

Large-scale geographic patterns of diversity and community structure of pelagic crustacean zooplankton in Canadian lakes

Bernadette Pinel-Alloul^{1,2*}, Adrien André^{1,4}, Pierre Legendre^{1,2}, Jeffrey Cardille^{1,3}, Kasimierz Patalas⁵, Alex Salki⁵

¹, GRIL, Groupe de recherche Interuniversitaire en Limnologie et en Environnement Aquatique, Université de Montréal, C.P. 6128, Succ. Centre-ville, Montréal (QC), Canada, H3C 3J7.

², Département de sciences biologiques, Université de Montréal, C.P. 6128, Succ. Centre-ville, Montréal (QC), Canada, H3C 3J7. E-mails: bernadette.pinel-alloul@umontreal.ca; pierre.legendre@umontreal.ca

³, Département de géographie, Université de Montréal, C.P. 6128, Succ. Centre-ville, Montréal (QC), Canada, H3C 3J7. E-mail : jeffrey.cardille@umontreal.ca

⁴, Département des sciences et génie de l'environnement, Faculté des sciences, Université de Liège, 15 Allée du 6 août, B-4000, Liège, Belgique. E-mail: adrienandre@hotmail.com

⁵, retired, formerly of Freshwater Institute, Department of Fisheries and Oceans, 501 University Crescent, Winnipeg (MA), Canada R3T 2N6.

*Correspondence : Bernadette Pinel-Alloul, GRIL, Groupe de recherche Interuniversitaire en Limnologie et en Environnement Aquatique, Département de sciences biologiques, Université de Montréal, C.P. 6128, Succ. Centre-ville, Montréal, (QC), Canada, H3C 3J.

E-mail: bernadette.pinel-alloul@umontreal.ca

Keywords

Large-scale patterns, pelagic crustacean zooplankton, diversity, community structure, lakes, ecoprovinces, continental scale, Canada

Short running header

Biodiversity patterns of crustacean zooplankton in Canada

ABSTRACT

Aim We tested the energy and metabolic theories for explaining diversity patterns of crustacean zooplankton in Canadian lakes, and evaluated the influence of regional and local environments on community structure.

Location The 1665 studied lakes are distributed across Canada in 47 ecoprovinces.

Methods Our database included the occurrence of 83 pelagic crustacean species. The regional species richness in each ecoprovince was estimated using the averaged local species richness per lake and the first-order jackknife diversity index. Using a principal component plot and forward selection in a multiple regression we identified the most important predictors of regional species richness estimates. We tested the predictions of the species richness-energy hypothesis using climatic variables at regional scale, and of the metabolic theory using the inverse of temperature. To evaluate the influence of regional and local environmental drivers, we carried out a redundancy analysis between crustacean species occurrences and regional climate and lake environmental factors on a subset of 458 lakes.

Results Estimates of pelagic crustacean species richness in Canadian ecoprovinces varied from 3 to 10 species per lake (averaged local species richness) or 8 to 52 species per ecoprovince (Jackknife diversity index). Our study fully supports the species richness-energy hypothesis and partially the metabolic theory. Mean daily global solar radiation was the most important regional predictor, explaining 51% of the variation in the regional species richness among ecoprovinces. Together, regional climate and local lake environment accounted for 31% of the total variation in community structure. Regional-scale energy variables accounted for 24% of the total explained variation, whereas local-scale lake conditions had less influence (2%).

51 **Main conclusions** The richness-energy theory explains diversity patterns of freshwater
52 crustacean zooplankton in Canadian ecoprovinces. Solar radiation is the best predictor
53 explaining regional species richness in ecoprovinces and community structure of pelagic
54 crustaceans in Canadian lakes.

55

INTRODUCTION

Since the eighteenth century, species distribution patterns have been of primary interest in the fields of biogeography, evolutionary biology, and ecology. Exploring how and why species are currently distributed in their geographical range are two fundamental issues in ecology and biogeography (Gaston, 2000; Lomolino *et al.*, 2010). Perhaps the most interesting property of species diversity is its organization through space or beta diversity (Whittaker *et al.*, 2001; Legendre *et al.*, 2005). Beta diversity is a key concept for understanding the function of ecosystems, for the conservation of biodiversity, and for ecosystem management (Legendre *et al.*, 2005). It can tell us which species are habitat generalists or specialists, which ones present similar or different competitive ability, and how species composition varies between sites or regions in response to environmental changes (Tuomisto & Ruokolainen, 2006). Thus, understanding the mechanisms controlling spatial patterns in species richness and community structure will help efforts to conserve biodiversity and functions of ecosystems in the face of climate change and increased human disturbances.

The multiple meanings of beta diversity have been discussed recently (Anderson *et al.*, 2011; Legendre & Legendre, 2012). Beta diversity studies can focus on two aspects of community structure. The first one is *turnover*, or the change in community composition between adjacent sampling units, explored by sampling along a spatial, temporal, or environmental gradient. The second is a non-directional approach to the study of community variation through space; it does not refer to any specific gradient but centres on the *variation in community composition among the study units*. In the present paper, we focused on the second approach, where spatial variation in species richness and community structure among lakes at the scale of Canadian ecoprovinces (beta diversity) will be analysed (in the statistical sense) through linear models involving regional and local environmental variables.

Biologists have studied large-scale diversity patterns in macroorganisms for centuries, leading to many insights on the biogeography of terrestrial species (Allen *et al.*, 2002, Hawkins *et al.*, 2003a, 2007a). In recent decades, large-scale patterns of beta diversity have been widely assessed for trees and plants (Jiménez *et al.*, 2009; Blach-Overgaard *et al.*, 2010), mammals and birds (Badgley & Fox, 2000; Hawkins *et al.*, 2003b; Melo *et al.*, 2009), and insects (Kerr & Currie, 1999). Many hypotheses have been proposed to explain large-scale diversity patterns of terrestrial species. For decades, the *richness-energy hypothesis* has provided a strong explanation related to productivity-diversity relationships (Chase & Ryberg, 2004). Energy (temperature or solar radiation) or water-energy (precipitation or evapotranspiration) variables typically drive diversity patterns of terrestrial organisms (Hawkins *et al.*, 2003a; Currie *et al.*, 2004). Areas with higher energy and water inputs are able to support higher species richness because productivity is strongly affected by the quantity of energy and water coming into terrestrial ecosystems. The *spatial heterogeneity hypothesis* can provide a supplemental explanation for plants for which diversity patterns also reflect heterogeneity in habitat and topography (Jimenez *et al.*, 2009). Another important theory is the physiologically-oriented *metabolic theory* (Allen *et al.*, 2002; Brown *et al.*, 2004; Hawkins *et al.*, 2007a, 2007b). According to this theory, large-scale diversity patterns result from the dependence of the metabolism of terrestrial ectotherms to the ambient solar radiation that controls their body temperature. Because species diversity of ectotherms is strongly influenced by metabolic processes, much of the variation in species diversity is due to air temperature, and higher species richness is expected in warmer areas. This hypothesis predicts that log-transformed species richness is linearly associated to the inverse of annual temperature, and that the slope of the relationship varies between -0.55 and -0.75 . This theory has received support from many studies conducted in terrestrial ecosystems (Algar *et al.*, 2007), though it has also been criticized (Hawkins *et al.*, 2007b).

In contrast to macroorganisms, the field of microbial biogeography is currently immature (Fierer, 2008). There is still debate as to whether microorganisms also exhibit biogeographical patterns, and whether established ecological theory can explain spatial diversity patterns of microbial communities (Fontaneto *et al.*, 2006; Martiny *et al.*, 2006). Microorganisms are expected to show weak geographic variation in diversity compared to macroorganisms because of their small size, high abundance, fast population growth and higher dispersal rates. However, recent reviews showed that spatial diversity patterns do exist for free-living microorganisms in soil and waters (Martiny *et al.*, 2006; Fontaneto, 2011). Strong biogeographic patterns in plankton biodiversity were observed at global scale in Northern America (Vyverman *et al.*, 2007; Stomp *et al.*, 2011) and in Europe (Hessen *et al.*, 2006, 2007; Ptacnik *et al.*, 2010); they were related to environmental gradients in productivity, habitat area and temperature as observed for terrestrial macroorganisms. However, there is no consensus on the importance of the *species-richness theory* for the regulation of diversity patterns of microorganisms in aquatic systems, where the water environment can buffer the effects of temperature and solar radiation. Air temperature and energy-related factors were found the major determinants of large-scale latitudinal pattern of crustacean species richness in Norwegian lakes (Hessen *et al.*, 2007) and of copepods in the Atlantic and Pacific Oceans (Rombouts *et al.*, 2009). However, Stomp *et al.* (2011) did not found a positive effect of water temperature on large-scale diversity patterns of freshwater phytoplankton. Also, the *metabolic theory* is not universally well supported in aquatic systems. It has little predictive power for the metabolism of lacustrine plankton (De Castro & Gaedke, 2008), and it weakly explains zooplankton species richness pattern in Norwegian lakes (Hessen *et al.*, 2007). Support to the *spatial heterogeneity theory* came from studies on zooplankton diversity patterns, which are driven by multiple regional and local processes (*the multiple force hypothesis*) acting differently across spatial scales (Pinel-Alloul, 1995; Pinel-

Alloul *et al.*, 1995; Pinel-Alloul & Ghadouani, 2007). However, it is difficult to disentangle the effects of energy-related variables from that of local environmental variables, because they are all correlated to geographical gradients. A better prediction of the patterns of aquatic biodiversity and its environmental drivers is a fundamental issue for ecologists and may be the most important scientific challenge to face in the twenty-first century (Willig & Bloch, 2006). This is especially important given current concerns about the loss of biodiversity in freshwater ecosystems due to the multiple stressors of climate changes, watershed land use, alteration of nutrient cycles, invasion of exotic species, and overexploitation of halieutic resources. Our ability to detect the effects of those stressors on biodiversity is often low, in part because of poor knowledge of baseline data on biodiversity patterns and their environmental drivers.

In Canada, lakes are a dominant feature of the landscape, but knowledge of biodiversity of freshwater microorganisms and its environmental control is still in its infancy. Our study provides the first comprehensive model relating diversity patterns of pelagic crustaceans across Canada to environmental gradients along a wide range of physiographic, climatic and lake environments. The aims of the study are fourfold: (1) document the spatial pattern of diversity, community structure and species distribution of lake crustacean zooplankton in Canadian ecoprovinces, (2) test the main predictions of the diversity-energy and metabolic ecological theories for large-scale diversity patterns of freshwater microorganisms, (3) determine the environmental drivers of spatial variation in crustacean diversity at regional scale across Canadian ecoprovinces, and (4) evaluate the relative influence of regional climatic features and local lake environmental factors on crustacean community structure.

