–815. <https://doi.org/10.1139/f78-130>.
 80. Delgado A, Gómez JA. 2016. The soil. Physical, chemical and biological properties, p 15–26. *In* Villalobos FJ, Fereres E (ed), *Principles of agronomy for sustainable agriculture*. Springer International Publishing, Cham, Switzerland.
 81. Fierer N. 2017. Embracing the unknown: disentangling the complexities of the soil microbiome. *Nat Rev Microbiol* 15:579–590. <https://doi.org/10.1038/nrmicro.2017.87>.
 82. Nearing MA, Pruski FF, O'Neal MR. 2004. Expected climate change impacts on soil erosion rates: a review. *J Soil Water Conserv* 59:43–50.
 83. Jeppesen E, Kronvang B, Meerhoff M, Søndergaard M, Hansen KM, Andersen HE, Lauridsen TL, Liboriussen L, Beklioglu M, Özen A, Olesen JE. 2009. Climate change effects on runoff, catchment phosphorus loading and lake ecological state, and potential adaptations. *J Environ Qual* 38:1930–1941. <https://doi.org/10.2134/jeq2008.0113>.
 84. Foley JA, DeFries R, Asner GP, Barford C, Bonan G, Carpenter SR, Chapin FS, Coe MT, Daily GC, Gibbs HK, Helkowski JH, Holloway T, Howard EA, Kucharik CJ, Monfreda C, Patz JA, Prentice IC, Ramankutty N, Snyder PK. 2005. Global consequences of land use. *Science* 309:570–574. <https://doi.org/10.1126/science.1111772>.
 85. Barak P, Jobe BO, Krueger AR, Peterson LA, Laird DA. 1997. Effects of long-term soil acidification due to nitrogen fertilizer inputs in Wisconsin. *Plant Soil* 197:61–69. <https://doi.org/10.1023/A:1004297607070>.
 86. Ontl TA, Schulte LA. 2012. Soil carbon storage. *Nat Educ Knowl* 3:35.
 87. Kaushal SS, Mayer PM, Vidon PB, Smith RM, Pennino MJ, Newcomer TA, Duan S, Welty C, Belt KT. 2014. Land use and climate variability amplify carbon, nutrient, and contaminant pulses: a review with management implications. *J Am Water Resour Assoc* 50:585–614. <https://doi.org/10.1111/jawr.12204>.
 88. Oliva A, Garner RE, Walsh D, Huot Y. 2022. The occurrence of potentially pathogenic fungi and protists in Canadian lakes predicted using geomatics, *in situ* and satellite-derived variables: towards a tele-epidemiological approach. *Water Res* 209:117935. <https://doi.org/10.1016/j.watres.2021.117935>.
 89. Wiken EB, Gauthier D, Marshall I, Lawton K, Hirvonen H. 1996. A perspective on Canada's ecosystems: an overview of the terrestrial and marine ecozones. *In* Canadian Council on Ecological Areas occasional papers. Canadian Council on Ecological Areas, Ottawa, Canada.
 90. Varin MP, Beaulieu ML, Huot Y. 2021. NSERC Canadian Lake Pulse Network field manual 2017 – 2018 – 2019 surveys. NSERC Canadian Lake Pulse Network, Ottawa, Canada.
 91. Gast RJ, Dennett MR, Caron DA. 2004. Characterization of protistan assemblages in the Ross Sea, Antarctica, by denaturing gradient gel electrophoresis. *Appl Environ Microbiol* 70:2028–2037. <https://doi.org/10.1128/AEM.70.4.2028-2037.2004>.
 92. Van de Peer Y, De Rijk P, Wuyts J, Winkelmans T, De Wachter R. 2000. The European Small Subunit Ribosomal RNA database. *Nucleic Acids Res* 28:175–176. <https://doi.org/10.1093/nar/28.1.175>.
 93. Capo E, Debroas D, Arnaud F, Guillemot T, Bichet V, Millet L, Gauthier E, Massa C, Davelle AL, Pignol C, Lejzerowicz F, Domaizon I. 2016. Long-term dynamics in microbial eukaryotes communities: a palaeolimnological view based on sedimentary DNA. *Mol Ecol* 25:5925–5943. <https://doi.org/10.1111/mec.13893>.
 94. Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *Embnet J* 17:10–12. <https://doi.org/10.14806/ej.17.1.200>.
 95. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. 2016. DADA2: high-resolution sample inference from Illumina amplicon data. *Nat Methods* 13:581–583. <https://doi.org/10.1038/nmeth.3869>.
 96. Guillou L, Bachar D, Audic S, Bass D, Berney C, Bittner L, Boutte C, Burgaud G, de Vargas C, Decelle J, del Campo J, Dolan JR, Dunthorn M, Edvardsen B, Holzmann M, Kooistra WHCF, Lara E, Le Bescot N, Logares R, Mahé F, Massana R, Montresor M, Morard R, Not F, Pawlowski J, Probert I, Sauvadet A-L, Siano R, Stoeck T, Vaulot D, Zimmermann P, Christen R. 2013. The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy. *Nucleic Acids Res* 41:D597–D604. <https://doi.org/10.1093/nar/gks1160>.
