Using empirical and simulated data to study the influence of environmental heterogeneity on fish species richness in two biogeographic provinces

Loss of species richness in aquatic ecosystems is occurring rapidly and many factors, including habitat heterogeneity, have been suggested to affect the diversity of aquatic communities. We used fish community data (> 200 species) from extensive surveys conducted in two biogeographic provinces (extent > 1000 km) in North America to test the hypothesis that fish species richness is greater in more heterogeneous habitats (grain < 10 km^2). Our tests are based on samples collected at nearly 800 stations over a period of five years. Using a set of environmental variables routinely measured by monitoring programs and a random placement model of community assembly, we demonstrate that fish species richness in coastal ecosystems is associated locally with the spatial heterogeneity of environmental variables but not with their magnitude. The observed effect of heterogeneity on species richness was substantially greater than that generated by simulations. Our modeling framework opens avenues for targeted conservation of habitat heterogeneity at broader temporal and spatial scales.

Using empirical and simulated data to study the influence of environmental heterogeneity on fish species richness in two biogeographic provinces

Philippe Massicotte[1,2], Raphaël Proulx[1], Gilbert Cabana[1], and Marco A. Rodríguez[1]

[1]Centre de Recherche sur les Interactions Bassins Versants-Écosystèmes aquatiques (RIVE). Université du Québec à Trois-Rivières, 3351, boul. des Forges, G9A 5H7
[2]Present address: Department of Bioscience. Aarhus University, Frederiksborgvej 399, DK-4000 Roskilde

ABSTRACT

Loss of species richness in aquatic ecosystems is occurring rapidly and many factors, including habitat heterogeneity, have been suggested to affect the diversity of aquatic communities. We used fish community data ($>$ 200 species) from extensive surveys conducted in two biogeographic provinces (extent $>$ 1000 km) in North America to test the hypothesis that fish species richness is greater in more heterogeneous habitats (grain $<$ 10 km^2). Our tests are based on samples collected at nearly 800 stations over a period of five years. Using a set of environmental variables routinely measured by monitoring programs and a random placement model of community assembly, we demonstrate that fish species richness in coastal ecosystems is associated locally with the spatial heterogeneity of environmental variables but not with their magnitude. The observed effect of heterogeneity on species richness was substantially greater than that generated by simulations. Our modeling framework opens avenues for targeted conservation of habitat heterogeneity at broader temporal and spatial scales.

Keywords: aquatic community assembly, conservation biology, diversity, heterogeneity hypothesis, random placement model, simulation

INTRODUCTION

The habitat heterogeneity hypothesis (MacArthur and MacArthur, 1961; MacArthur and Wilson, 1967) states that species richness increases with the number of ecological niches; that is, species coexistence is facilitated in more heterogeneous habitats because different taxa can capitalize on different environmental conditions. The hypothesis has been tested using many taxonomic groups across different spatial grains (average distance among observations) and extents (size of the whole study area) ranging from meters to thousands of kilometers. An extensive meta-analysis by Field et al. (2009) found that environmental heterogeneity was the primary factor driving species richness for 63 of the 273 cases (23%) assessing the relative importance of environmental heterogeneity versus other environmental factors. Environmental heterogeneity, however, had a stronger effect on species richness in studies conducted at small grain sizes (39% of the cases), suggesting that the relationship is contingent on the spatial scale. Furthermore, only 4 of the 393 relationships (1%) were from surveys of aquatic ecosystems having small grain size ($<$ 10 km^2) and large geographical extent ($>$ 1000 km).

Aquatic ecologists have faced difficulties in quantifying heterogeneity across different temporal and spatial scales (Kovalenko et al., 2011; Tisseuil et al., 2012; Yeager et al., 2011) possibly reflecting the difficulties of achieving the data needs to quantify such relationship. As a consequence, the term ’heterogeneity’ has been used rather loosely, as it could refer to habitat complexity, habitat diversity or environmental variability in both space and time (Palmer et al., 2010). For example, Oberdorff et al. (2011) assessed habitat heterogeneity at the continental scale using the proportion of different biomes found within river drainage basins, whereas Guégan et al. (1998) used the mean annual flow discharge as a proxy

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.567v1 | CC-BY 4.0 Open Access | rec: 30 Oct 2014, publ: 30 Oct 2014

for environmental heterogeneity in 183 rivers throughout the world. Although these two studies found a positive relationship between heterogeneity and fish species richness, their measures of environmental heterogeneity were confounded with biogeographic factors, such as the size of the drainage area, and with other global environmental descriptors including seasonality of rainfall. More recent studies of aquatic ecosystems investigated the heterogeneity hypothesis at smaller spatial grains and reported both positive (Buhl-Mortensen et al., 2010; Mellin et al., 2012) and negative (Kadmon and Allouche, 2007; Palmer et al., 2010) relationships between heterogeneity and the taxonomic richness of aquatic communities. Recent meta-analyses on the topic concluded that decrease in environmental heterogeneity always had a negative impact on diversity (Smokorowski and Pratt, 2007; Seiferling et al., 2014).