MATERIALS AND METHODS

Study sites, zooplankton sampling, and database

A long-term sampling program (1961-1991) carried out by the Freshwater Institute of Fisheries and Oceans Canada and an extensive literature review (1891-1990) provided a large database of crustacean species occurrence in the pelagic zone of nearly 2000 lakes across the entire mainland of Canada (42-80° N and 52-139° W) (Patalas *et al.*, 1994). The studied lakes are distributed across the 15 Canadian terrestrial ecozones and in 47 of Canada's 53 ecoprovinces (Fig. 1). Ecozones and ecoprovinces represent high- and intermediate-level divisions of the Canadian land mass according to climatic and vegetation patterns (Marshall & Schut, 1999). They are useful geographic units for general national reporting (<http://atlas.nrcan.gc.ca>) and for placing Canada's ecosystem diversity assessment in an ecologically meaningful context (McMahon *et al.*, 2004).

The number of lakes sampled during the long-term field survey varied widely among ecoprovinces, from 326 lakes in the Southern Boreal Shield to a single lake in the Sverdrup Islands in the Northern Arctic, and the Whale River Lowland in the Taiga Shield (Appendix S1). Zooplankton sampling took place during mid-summer near the centre of each lake; per-lake sampling effort ranged from a single site in small lakes to 30 to 50 sites in large lakes. Zooplankton was collected with a Wisconsin plankton net (25 cm in diameter, mesh size of 53-77 µm) by vertical hauls from the lake bottom to the surface, or from a depth of 50 m in the deepest lakes. Zooplankton was preserved in 4% formalin, and analysed for species identification (see Patalas, 1990; Patalas *et al.*, 1994).

Given the opportunistic nature of the survey and changing technology over the sampling period, we encountered a considerable range of precision and accuracy for lake positional data. The locations of lakes sampled during the earlier decades in the survey were not reported clearly whereas those sampled during the later decades were referenced with relatively clear positional information. Thus, we applied modern techniques to determine the

location of each lake in the data set, as accurately as possible. Using their descriptions in the original 2000-lake data set, the GeoNames database (<http://www.geonames.org>) to convert text descriptions to potential latitude-longitude positions, and Google Earth (<http://earth.google.com>) to verify and choose among proposed locations, we relocated 1665 of the lakes with enough precision to be included in the analysis.

Regional and local environmental data

Regional environmental data on climate and precipitation were collected for the 1951-1980 period from the Ecological Framework of Canada website (<http://ecozones.ca>), and from the Climate Atlas of Canada (Phillips, 1990; Environment Canada, 1986). The selected regional environmental descriptors were the growing season length in days, the number of growing days above 10°C, the effective growing degree days above 5°C, the mean elevation, the total annual precipitation, the mean daily global solar radiation, the mean duration of bright sunshine, the daily air temperatures (mean, minimum, maximum) estimated over the entire year, the mean annual vapour pressure, the annual potential evapotranspiration, and the ecoprovince area (Table 1). Using ArcGIS (ESRI software, 2009), we estimated the values of regional environmental descriptors for each lake by intersecting them with the sampled points in the lake-zooplankton database. We also calculated the area-weighted mean for each regional environmental descriptor in each ecoprovince using the values in each polygon.

Our lake survey covered a wide range of regional climatic conditions across most of the Canadian ecoprovinces (Table 1). In addition to regional descriptors, we used previous reports (Salki & Patalas, unpublished data) to obtain local descriptors of the lake environments (July air temperature, lake area and depth, total dissolved solids, Secchi depth) in 458 lakes among the 1665-lake data set. The lakes of this subset were distributed across Canada and covered the same broad range of physiographic, climatic and limnological conditions as the 1665-lake dataset (Fig. 1, Appendix S5).

Biodiversity patterns, community structure, and species distribution

The survey provided us with a unique database on the occurrence of 83 crustacean species (Patalas *et al.*, 1994) in 1665 lakes distributed across Canada. To document biodiversity patterns, local species richness was estimated as the total number of crustacean species found in each lake, and averaged within each ecoprovince. Because sampling effort varied among ecoprovinces, we also calculated the first-order jackknife estimator of species richness, $\hat{S} = S_{\text{obs}} + r(n-1)/n$, where S_{obs} is the number of species observed in n lakes and r is the number of species present in only one lake, to estimate regional species richness within each ecoprovince, as recommended by Palmer (1990) and Arnott *et al.* (1998). We were able to calculate the two estimates of species richness for 35 of the 47 ecoprovinces. The 12 ecoprovinces, which had five or fewer than five sampled lakes, were excluded (Appendix S1). Community structure was established on information about the presence or absence of each species in each lake. Occurrence of crustacean species across Canada was illustrated by a rank-frequency diagram. We produced distribution maps of each species in ArcGIS using the KML format in ET GeoWizards extension of the Google Earth Web software.

Diversity-energy relationships

To test the species richness-energy theory, we evaluated the relationships between the estimates of species richness (the averaged local species richness and the Jackknife diversity index) and regional environmental descriptors at the scale of the ecoprovinces using data from the 1665 lakes and the 35 selected ecoprovinces. First, a Principal Component Analysis (PCA) was computed to identify and illustrate groups of correlated variables among the standardized regional environmental descriptors. The averaged local species richness and the Jackknife diversity index were added to the plot by passive ordination (Legendre & Legendre, 2012). Forward selection in multiple regression analysis was also used to determine which environmental variables were the best predictors of the species richness estimates in the

35 ecoprovinces. We then computed linear regressions between the species richness estimates and selected energy or water-energy related descriptors (mean daily solar radiation, effective growing degree days above 5°C, mean daily air temperature over the entire year, annual potential evapotranspiration and mean duration of bright sunshine) and tested their significance. Finally, acknowledging the fact that the energy variables were collinear, we used forward selection in linear regression to show which energy-related variables contributed significantly to explain the variation of the species richness estimates.

To test the metabolic theory, we related the log-transformed estimates of species richness to the inverse of annual temperature (in Kelvin) multiplied by the Boltzmann constant, using linear regressions. Then, as suggested by Algar *et al.* (2007), we tested whether this relationship was best fitted by a linear or a curvilinear model using the Akaike Information Criterion (AIC).

Relationships between regional and local environments and community structure

To test the hypothesis of multiple forcing of the crustacean community structure by regional and local environmental drivers (Pinel-Alloul, 1995; Pinel-Alloul *et al.*, 1995), we used the subset of 458 lakes for which local environmental descriptors were available. The regional factors were represented by climatic features of the ecoprovinces whereas the local environmental factors featured morphometric and water quality lake variables. Both local and regional environmental variables were used to construct third-degree polynomial equations, for a total of forty-eight monomials. The quadratic and cubic monomials enabled us to model nonlinear relationships between the environmental descriptors and each species presence-absence variable (Legendre & Legendre, 2012). We performed a redundancy analysis (RDA) of the species occurrence matrix with forward selection of the significant monomials. The species data had previously been transformed using the Hellinger method (Legendre & Gallagher, 2001). We produced a RDA biplot of the crustacean species with the five most

important environmental descriptors retained by forward selection. We then used variation partitioning to assess the relative contributions of the regional and local descriptors (Borcard *et al.*, 1992; Peres-Neto *et al.*, 2006). The statistical analyses were performed using software available in the VEGAN and PACKFOR packages in the R language (R Development Core Team, 2012).

RESULTS

Species richness and distribution patterns

Among the 83 crustacean species recorded in the 1665 lakes, the most diversified groups were the Calanoida (33 species) and the Cladocera (33 species), followed by the Cyclopoida (17 species) (Appendix S2). Only nine species were found in more than 25% of the lakes (Fig. 2). *Bosmina longirostris* was the most common species, found in almost half of the lakes. The other common species found in more than 25% of lakes were: *Holopedium gibberum*, *Diacyclops thomasi*, *Mesocyclops edax*, *Leptodiaptomus minutus*, *Daphnia mendotae*, *Cyclops scutifer scutifer*, *Daphnia longiremis*, and *Diaphanosoma leuchtenbergianum*. Two thirds of the species (54/83) were present in less than 5 % of the sampled lakes.

The total number of crustacean species found in Canadian ecoprovinces varied from 7 to 44, with the median and mean values respectively of 29 and 27 species; the averaged local species richness per lake ranged from 3 to 10 species (median: 5; mean: 6) while the Jackknife index ranged from 8 to 52 species (median: 36; mean: 34) (Table 1). Species-rich lakes were typically located in the ecoprovinces of the Boreal Shield and Plains and in the Hudson-Erie and the Great Lakes-Saint Lawrence Plains, whereas the species-poor lakes were found in the northern and arctic ecoprovinces (Fig. 1). Distribution maps of each crustacean

species across Canada are available in the KML format compatible with Google Earth Web software: (<https://sites.google.com/site/canadianzooplankton/maps/distribution>).

Regional environment and diversity-energy relationships

The first two axes of the PCA ordination accounted for 73% of the observed variance in regional environmental descriptors among the ecoprovinces (Fig. 3). The first axis (61%) showed strong positive correlations with air temperature, the energy- and water-energy related descriptors, and an inverse correlation with latitude, reflecting the latitudinal gradient of decreasing temperature and solar irradiance. The Jackknife diversity index was strongly and positively associated with this axis. Higher regional species richness was observed in southern ecoprovinces located at lower latitude where mean global daily solar variation was the highest. The second axis (12%) represents the longitudinal and altitudinal gradients and was related to the averaged local species richness; it opposes species-poor lakes located at higher altitude in small ecoprovinces in western Canada to species-rich lakes located at low altitude in large ecoprovinces with higher duration of bright sunshine.

Forward selection in multiple regression modelling retained four statistically significant regional descriptors (Table 2). It showed that the single energy variable *Mean daily global solar radiation* accounted alone for 52% of the variation of regional species richness among ecoprovinces (Jackknife index) and that two other energy-related variables (*Effective growing degree days above 5°C*, and *Mean duration of bright sunshine*) also contributed significantly to explain the spatial variation of regional species richness. The ecoprovince area was also included as a significant regional descriptor of crustacean species richness in multiple regression modelling; however, its contribution was weak as suggested by the lack of correlation detected with the PCA ordination. The adjusted R-square (R^2_{adj}) of the combined effect of these four variables on the Jackknife index was 0.72 (Table 2). In contrast, no energy variable was related to the variation of the averaged local species richness

among ecoprovinces which was significantly explained by the latitude, altitude, and area of ecoprovinces ($R^2_{adj} = 0.40$) (Appendix S3-A).

At the ecoprovince level, simple linear regression models computed between the Jackknife diversity index and selected environmental variables indicated significant positive correlations with the energy and water-energy descriptors measured in the ecoprovinces. Higher values of mean daily global solar radiation, annual potential evapotranspiration, effective growing degree days above 5°C, mean annual air temperature, and mean duration of bright sunshine were all significantly associated with higher values of the Jackknife index of regional species richness (Table 3). However, there was no significant correlation with the area of ecoprovinces considered alone. In comparison, simple linear regression models between the averaged local species richness and the selected environmental variables were weak (Appendix S3-B).