 97. Katoh K, Rozewicki J, Yamada KD. 2019. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief Bioinform* 20:1160–1166. <https://doi.org/10.1093/bib/bbx108>.
 98. Adl SM, Bass D, Lane CE, Lukeš J, Schoch CL, Smirnov A, Agatha S, Berney C, Brown MW, Burki F, Cárdenas P, Čepička I, Chistyakova L, Campo J, Dunthorn M, Edvardsen B, Eglit Y, Guillou L, Hampl V, Heiss AA, Hoppenrath M, James TY, Karnkowska A, Karpov S, Kim E, Kolisko M, Kudryavtsev A, Lahr DJG, Lara E, Le Gall L, Lynn DH, Mann DG, Massana R, Mitchell EAD, Morrow C, Park JS, Pawlowski JW, Powell MJ, Richter DJ, Rueckert S, Shadwick L, Shimano S, Spiegel FW, Torruella G, Yousef N, Zlatogursky V, Zhang Q. 2019. Revisions to the classification, nomenclature, and diversity of eukaryotes. *J Eukaryot Microbiol* 66:4–119. <https://doi.org/10.1111/jeu.12691>.
 99. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* 10:421. <https://doi.org/10.1186/1471-2105-10-421>.
 100. Canadian Council of Ministers of the Environment. 2004. Canadian water quality guidelines for the protection of aquatic life: phosphorus: Canadian guidance framework for the management of freshwater systems. *In* Canadian environmental quality guidelines. Canadian Council of Ministers of the Environment, Winnipeg, Canada.
 101. Muñoz SJ. 2019. ERA5-Land hourly data from 1981 to present. Copernicus Climate Change Service Climate Data Store. <https://doi.org/10.24381/cds.e2161bac>.
 102. Hengl T, Mendes de Jesus J, Heuvelink GBM, Ruiperez Gonzalez M, Kilibarda M, Blagotić A, Shangguan W, Wright MN, Geng X, Bauer-Marschallinger B, Guevara MA, Vargas R, MacMillan RA, Batjes NH, Leenaars JGB, Ribeiro E, Wheeler I, Mantel S, Kempen B. 2017. SoilGrids250m: global gridded soil information based on machine learning. *PLoS One* 12:e0169748. <https://doi.org/10.1371/journal.pone.0169748>.
 103. Becker RA, Wilks AR, Brownrigg R, Minka TP, Deckmyn A. 2018. maps: draw geographical maps. CRAN.
 104. Pruesse E, Peplies J, Glöckner FO. 2012. SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* 28:1823–1829. <https://doi.org/10.1093/bioinformatics/bts252>.
 105. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 41:D590–D596. <https://doi.org/10.1093/nar/gks1219>.
 106. Price MN, Dehal PS, Arkin AP. 2010. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS One* 5:e9490. <https://doi.org/10.1371/journal.pone.0009490>.
 107. Chen J, Bittinger K, Charlson ES, Hoffmann C, Lewis J, Wu GD, Collman RG, Bushman FD, Li H. 2012. Associating microbiome composition with environmental covariates using generalized UniFrac distances. *Bioinformatics* 28:2106–2113. <https://doi.org/10.1093/bioinformatics/bts342>.
 108. Oksanen J, Blanchet FG, Kindt R, Legend P, Minchin PR, Hara RBO, Simpson GI, Soly P, Stevens MHH, Wagner H. 2020. vegan: Community ecology package. CRAN.
 109. Kembel SW, Cowan PD, Helmus MR, Cornwell WK, Morlon H, Ackerly DD, Blomberg SP, Webb CO. 2010. Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* 26:1463–1464. <https://doi.org/10.1093/bioinformatics/btq166>.

110. Dray S, Bauman D, Blanchet G, Borcard D, Clappe S, Guenard G, Jombart T, Larocque G, Legendre P, Madi N, Wagner HH. 2021. *adespatial*: multivariate multiscale spatial analysis. CRAN.
111. Lê S, Josse J, Husson F. 2008. *FactoMineR*: an R package for multivariate analysis. *J Stat Softw* 25:253–258. <https://doi.org/10.18637/jss.v025.i01>.
112. Ferrier S, Manion G, Elith J, Richardson K. 2007. Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Divers Distrib* 13:252–264. <https://doi.org/10.1111/j.1472-4642.2007.00341.x>.
113. Rosauer DF, Ferrier S, Williams KJ, Manion G, Keogh JS, Laffan SW. 2014. Phylogenetic generalised dissimilarity modelling: a new approach to analysing and predicting spatial turnover in the phylogenetic composition of communities. *Ecography* 37:21–32. <https://doi.org/10.1111/j.1600-0587.2013.00466.x>.
114. Fitzpatrick MC, Mokany K, Manion G, Lisk M, Ferrier S, Nieto-Lugilde D. 2021. *gdm*: generalized dissimilarity modeling. CRAN.
115. R Core Team. 2021. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.