The relationship between the natural logarithm of the Jackknife diversity index in each ecoprovince and the inverse of ambient temperature (in Kelvin) corrected by the Boltzmann constant supported the first prediction of the metabolic theory (Fig. 4). Both the linear model ($y = 14.63 - 0.26 \cdot x$; $R^2_{adj} = 0.46$; p-value: $4.06 \cdot 10^{-8}$, AIC = 18.13) and the quadratic model ($y = -165.91 + 8.09 \cdot x - 0.096 \cdot x^2$; $R^2_{adj} = 0.54$, p-value: $1.39 \cdot 10^{-6}$, AIC = 25.85) were highly significant but AIC, whose minimum indicates the best predictive model, showed that the linear model was best. However, the second prediction of the metabolic theory was not supported: the slope of the linear regression was -0.26 , far out of the range of slopes (-0.55 to -0.75) predicted by the metabolic scaling law for species richness. When using the averaged local species richness, the relationships with the inverse of ambient temperature were very weak both for the linear ($R^2_{adj} = 0.10$; p-value: 0.04) and the quadratic models ($R^2_{adj} = 0.13$; p-value: 0.04) (Appendix S4).

Local and regional control of crustacean community structure

Among the regional descriptors of the ecoprovinces and the local lake descriptors measured on the subset of 458 lakes, 32 monomials were retained as significant predictors of the crustacean community structure (Table 4). Those variables accounted together for 31% of the total variation in species assemblages. The most important regional descriptors were the mean daily global solar radiation followed by the mean duration of bright sunshine, mean elevation, mean annual vapour pressure, and mean annual air temperature. Some local descriptors of lake environments such as total dissolved solids, July air temperature, Secchi depth transparency and lake depth were also retained in the RDA model. The first five selected descriptors (4 regional and 1 local) were first-degree monomials, indicating predominance of linear effects. The RDA ordination of the species occurrences by the five most important descriptors ($R^2_{adj} = 0.22$) showed strong associations with species distributions across lakes (Fig. 5). The first canonical axis accounted for 14.3% of the total variation and was mainly driven by higher mean daily global solar radiation and duration of bright sunshine. The second axis was related to gradients of altitude and productivity (via total dissolved solids) and accounted only for 5.4% of the total variation. Western species such as *Leptodiaptomus angustilobus* and *Heterocope septentrionalis* were associated with lakes located at higher altitude, and eastern species such as *Holopedium gibberum*, *Leptodiaptomus minutus* and *Tropocyclops prasinus mexicanus* to lakes located at lower elevation. *Mesocyclops edax*, *Bosmina longirostris* and *Diaphanosoma leuchtenbergianum* were associated with lakes receiving the highest solar radiation, and *Skistodiaptomus oregonensis* and *Diacyclops thomasi* to the highest duration of sunshine. Variation partitioning between significant regional and local environmental variables indicated that 24% of the total variation in community structure was due to regional factors, and only 2% to local factors (Appendix S6).

DISCUSSION

In total, our records indicate that at least 83 crustacean species inhabit the pelagic zone of Canadian lakes. The variation in species richness of crustaceans in lakes across Canadian ecoprovinces is comparable to the range observed in other large-scale and multiple-year surveys in European and American lakes (Arnott *et al.*, 1998; Shurin *et al.*, 2007; Hessen *et al.*, 2006, 2007).

Our study provided the first model of continental-scale distribution patterns of freshwater crustacean species in relation to energy-related gradients in Canadian ecoprovinces. The richness-energy theory is a fundamental process explaining diversity patterns of freshwater zooplankton at the continental scale of Canada. Our results are consistent with monotonically increasing species-energy relationships found in macro-scale studies of fish (Kerr & Currie, 1999) and terrestrial organisms (Hawkins *et al.*, 2003a; Evans *et al.*, 2005). Crustacean species richness per ecoprovince, estimated by the Jackknife diversity index is best predicted by global solar radiation, meaning that in ecoprovinces with higher energy inputs, the lakes support globally more crustacean species. The identification of a pure energy variable (mean daily global solar radiation) as the first predictor of diversity pattern rather than a water-energy (e.g., annual potential evapotranspiration) or a water-only variable (e.g., total annual precipitation) reflects the obvious fact that water availability is not a limiting factor in lake ecosystems. Our results also pointed to the importance of other energy-related variables, mainly the effective growing degree days above 5°C and the mean duration of bright sunshine, because they are closely correlated with the global solar radiation variable. In contrast, averaged local species richness in lakes of Canadian ecoprovinces does not vary with energy-related variables as observed by Hessen *et al.* (2007) at the continental scale of Norway. This means that while solar radiation appears to control regional diversity of freshwater zooplankton at large continental scale, it does not control local species richness

within a lake. At local scale, zooplankton species richness may to be controlled both by abiotic and biotic processes (Pinel-Alloul, 1995; Hessen *et al.*, 2006; Pinel-Alloul & Ghadouani, 2007).

Our study gave partial support for the metabolic ecological theory for explaining diversity patterns of freshwater crustaceans in Canada. We did find a linear significant relationship ($R^2_{adj} = 0.46$) between the log-transformed Jackknife index and the inverse transformation of annual temperature, yet the slope (-0.26) was less steep than predicted by the theory (between -0.55 and -0.75). In comparison, the linear relationship with the averaged local species richness was weak ($R^2_{adj} = 0.10$). In Norwegian lakes, the metabolic theory explained only 21% of the variation in zooplankton local species richness (Hessen *et al.*, 2007), but the slope (-0.78) was in the expected range. Several reasons may account for these differences: (i) our broad-scale study covered a wider air temperature gradient (29° vs 12°), (ii) we used regional species richness per ecoprovince in our Canadian survey whereas local species richness in each lake was used in the Norwegian survey, (iii) the weaker relationships with the local species richness in Norwegian and Canadian lakes suggest that there are other important factors influencing species richness at local scale than at regional scale, (iv) there is still controversy over the value of the slope of the relationship of the metabolic theory, which may vary close to multiples of 0.25. In our large-scale study, we suggested that lake water stratification may buffer the effect of ambient air temperature on metabolism of crustacean zooplankton that can migrate in deeper cold waters. Furthermore, mean daily air temperature over the entire year used as the explanatory variable in our model does not directly correspond to the realized temperature experienced by crustacean species during their life cycle.

Our study of crustaceans in Canadian lakes gave additional support to the spatial heterogeneity theory and the multiple forces hypothesis which stated that abiotic environmental gradients will be the most important drivers of spatial variation in community

structure at large scale (Pinel-Alloul, 1995; Pinel-Alloul *et al.* 1995; Pinel-Alloul & Ghadouani, 2007; Shurin *et al.*, 2000). Indeed, latitudinal gradients in global solar radiation and ambient air temperature were the most important drivers of biodiversity patterns of freshwater crustaceans across Canada. Lakes situated in northern and arctic ecoprovinces experienced almost no light and extremely low temperatures during winter, and with such limiting factors they showed the lowest crustacean species richness. Crustaceans need to develop special physiological or behavioural responses to survive in this kind of environment. In lakes located in southern ecoprovinces, the higher inputs of energy by solar radiation can sustain primary productivity via efficient photosynthesis and carbon fixation of producers. Higher solar radiation can also be responsible for an increased euphotic zone depth. In consequence, a positive bottom-up control of primary producers by solar radiation can promote more abundant consumer populations, and therefore a reduction of the extinction rate which leads to increased species richness (Hessen *et al.*, 2007). Moreover, areas receiving higher energy inputs contain higher abundances of relatively rare resources that are exploited by niche position specialists. This should allow those specialists to maintain larger viable populations, thus increasing species richness. In our study, regional climate was the most important predictor of both crustacean diversity and community structure, while local factors appear to be poor predictors of the community structure. The weak signal of local predictors may probably be caused by the low number of local variables available in our survey, which reduced our power to detect strong effects at local scale. As in other large-scale studies of zooplankton-environment linkages (Pinel-Alloul *et al.*, 1995; Hessen *et al.*, 2007), more than half of the total variance in crustacean community structure remains unexplained.

As with most large-scale surveys, there are caveats in our study which might limit our conclusions. Sampling did not cover the whole summer growing season and the survey extended through more than 90 years. However, we are confident that our data provide an

accurate evaluation of crustacean species richness patterns in Canada: a recent study conducted at spatial and temporal scales on an extensive data set of zooplankton showed that species richness evaluated on a daily basis was linearly related to mean annual species richness (Shurin *et al.*, 2007). Pelagic zooplankton is a key component of food webs in large and deep lakes. However, because littoral crustaceans can make a major contribution to zooplankton species richness in small lakes (Walseng *et al.*, 2006), it is important to note that our study does not assess species richness of crustaceans in littoral habitats of Canadian lakes. Our modelling effort for assessing regional and local control of crustacean assemblages in the subset of 458 lakes was also limited by the unavailability of local descriptors beyond a few morphometric and trophic variables. Other unmeasured descriptors such as nutrients, chlorophyll biomass and fish predation may have important effects on crustacean community structure at local scale (Hessen *et al.*, 2006, 2007).

In the coming years, scientists may well discover that zooplankton species distribution and community structure are changing with climate warming and land use, perhaps at a variety of scales from local lakes to large ecoprovinces. Efforts to compile and modernize existing but disparate records will serve a key role in understanding changes inside lakes. Meanwhile, modern techniques allow the assembly and sharing of such databases for future researchers. In this setting, we created an extensive and modern spatial lake database, addressed it with models built from wide-area environmental data, and tested basic ecological theories at large scales. We hope that our findings on continental-scale distribution of zooplankton can serve as a baseline for understanding potential future departures from current and recent conditions.

ACKNOWLEDGEMENTS

We are grateful to researchers, professionals and technicians who carried out the sampling survey, and literature review at the Freshwater Institute in Winnipeg. We also express our thanks to Danielle Defaye at the Muséum National d'Histoire Naturelle in Paris who checked the nomenclature of the 83 crustacean species. A.A. carried out the study as a Master 2 student within the Groupe de Recherche Interuniversitaire en Limnologie et Environnement Aquatique at the Université de Montréal, which provided financial support. The study was supported by Discovery grants from the National Science and Engineering Research Council (NSERC) to B.P.A., P.L. and J.C.

REFERENCES

- Algar, A.C., Kerr, J.T. & D.J. Currie (2007) A test of Metabolic Theory as the mechanism underlying broad-scale species-richness gradients. *Global Ecology and Biogeography*, **16**, 170-178.
- Allen, A.P., Brown, J.H. & Gillooly, J.F. (2002) Global biodiversity, biochemical kinetics, and the energetic-equivalence rule. *Science*, **297**, 1545-1548.
- Anderson, M.J., Crist, T.O., Chase, J.M., Vellend, M., Inouye, B.D., Freestone, A.L., Sanders, N.J., Cornell, H.V., Comita, L.S., Davies, K.F., Harrison, S.P., Kraft, N.J.B., Stegen, J.C. & Swenson, N.G. (2011) Navigating the multiple meanings of β diversity: a roadmap for the practicing ecologist. *Ecology Letters*, **14**, 19-28.
- Arnott, S.E., Magnuson, J.J. & Yan, N.D. (1998) Crustacean zooplankton species richness: single- and multiple-year estimates. *Canadian Journal of Fisheries and Aquatic Sciences*, **55**, 1573-1582.
- Badgley, C. & Fox, D.L. (2000). Ecological biogeography of North American mammals: species density and ecological structure in relation to environmental gradients. *Journal of Biogeography*, **27**, 1437-1467.
- Blach-Overgaard, A., Svenning, J.-C. Dransfield, J., Greve, M. & Balslev, H. (2010) Determinants of palm species distributions across Africa: the relative roles of climate, non-climatic environmental factors, and spatial constraints. *Ecography*, **33**, 380-391.
- Borcard, D., Legendre, P. & Drapeau, P. (1992) Partialling out the spatial component of ecological variation. *Ecology*, **73**, 1045-1055.
- Brown, J.H., Gillooly, J.F., Allen, A.P., Savage, V.M. & West, G.B. (2004) Toward a metabolic theory of ecology. *Ecology*, **85**, 1771-1789.
- Chase, J.M. & Ryberg, W.A. (2004). Connectivity, scale-dependence, and the productivity–diversity relationship. *Ecology Letters*, **7**, 676-683.

- 483 Currie, D.J., Mittelbach, G.G., Cornell, H.V., Field, R., Guégan, J.-F., Hawkins, B.A.,
 484 Kaufman, D.M., Kerr, J.T., Oberdorff, T., O'Brien, E. & Turner, J.R.G. (2004) Predictions
 485 and tests of climate-based hypotheses of broad-scale variation in taxonomic richness.
 486 *Ecology Letters*, **7**, 1121-1134.
- 487 De Castro, F. & Gaedke, U. (2008) The metabolism of lake plankton does not support the
 488 metabolic theory of ecology. *Oikos*, **117**, 1218-1226.
- 489 Environment Canada (1986) *Climate Atlas climatique Canada: Map Series 2 - Precipitation*.
 490 Atmospheric environment Service, Ottawa, Ont.
- 491 ESRI (Environmental Systems Resource Institute). 2009. *ArcMap 9.2*. ESRI, Redlands,
 492 California. <http://forums.esri.com>.
- 493 Evans, K.L., Warren, P.H. & Gaston, K.J. (2005) Species-energy relationships at the
 494 macroecological scale: a review of the mechanisms. *Biological Reviews*, **80**, 1-25.
- 495 Fierer, N. (2008) Microbial biogeography: patterns in microbial diversity across space and
 496 time. In: *Accessing uncultivated microorganisms: from the environment to organisms and*
 497 *genomes and back*. K. Zengler (ed.). ASM Press, Washington DC. pgs. 95-115.
- 498 Fontenato, D. (ed.) (2011). *Biogeography of microscopic organisms. Is everything small*
 499 *everywhere?*. Cambridge University Press 365pp.
- 500 Fontaneto, D., Ficetola, G.F., Ambrosini, R. & Ricci, C. (2006) Patterns of diversity in
 501 microscopic animals: are they comparable to those in protists or in larger animals. *Global*
 502 *Ecology and Biogeography*, **15**, 153-162.
- 503 Gaston, K.J. (2000) Global patterns in biodiversity. *Nature*, **405**, 220-227.
- 504 Hawkins, B.A., Albuquerque, F., Araújo, M.B., Bini, L.M., Cabrero-Sañudo, F.J., Castro-
 505 Parga, I., Diniz-Filho, A.F., Ferrer-Castan, D., Field, R., Gómez, J.F., Hortal, J., Kerr, J.T.,
 506 Kitching, I.J., León-Cortés, J.L., Lobo, J.M., Montoya, D., Moreno, J.C., Olalla-Tárraga,
 507 M.Á., Pausas, J.G., Qian, H., Rahbek, C., Rodríguez, M.Á., Sanders, N.J. & Willians, P.

- 508 (2007a) A global evaluation of metabolic theory as an explanation for terrestrial species
509 richness gradients. *Ecology*, **88**, 1877-1888.
- 510 Hawkins, B.A., Diniz-Filho, A.F., Bini, L.M., Araùjo, M.B., Field, R., Hortal, J., Kerr, J.T.,
511 Rahbek, C., Rodriguez, M.A. & Sanders, N.J. (2007b) Metabolic theory and diversity
512 gradients: where do we go from here? *Ecology*, **88**, 1898-1902.
- 513 Hawkins, B.A., Field, R., Cornell, H.V., Currie, D.J., Guégan, J.F., Kaufman, D.M., Kerr,
514 J.T., Mittelback, G.G., Oberdorff, T., O'Brien, E.M., Porter, E.E. & Turner, J.R.G. (2003a)
515 Energy, water, and broad-scale geographic patterns of species richness. *Ecology*, **84**, 3105-
516 3117.
- 517 Hawkins, B.A., Porter, E.E. & Diniz-Filho, J.A.F. (2003b) Productivity and history as
518 predictors of the latitudinal diversity gradient of terrestrial birds. *Ecology*, **84**, 1608-1623.
- 519 Hessen, D.O., Faafeng, B.A., Smith, V.H., Bakkestuen, V. & Walseng, B. (2006) Extrinsic
520 and intrinsic controls of zooplankton diversity in lakes. *Ecology*, **87**, 433-443.
- 521 Hessen, D.O., Bakkestuen, V. & Walseng, B. (2007). Energy input and zooplankton species
522 richness. *Ecography*, **30**, 749-758.
- 523 Jiménez, I., Distler, T. & Jørgensen, P.M. (2009), Estimated plant richness pattern across
524 northwest South America provides similar support for the species-energy and spatial
525 heterogeneity hypotheses. *Ecography*, **32**, 433-448.
- 526 Kerr, J.T. & Currie, D. J. (1999) The relative importance of evolutionary and environmental
527 controls on broad-scale patterns of species richness in North America. *Ecoscience*, **6**, 329-
528 337.
- 529 Legendre, P. & Gallagher, E.D. (2001) Ecologically meaningful transformations for
530 ordination of species data. *Oecologia*, **129**, 271-280.
- 531 Legendre, P. & Legendre, L. (2012) *Numerical ecology*, 3rd english edition. Elsevier Science
532 BV, Amsterdam.

- 533 Legendre, P., Borcard, D. & Peres-Neto, P.R. (2005) Analyzing beta diversity: partitioning
534 the spatial variation of community composition data. *Ecological Monographs* **75**, 435-450.
- 535 Lomolino, M.V., Riddle, B.R., Whittaker, R.J. & Brown, J.H. (2010) *Biogeography*. Fourth
536 edition. Sinauer Associates, Inc. Publishers, Sunderland, Massachusetts, USA.
- 537 Marshall, I. & Schut, P. (1999) *A national ecological framework for Canada*. Agriculture and
538 Agrifood Canada. <http://sis.agr.gc.ca/cansis/nsdb/ecostrat/intro.html>.
- 539 Martiny, J.B. Hugues, Bohannan, B.J.M., Brown, J.H., Colwell, R.K., Fuhrman, J.A., Green,
540 J.L., Horner-Devine, M.C., Kane, M., Krumins, J.A., Kuske, C.R., Morin, P.J., Naeem, S.,
541 Øvreås, L., Reysenbach, A-L., Smith, V.H. & Staley, J.T (2006) Microbial biogeography:
542 putting microorganisms on the map. *Nature Review Microbiology*, **4**, 102–112.
- 543 McMahon, G., Wiken E.B. & Gauthier, D.A. (2004) Toward a scientifically rigorous basis for
544 developing mapped ecological regions. *Environmental Management*, 34, Suppl. 1, S111–
545 S124.
- 546 Melo, A.S., Rangel, T.F.L.V.B. & Diniz-Filho, A.F. (2009) Environmental drivers of beta-
547 diversity patterns in New-World birds and mammals. *Ecography*, **32**, 226-236.
- 548 Palmer, M.W. (1990) The estimation of species richness by extrapolation. *Ecology*, **71**, 1195-
549 1198.
- 550 Patalas, K. (1990) Diversity of the zooplankton communities in Canadian lakes as a function
551 of climate. *Verhandlungen des Internationalen Verein Limnologie*, **24**, 360-368.
- 552 Patalas, K., Patalas, J. & Salki, A. (1994) Planktonic crustaceans in lakes of Canada
553 (Distribution of Species, Bibliography). *Canadian Technical Report of Fisheries and*
554 *Aquatic Sciences* **1954**. 218p.
- 555 Peres-Neto, P.R., Legendre, P., Dray, S & Borcard, D. (2006) Variation partitioning of
556 species data matrices: estimation and comparison of fractions. *Ecology*, **87**, 2614-2625.
- 557 Phillips, D. (1990) *Climates of Canada*. Canadian Government Publishing, 176 pp.

- 558 Pinel-Alloul, B. & Ghadouani, A. (2007) Spatial heterogeneity of planktonic microorganisms
 559 in aquatic systems: multiscale patterns and processes. Chapter 8. *In: The importance of*
 560 *spatial scale on the analysis of patterns and processes in microbial communities*. R. B.
 561 Franklin & A. L. Mills. (Eds.). Kluwer Publishers, p 203-309.
- 562 Pinel-Alloul, B. (1995) Spatial heterogeneity as a multiscale characteristic of zooplankton
 563 community. *Hydrobiologia*, **300/301**, 17-42.
- 564 Pinel-Alloul, B., Niyonsenga, T. & Legendre, P. (1995) Spatial and environmental
 565 components of freshwater zooplankton structure. *Ecoscience* **2**: 1-19.
- 566 Ptacnik, R., Andersen, T., Brettum, P., Lepistö, L. & Willén, E. (2010) Regional species pools
 567 control community saturation in lake phytoplankton. *Proceedings of the Royal Society B -*
 568 *Biological Sciences*, **277**, 3755-3764.
- 569 R Development Core Team (2012) R: A language and environment for statistical computing,
 570 Vienna, Austria.
- 571 Rombouts, I., Beaugrand, G., Ibañez, F., Gasparini, S., Chiba, S. & Legendre, L. (2009)
 572 Global latitudinal variations in marine copepod diversity and environmental factors.
 573 *Proceedings of the Royal Society B - Biological sciences*, **276**, 3053-3062.
- 574 Shurin, J.B., Havel, J.E., Leibold, M.A. & Pinel-Alloul, B. (2000) Local and regional
 575 zooplankton species richness: a scale-independent test for saturation. *Ecology*, **81**, 3062-
 576 3073.
- 577 Shurin, J.B., Arnott, S.E., Hillebrand, H., Longmuir, A., Pinel-Alloul, B., Winder, M. & Yan,
 578 M.D. (2007) Diversity-stability relationships varies with latitude in zooplankton. *Ecology*
 579 *Letters*, **10**, 1-8.
- 580 Stomp, M., Huisman, J., Mittelbach, G.G., Lichman, E. & Klausmeier, C.A. (2011) Large-
 581 scale biodiversity patterns in freshwater phytoplankton. *Ecology*, **92**, 2096-2107.

- 582 Tuomisto, H. & Ruokolainen, K. (2006) Analyzing or explaining beta diversity?
583 Understanding the targets of different methods of analysis. *Ecology*, **87**, 2697-2708.
- 584 Vyverman, W., Vierleyen, E., Sabbe, K., Vanhoutte, K., Sterken, M., Hodgson, D.A., Mann,
585 G.G., Juggins, S., Van de Vijver, B., Jones, V., Flower, R., Roberts, D., Chepurnov, V.A.,
586 Kilroy, C., Vanormelingen, P. & De Wever, A. (2007) Historical processes constrain
587 patterns in global diatom diversity. *Ecology*, **88**, 1924–1931.
- 588 Walseng, B., Hessen, D.O., Halvorsen, G. & Schartau, A.K. (2006) Major contribution from
589 littoral crustaceans to zooplankton species richness in lakes. *Limnology & Oceanography*,
590 **51**, 2600-2606.
- 591 Whittaker, R.J., Willis, K.J. & Field, R. (2001) Scale and species richness: towards a general,
592 hierarchical theory of species diversity. *Journal of Biogeography*, **28**, 453-470.
- 593 Willig, M.R. & Bloch, C.P. (2006) Latitudinal gradients of species richness: a test of the
594 geographic area hypothesis at two ecological scales. *Oikos*, **112**, 163-173.
- 595
- 596

BIOSKETCH

B.P.A., P.L. and J.C. are research scientists in the Group for Interuniversity Research in Limnology and aquatic environment (GRIL), a major Canadian research centre in freshwater ecology. They co-supervised the master student (A.A). B.P.A research aims at studying patterns and processes of biodiversity and community structure of freshwater zooplankton at multiple scales; she developed the general research questions and hypotheses and led the writing. P.L. develops quantitative methods of spatial analysis in Numerical Ecology; he supervised the statistical analyses and contributed to ideas and writing. J.C. uses GIS applications to study patterns and processes at large spatial scales in the landscape; he supervised the GIS analysis and mapping, and contributed to ideas and writing. K.P. and A.S. provided the database on crustacean species occurrences. A.A. conducted the day-to-day work, validated the lake-zooplankton data, performed GIS and statistical analyses, wrote the initial draft of the article and participated in ideas and final writing and revisions. His contribution was crucial to the completion of the study and he should be considered as the first author of the paper.

612

Table 1 Minimum, maximum, median and mean values of regional species richness estimates (total number of species, averaged local species richness, Jackknife diversity index) and environmental descriptors in 35 ecoprovinces across Canada ($n = 1665$ lakes).

	Minimum	Maximum	Median	Mean
Total number of species	7	44	29	27
Averaged local species richness	3	10	5	6
Jackknife diversity index	8	52	36	34
Ecoprovince area (km² 10⁴)	1.9	62.6	15.8	23.2
Number of lakes	6	326	32	47
Growing season length (day)	19.4	258.8	163.1	144.8
Growing degree days above 10°C	0.0	1199.9	423.2	435.5
Effective growing degree days above 5°C	7.6	2133.2	1117.1	1042.6
Mean elevation (m)	36.1	1624.9	349.2	499.2
Total annual precipitation (mm)	101.2	2258.2	494.9	636.9
Mean daily global solar radiation (megajoules/m²/day)	8.3	13.6	11.4	11.2
Mean duration of bright sunshine (hrs)	1487.6	2338.0	1873.2	1879.3
Minimum annual air temperature (°C)*	-22.1	5.1	-5.6	-6.8
Maximum annual air temperature (°C)*	-15.6	13.8	6.0	3.0
Mean annual air temperature (°C)*	-18.8	9.4	0.3	-1.8
Mean annual vapour pressure	0.4	1.0	0.7	0.7
Annual potential evapotranspiration**	135.0	864.4	493.5	467.4
Longitude	-132.61	-64.09	-99.75	-100.22
Latitude	42.82	80.80	54.80	57.02

* over the entire year

** Penman method

621

622 **Table 2** Forward selection of regional environmental descriptors in multiple regression
 623 analysis of regional species richness (Jackknife diversity index; $n = 1665$ lakes).

624

625

Regional descriptors - Jackknife diversity index	R^2	R^2 Cum	R^2_{adj} Cum	F	p -value
Mean daily global solar radiation (megajoules/m ² /day)	0.52	0.52	0.51	36.09	0.00001
Ecoprovince area (km ² 10 ⁴)	0.08	0.60	0.58	6.66	0.01491
Effective growing degree days above 5°C	0.12	0.72	0.69	12.84	0.00111
Mean duration of bright sunshine (hours)	0.03	0.75	0.72	4.18	0.04949

Table 3 Linear regression models between regional species richness (Jackknife diversity index) and selected energy-related regional descriptors and the area of ecoprovinces ($n = 1665$ lakes)

Regional descriptors – Jackknife diversity index	R^2_{adj}	p -value
Mean daily global solar radiation (megajoules/m ² /day)	0.51	< 0.001
Annual potential evaporation **	0.48	< 0.001
Effective growing degree days above 5°C	0.44	< 0.001
Mean annual air temperature (°C)*	0.41	< 0.001
Mean duration of bright sunshine (hours)	0.32	< 0.001
Ecoprovince area (km ² 10 ⁴)	/	0.28

* over the entire year;

** Penman method

635

636 **Table 4** Forward selection in RDA model relating crustacean species community structure to
 637 polynomial terms of the regional and local environmental descriptors after a forward
 638 regression ($n = 458$ lakes; R^2_{adj} : 0.31, p -value = 0.008).

Regional and local descriptors	R^2	R^2_{adj} Cum	F	p -value
Mean daily global solar radiation 1	0.130	0.128	65.43	0.001
Mean duration of bright sunshine 1	0.042	0.169	22.09	0.001
Mean elevation 1	0.026	0.193	14.29	0.001
Mean annual vapour pressure 1	0.023	0.214	12.67	0.001
Total dissolved solids 1	0.012	0.224	6.60	0.001
July air temperature 2	0.008	0.231	4.57	0.001
July air temperature 1	0.007	0.236	4.16	0.001
July air temperature 3	0.008	0.243	4.66	0.001
Mean duration of bright sunshine 2	0.007	0.249	4.25	0.001
Mean daily global solar radiation 2	0.007	0.254	3.95	0.001
Mean annual air temperature 2	0.007	0.259	3.99	0.001
Mean daily global solar radiation 3	0.007	0.265	4.38	0.001
Secchi depth 1	0.006	0.269	3.54	0.001
Mean annual vapour pressure 2	0.005	0.273	3.25	0.001
Mean annual air temperature 1	0.006	0.277	3.60	0.001
Maximum annual air temperature 3	0.004	0.280	2.59	0.001
Mean duration of bright sunshine 3	0.004	0.283	2.43	0.001
Mean annual vapour pressure 3	0.004	0.285	2.50	0.001
Mean elevation 2	0.004	0.287	2.36	0.001
Total annual precipitation 1	0.004	0.290	2.40	0.001
Annual potential evapotranspiration 1	0.003	0.292	2.10	0.003
Total annual precipitation 3	0.004	0.294	2.27	0.002
Mean elevation 3	0.004	0.296	2.32	0.001
Annual potential evapotranspiration 3	0.003	0.298	1.98	0.002
Growing season length 2	0.004	0.300	2.22	0.001
Mean annual air temperature 1	0.003	0.301	1.92	0.003
Secchi depth 2	0.003	0.303	1.91	0.004
Lake depth 2	0.003	0.304	1.90	0.011
Lake depth 1	0.004	0.307	2.22	0.001
Lake depth 3	0.003	0.308	1.92	0.003
Growing season length 3	0.003	0.310	1.85	0.009
Effective growing degree day above 5°C 1	0.003	0.311	1.66	0.008

639

640

Figures captions

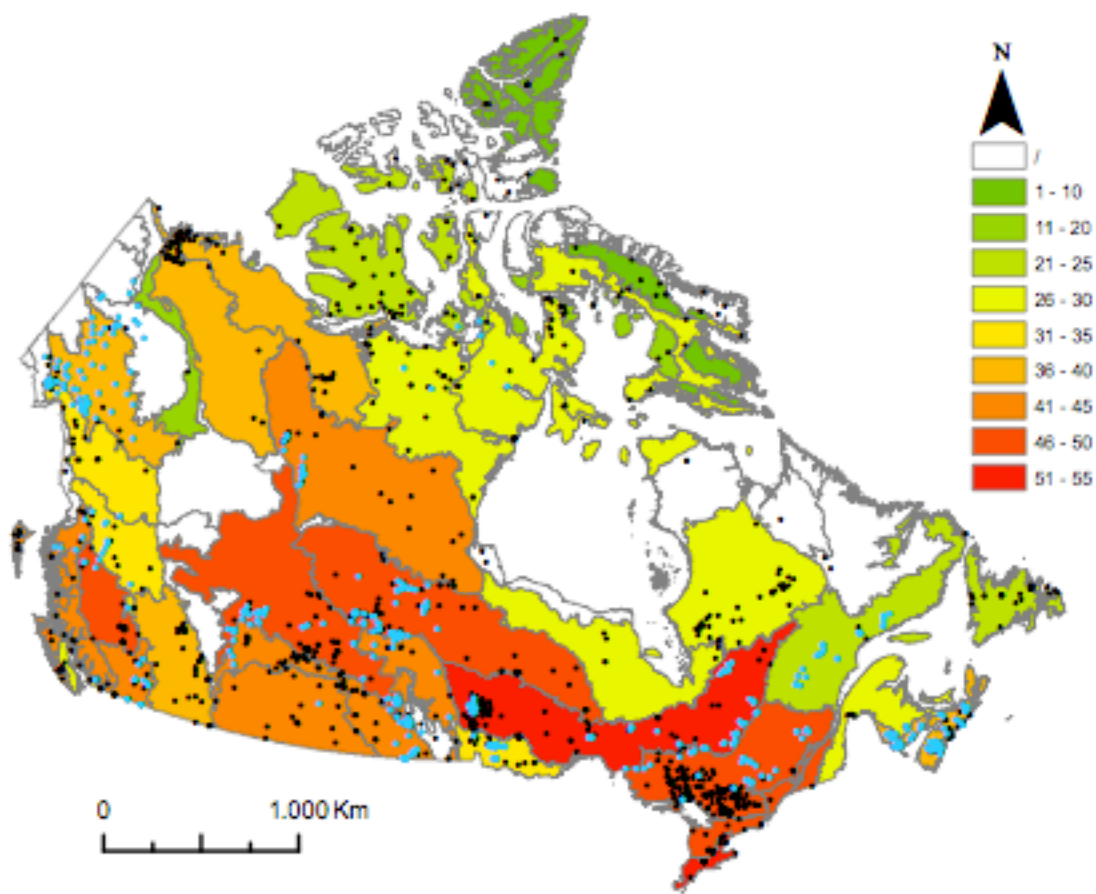
Figure 1 Distribution of the 1665 lakes used for describing lake crustacean diversity pattern across Canadian ecoprovinces. Colors of ecoprovinces are ranged according to increasing regional species richness based on the Jackknife diversity index. The localisation of the 458 lakes subset are indicated by light blue dots. Blank color (/): ecoprovinces having 5 or less than 5 sampled lakes.

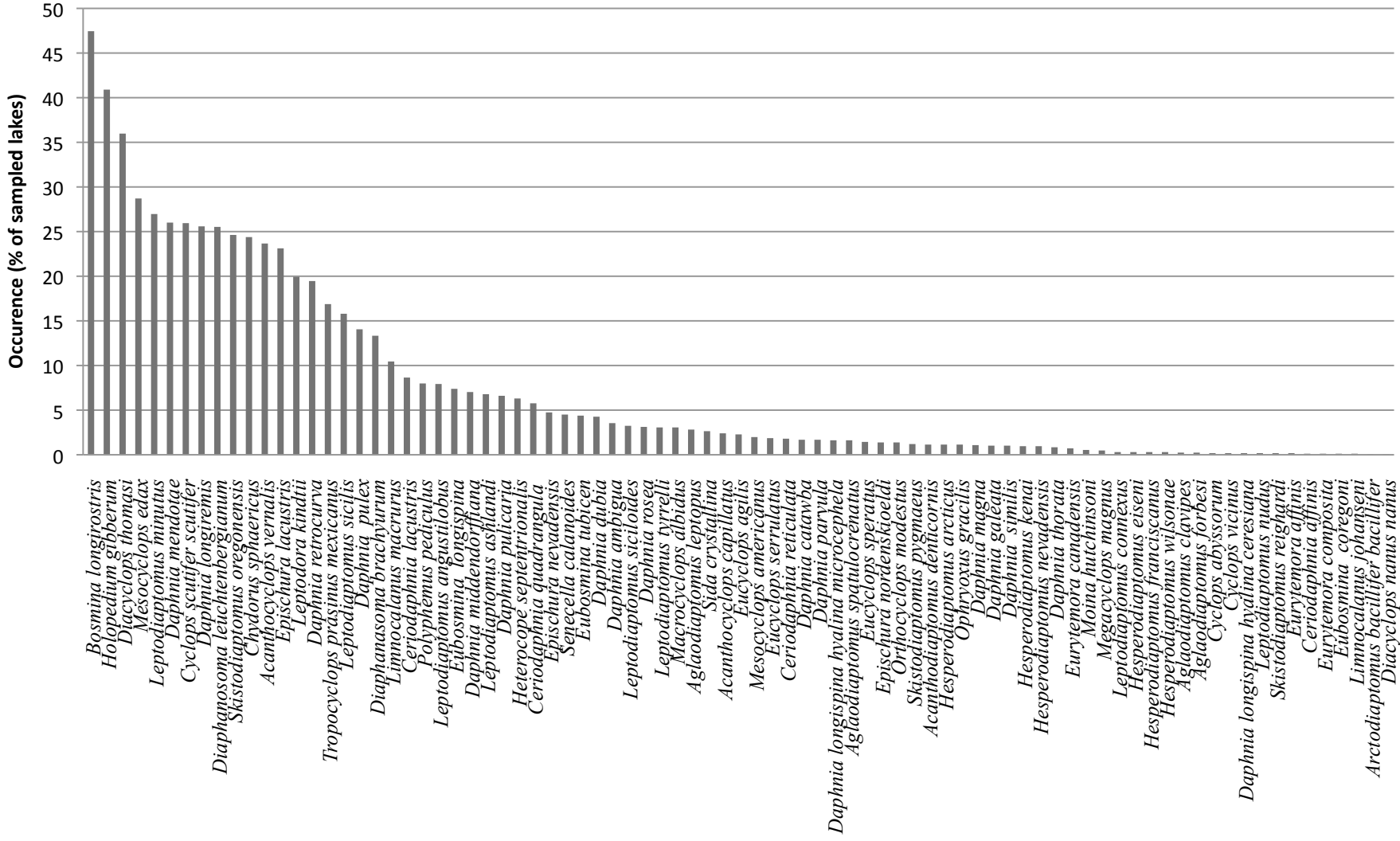
Figure 2 Rank-frequency occurrences of the 83 pelagic crustacean species collected in the 1665 lakes across Canada.

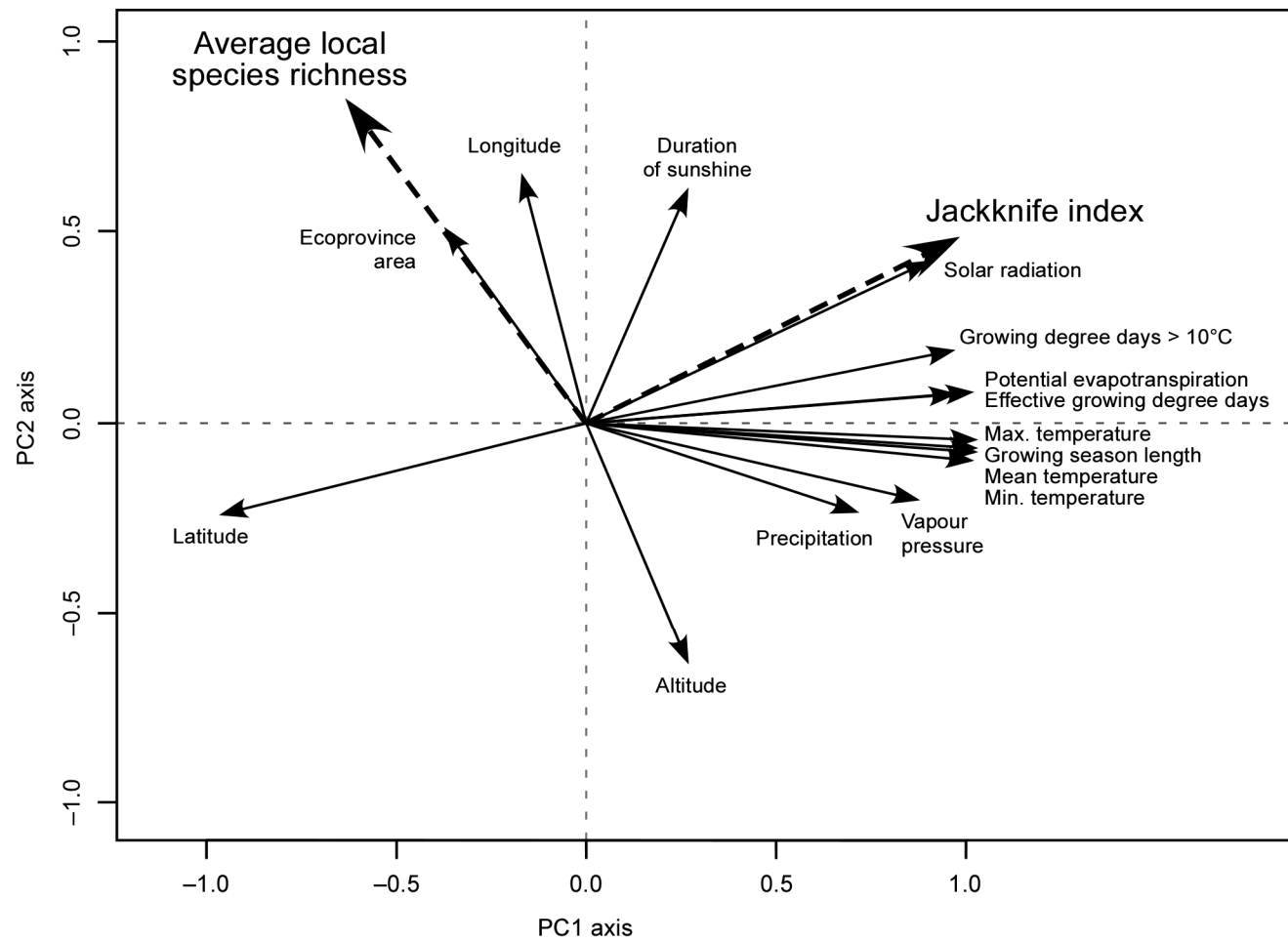
Figure 3 PCA ordination plots of the regional environmental descriptors selected by forward selection and the estimates of regional species richness (Averaged local species richness, Jackknife index).

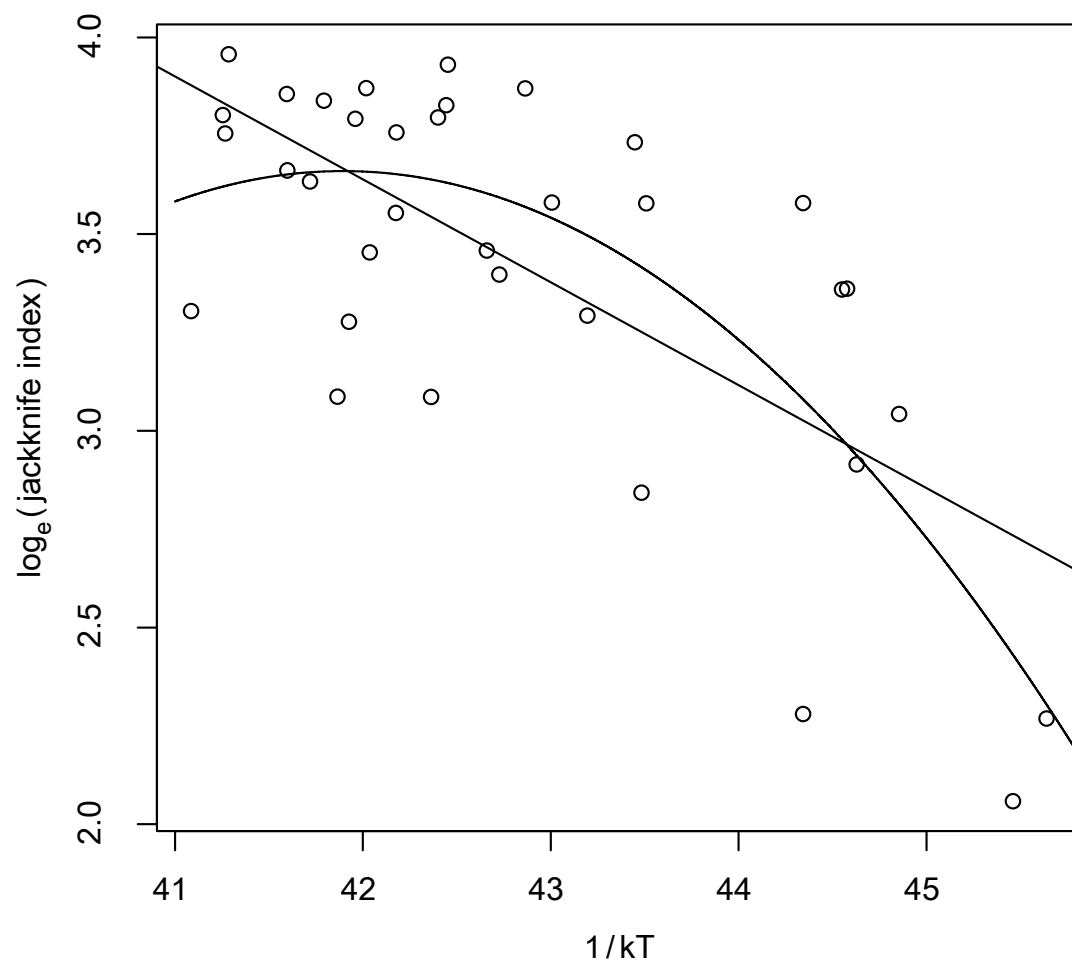
Figure 4 Linear and quadratic regressions between the natural logarithm of estimates of regional species richness (Jackknife index) and the inverse of temperature (in Kelvin) corrected by the Boltzmann constant.

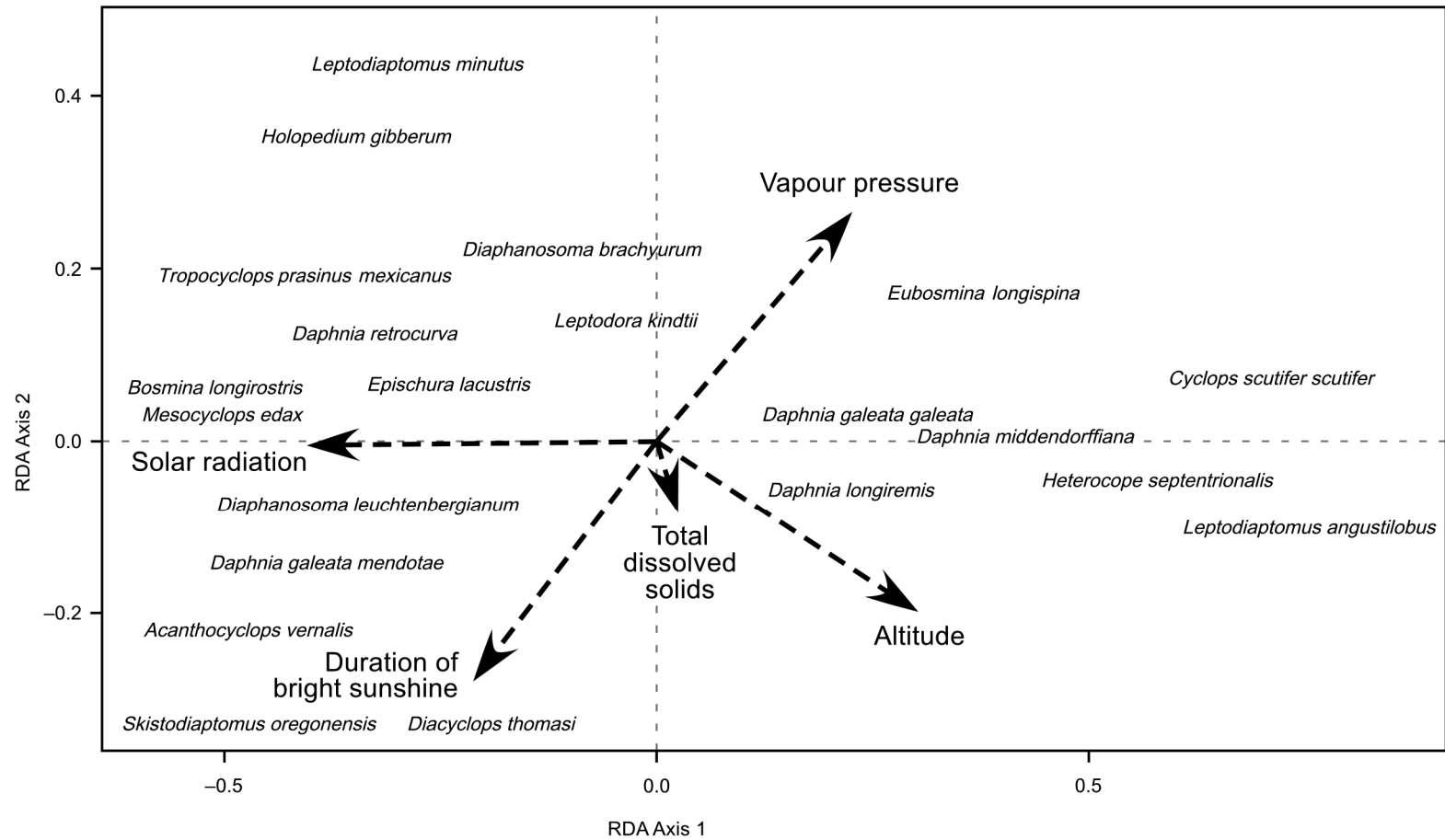
Figure 5 Redundancy analysis (RDA) between the matrix of occurrences of crustacean species in the 458 lakes and the five most important environmental descriptors selected by forward selection at regional and local scales.











Appendices and Supporting Information

Appendix S1 – Estimates of regional species richness (Average local species richness, Jackknife diversity index, Total number of crustacean species) and number of lakes in each of the 47 ecoprovinces located in the 15 ecozones of Canada. (Total numbers: $n = 1665$ lakes; $n = 83$ species). Data not calculated for 12 ecoprovinces with total number of species ≤ 10 or total number of lakes ≤ 5 .

Ecozone	Ecoprovince	Average local species richness	Jackknife diversity index	Total number of species	Number of lakes
Arctic Cordillera	Northern Arctic Cordillera	4	8	7	6
	Southern Arctic Cordillera	–	–	5	3
Atlantic Maritime	Appalachian-Acadian Highlands	8	27	22	10
	Fundy Uplands	8	39	34	74
Boreal Cordillera	Northern Boreal Cordillera	5	36	29	62
	Southern Boreal Cordillera	3	32	24	32
	Wrangel Mountains	–	–	7	3
Boreal Plains	Central Boreal Plains	7	46	41	102
	Eastern Boreal Plains	9	45	36	20
Boreal Shield	Eastern Boreal Shield	6	22	19	26
	Lake of the Woods	7	35	31	70
	Mid-Boreal Shield	9	51	43	126
	Newfouldland	3	22	19	31
	Southern Boreal Shield	6	48	44	326
	Western Boreal Shield	10	48	43	127
Columbia Montane	Columbia Montane Cordillera	4	38	32	42
Hudson Plains	Hudson Bay Coastal Plains	–	–	5	2
	Hudson-James Lowlands	7	30	22	8
Mixedwood Plains	Great Lakes-St.Lawrence	4	47	44	21
	Huron-Erie Plains	6	52	39	20
Montane Cordillera	Central Montane Cordillera	4	46	34	25
	Northern Montane Cordillera	5	32	25	18
	Southern Montane Cordillera	5	45	37	46
Northern Arctic	Baffin Uplands	3	10	8	9
	Boothia-Foxe Shield	4	29	20	39
	Ellesmere Basin	4	10	8	6
	Foxe-Boothia Lowlands	4	18	14	9
	Parry Channel Plateaux	–	–	5	4
	Sverdrup Islands	–	–	2	1
	Victoria Lowlands	4	21	19	44
	Georgia Depression	10	27	21	9
Pacific Maritime	Northern Coastal Mountains	–	–	3	2
	Southern Coastal Mountains	6	43	33	41
Prairies	Central Grassland	4	44	32	21
	Eastern Prairies	–	–	10	3
	Parkland Prairies	5	43	35	66
Southern Arctic	Amundsen Lowlands	6	36	29	44

	Keewatin Lowlands	5	29	24	28
Southern Arctic	Ungava-Belcher	–	–	3	2
Taiga Cordillera	Mackenzie-Selwyn Mountains	–	–	6	5
	Ogilvie Mountains	–	–	10	3
Taiga Plains	Great Bear Lowlands	5	36	29	36
	Mackenzie Foothills	4	17	13	6
Taiga Shield	Eastern Taiga Shield	7	27	24	38
	Labrador Uplands	–	–	12	4
	Western Taiga Shield	8	42	34	44
	Whale River Lowland	–	–	6	1

Appendix S2 – List of crustacean species recorded in the field and literature surveys based on the GIS-validated 1665 lakes.

Crustacean species	Lakes (1665)	Ecozones (15)	Écoprovinces (47)	Canadian provinces (15)
<i>Acanthocyclops capillatus</i> (G.O. Sars, 1863)	40	6	13	8
<i>Acanthocyclops vernalis</i> (Fischer, 1853)	394	13	33	15
<i>Acanthodiaptomus denticornis</i> (Wierzejski, 1887)	19	7	10	5
<i>Aglaodiaptomus clavipes</i> (Schacht, 1897)	4	2	2	1
<i>Aglaodiaptomus forbesi</i> Light, 1938	4	2	2	2
<i>Aglaodiaptomus leptopus</i> (Forbes, 1882)	47	7	16	7
<i>Aglaodiaptomus spatulocrenatus</i> (Pearse, 1906)	27	4	7	4
<i>Arctodiaptomus (Rhabdodiaptomus) bacillifer</i> (Koelbel, 1885)	1	1	1	1
<i>Bosmina (Bosmina) longirostris</i> (O.F. Müller, 1776)	790	15	39	15
<i>Ceriodaphnia affinis</i> Lilljeborg, 1900	2	2	2	2
<i>Ceriodaphnia lacustris</i> Birge, 1893	144	8	16	11
<i>Ceriodaphnia quadrangula</i> O.F. Müller, 1785	96	9	18	9
<i>Ceriodaphnia reticulata</i> (Jurine, 1820)	30	5	11	6
<i>Chydorus sphaericus</i> (O.F. Müller, 1785)	406	13	31	14
<i>Cyclops abyssorum</i> G.O. Sars, 1863	3	1	1	1
<i>Cyclops scutifer scutifer</i> G.O. Sars, 1863	432	14	33	14
<i>Cyclops vicinus</i> Ulianine, 1875	3	1	1	1
<i>Daphnia pulex</i> Leydig, 1860	234	11	26	11
<i>Daphnia similis</i> Claus, 1876	17	3	4	3
<i>Daphnia ambigua</i> Scourfield, 1947	59	7	12	6
<i>Daphnia catawba</i> Coker, 1926	28	4	6	4
<i>Daphnia dubia</i> Herick, 1883	71	4	9	5
<i>Daphnia galeata</i> Sars, 1864	17	3	3	3
<i>Daphnia longiremis</i> G. O. Sars, 1862	426	13	34	14
<i>Daphnia longispina (hyalina) f. ceresiana</i> Burckhard 1899	3	1	3	1
<i>Daphnia longispina (hyalina) f. microcephala</i> Ekman 1904	27	8	13	5
<i>Daphnia magna</i> Straus, 1820	18	7	9	6
<i>Daphnia mendotae</i> Birge, 1918	433	12	24	11
<i>Daphnia middendorffiana</i> Fischer, 1851	117	10	20	10
<i>Daphnia parvula</i> Fordyce, 1901	28	5	9	5
<i>Daphnia pulicaria</i> Forbes, 1893	110	9	20	9
<i>Daphnia retrocurva</i> Forbes, 1882	324	8	16	10
<i>Daphnia rosea</i> G.O. Sars, 1862	52	7	12	7
<i>Daphnia thorata</i> Forbes, 1893	14	2	6	1
<i>Diacyclops nanus nanus</i> (G.O. Sars, 1863)	1	1	1	1
<i>Diacyclops thomasi</i> (Forbes, 1882)	599	12	28	10
<i>Diaphanosoma brachyurum</i> (Liévin, 1848)	222	8	19	10
<i>Diaphanosoma leuchtenbergianum</i> Fischer, 1850	425	8	21	11
<i>Epischura (Epischura) lacustris</i> Forbes, 1882	385	9	20	10

<i>Epischura (Epischura) nevadensis</i> Lilljeborg, 1889	79	9	14	6
<i>Epischura (Epischura) nordenskioeldi</i> Lilljeborg, 1889	23	2	2	2
<i>Eubosmina (Eubosmina) coregoni</i> Baird, 1857	2	2	2	1
<i>Eubosmina (Eubosmina) longispina</i> (Leydig, 1860)	123	11	21	14
<i>Eubosmina (Neobosmina) tubicen</i> (Brehm, 1953)	73	4	5	5
<i>Eucyclops agilis</i> (Koch, 1838)	38	8	17	12
<i>Eucyclops elegans</i> (Herrick, 1884)	24	8	11	7
<i>Eucyclops serrulatus</i> (Fischer, 1851)	31	7	10	6
<i>Eurytemora affinis</i> (Poppe, 1880)	3	1	2	1
<i>Eurytemora canadensis</i> Marsh, 1920	12	2	4	3
<i>Eurytemora composita</i> Keiser, 1929	2	2	2	1
<i>Hesperodiaptomus arcticus</i> (Marsh, 1920)	19	5	10	6
<i>Hesperodiaptomus eiseni</i> (Lilljeborg, 1889)	5	3	3	4
<i>Hesperodiaptomus franciscanus</i> (Lilljeborg, 1889)	5	2	2	2
<i>Hesperodiaptomus kenai</i> Wilson, 1953	16	2	4	1
<i>Hesperodiaptomus nevadensis</i> (Light, 1938)	16	3	4	4
<i>Hesperodiaptomus wilsonae</i> (Reed, 1958)	5	3	4	2
<i>Heterocope septentrionalis</i> Juday & Muttkowski, 1915	105	10	16	6
<i>Holopedium gibberum</i> Zaddach, 1855	681	14	32	15
<i>Leptodiaptomus angustilobus</i> (G.O. Sars, 1898)	132	9	16	8
<i>Leptodiaptomus ashlandi</i> (Marsh, 1893)	113	8	17	9
<i>Leptodiaptomus connexus</i> (Light, 1938)	5	2	3	3
<i>Leptodiaptomus minutus</i> (Lilljeborg, 1889)	449	9	21	12
<i>Leptodiaptomus nudus</i> (Lilljeborg, 1889)	3	3	3	2
<i>Leptodiaptomus sicilis</i> (Forbes, 1882)	263	12	25	12
<i>Leptodiaptomus siciloides</i> (Lilljeborg, 1889)	54	5	10	6
<i>Leptodiaptomus tyrrelli</i> (Poppe, 1888)	51	3	9	3
<i>Leptodora kindtii</i> (Focke, 1844)	332	13	30	13
<i>Limnocalanus johanseni</i> Marsh, 1920	2	1	1	1
<i>Limnocalanus macrurus</i> G.O. Sars, 1863	174	10	23	10
<i>Macrocyclops albidus</i> (Jurine, 1820)	51	10	17	9
<i>Megacyclops magnus</i> (Marsh, 1920)	8	3	7	4
<i>Mesocyclops americanus</i> Dussart, 1985	33	7	9	6
<i>Mesocyclops edax</i> (S.A. Forbes, 1891)	478	7	18	9
<i>Moina hutchinsoni</i> Brehm, 1937	9	4	5	2
<i>Ophryoxus gracilis</i> G. O. Sars, 1861	19	4	8	6
<i>Orthocyclops modestus</i> (Herrick, 1883)	23	6	10	6
<i>Polyphemus pediculus</i> (Linnaeus, 1761)	133	11	21	12
<i>Senecella calanoides</i> Juday, 1923	75	7	11	9
<i>Sida crystallina</i> (O.F. Müller, 1776)	44	9	13	7
<i>Skistodiaptomus oregonensis</i> (Lilljeborg, 1889)	410	12	23	10
<i>Skistodiaptomus pygmaeus</i> (Pearse, 1906)	20	1	1	2
<i>Skistodiaptomus reighardi</i> (Marsh, 1895)	3	1	1	1
<i>Tropocyclops prasinus mexicanus</i> (Kiefer, 1938)	281	6	13	8

Appendix S3 – (A) Forward selection of regional environmental descriptors in multiple regression. (B) Simple linear regression analyses of the Average local species richness in each ecoprovince ($n = 1665$ lakes); the adjusted R^2 is shown when the regression coefficient was significant ($\alpha = 0.05$).

A

Regional descriptors – Average local species richness	R^2	R^2 Cum	R^2_{adj} Cum	F	p -value
Latitude	0.22	0.22	0.20	9.47	0.003
Mean elevation (m)	0.15	0.37	0.33	7.57	0.011
Ecoprovince area (km ² 10 ⁴)	0.08	0.45	0.40	4.73	0.041

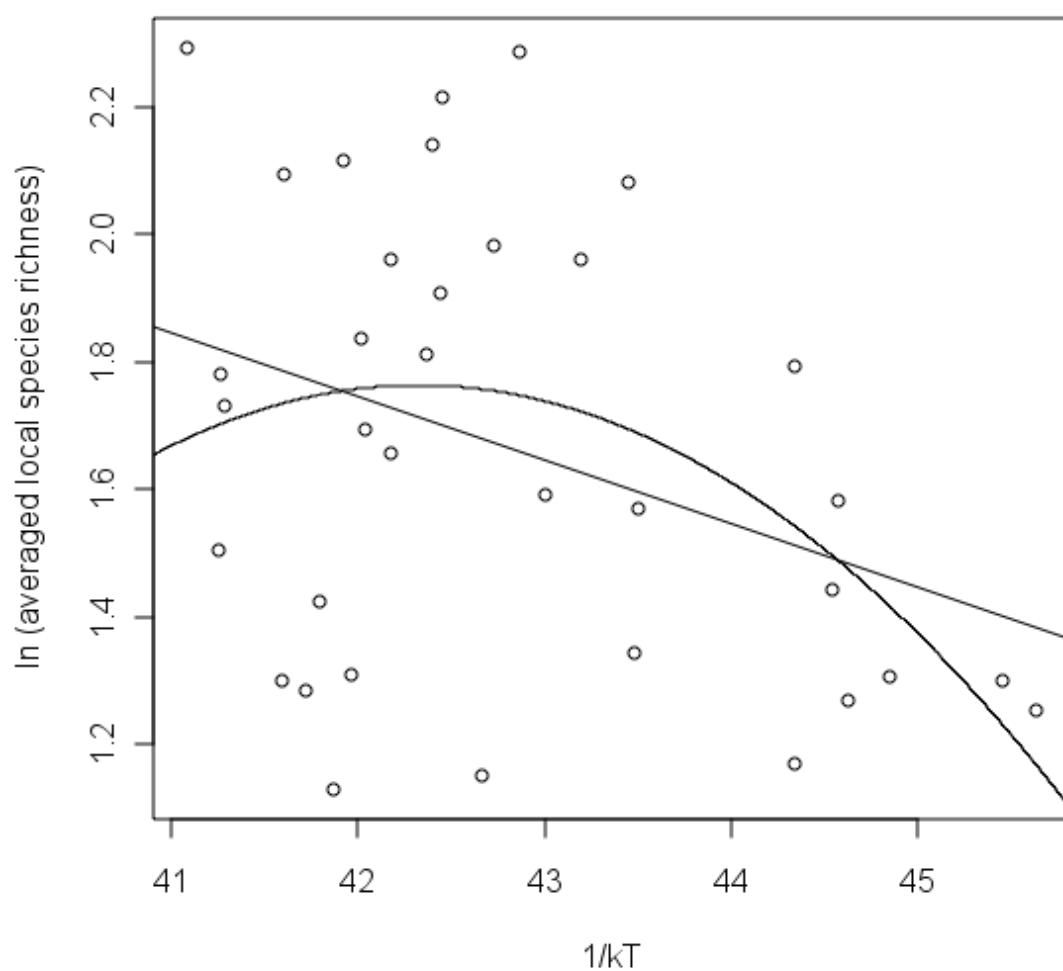
B

Regional descriptors – Average local species richness	R^2_{adj}	p -value
Mean daily global solar radiation (megajoules/m ² /day)	0.14	0.015
Annual potential evapotranspiration**	–	0.10
Effective growing degree days above 5°C	–	0.07
Mean annual air temperature (°C)*	0.09	0.04
Mean duration of bright sunshine (hours)	–	0.20
Ecoprovince area (km ² 10 ⁴)	–	0.10

* over the entire year

** Penman method

Appendix S4 – Linear and quadratic regressions between the natural logarithms of estimates of regional species richness (Average local species richness) and the inverse of temperature (in Kelvin) corrected by the Boltzmann constant.



Appendix S5 – Minimum, maximum, median and mean values of the local and regional environmental descriptors within ecoprovinces across Canada in the subset of 458 lakes.

	Minimum	Maximum	Median	Mean
Local variables				
July air temperature (°C)	10.0	19.5	17.2	16.6
Lake surface (km²)	0.0	19600.0	2.2	82.7
Lake depth (m)	0.6	283.0	10.5	18.1
Total dissolved solids (mg/L⁻¹)	3.7	9999.0	69.0	176.6
Secchi depth (m)	0.1	11.5	3.0	3.5
Longitude	-139.38	-61.27	-100.21	-99.14
Latitude	44.03	66.14	50.80	52.41
Regional variables				
Growing season length (day)	0.0	274.0	173.0	1698.0
Growing degree days above 10°C	0.0	1185.3	601.5	575.4
Effective growing degree days above 5°C	0.0	2198.4	1343.0	1320.1
Mean elevation (m)	8.0	1809.8	379.3	542.3
Total annual precipitation (mm)	0.0	2674.4	536.4	682.5
Mean daily global solar radiation (megajoules/m²/day)	9.3	13.7	12.4	12.00
Mean duration of bright sunshine (hrs)	1486.3	2335.3	1915.6	1964.8
Maximum annual air temperature (°C)*	-5.0	14.2	7.3	6.9
Mean annual air temperature (°C)*	-9.7	9.9	1.4	1.3
Mean annual vapour pressure	0.6	1.0	0.7	0.7
Annual potential evapotranspiration**	0.0	845.4	562.0	563.3

* over the entire year

** Penman method

Appendix S6 – Variation partitioning of the total explained variation (represented by the outside rectangle) of the crustacean community structure between regional (left box, $[a + b]$) and local (right box, $[b + c]$) environmental descriptors.