Given that species richness is declining in both freshwater and marine ecosystems (Ricciardi and Rasmussen, 1999; Worm et al., 2006), that coastal ecosystems are increasingly impacted by human activities, such as overfishing, oil drilling and regulation of river runoffs, and that conservation strategies are more easily enforced at local scales (Fausch et al., 2002), tests of the heterogeneity hypothesis under these circumstances are critically needed. The objective of this study was to evaluate the effect of environmental heterogeneity (spatial grain < 10 km^2) on fish species richness at the scale of biogeographic regions (spatial extent > 1000 km). We used data on fish communities (26 orders, 73 families, 136 genera, 204 species), obtained from extensive surveys in two coastal ecosystems of North America. Using a set of environmental variables routinely measured by monitoring programs, we demonstrate that fish species richness in coastal ecosystems responds positively to the spatial heterogeneity of environmental conditions. We further implemented a random placement model of community assembly to describe the relationship between environmental heterogeneity and species richness in the absence of explicit habitat selection mechanisms.

MATERIAL AND METHODS

Study site and data collection

Fish abundances and environmental measurements were obtained from two extensive surveys conducted by the by the U.S. Environmental Protection Agency's Environmental Monitoring and Assessment Program (EMAP). The first data set consisted of four sampling campaigns conducted in the Virginian biogeographic province between 1990 and 1993 (Hale et al., 2002). Stations were located along the coastline and in large river estuaries of the East Coast (Delaware, Hudson, Potomac, York; Fig. 1A). The second data set was assembled from four sampling campaigns conducted in the Louisiana biogeographic province between 1991 and 1994. Stations were located along the Gulf of Mexico from the Rio Grande, Texas, to Anclote Island, Florida (Fig. 1B). Field campaigns in the two biogeographic provinces were carried out between July and September of each year.

Fish were sampled using balloon trawls (funnel-shaped nets, 4.9 m wide with 2.5 cm stretched mesh) deployed from a research vessel using a hydraulic–powered boom in the vicinity of the sampling stations. The duration of the trawl was 10 ± 2 (mean $\pm$ SD) minutes at a speed of 2-3 knots. This corresponds to a length of 0.77 ± 0.15 (mean $\pm$ SD) km. Following a successful trawl, the net was hauled aboard and the catch was released into a plastic trough, or a fish sorting table, where species composition and abundance were recorded (see Appendix S1 in Supporting Information). A total of 2237 individuals (fork length: min. = 2.2 cm; max. = 91.18 cm; mean $\pm$ SD = 12.08 ± 7.33 cm) were captured from the Louisiana biogeographic province and 1883 individuals (fork length: min. = 2.5 cm; max. = 92.6 cm; mean $\pm$ SD = 16.03 ± 10.37 cm) were captured from the Virginian biogeographic province, yielding a total of 4120 individuals (Table 1, Appendix S1).

The environmental data comprised physical and chemical measurements. Dissolved oxygen concentrations (mg $\times$ L^{-1}) were determined using an air-calibrated oxygen meter (Yellow Springs Instruments) on surface water samples (625 mL) obtained with a Go-Flo bottle. Salinity (ppt), temperature (°C), pH, transmissivity (% of ambient light transmitted through the water column), photosynthetically active radiation (μE $\times$ m^{-2} $\times$ s^{-1}), fluorescence (unitless) and water density (σt, kg $\times$ m^{-3} – 1000) were measured using a SeaBird CTD meter lowered through the water column at a rate of approximately 0.25 m $\times$ s^{-1} until it reached the bottom (Table 1). Fluorescence and water density data were not available for the Louisiana surveys. Detailed information about the sampling and analytical procedures can be found on the EMAP web site (`http://www.epa.gov/emap/index.html`). Although other environmental variables such as macrophyte cover might be important determinants of environmental heterogeneity, the selected variables are known to affect the ecology of individual fish species (Mandrak, 1995).

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.567v1 | CC-BY 4.0 Open Access | rec: 30 Oct 2014, publ: 30 Oct 2014

Environmental heterogeneity

To represent the gradient of environmental conditions among stations of the same biogeographic province, we used the scores of a principal component analysis (PCA) performed on the environmental variables. The first three PCA axes (Table 1) were retained based on Kaiser's criterion and explained nearly 75% of the environmental variability in both Virginian (PC1 = 42.28%, PC2 = 19.7%, PC3 = 12.6%) and Louisiana (PC1 = 32.5%, PC2 = 23.8%, PC3 = 19.6%) biogeographic provinces. We quantified the degree of local spatial autocorrelation in environmental conditions near each station as a reciprocal measure of environmental heterogeneity. We calculated the local Moran I statistic on the scores of the first PCA axis using the `localmoran` function of the `spdep` package in R (Bivand et al., 2013). This statistic identifies station neighborhoods where environmental conditions of similarly high or low values cluster spatially (high I), as well as neighborhoods where environmental conditions are more contrasted (low I). High I values indicate low heterogeneity (positive autocorrelation), whereas values around zero indicate high heterogeneity. Negative I values indicate local over-dispersion patterns (i.e., negative autocorrelation), which are rarely observed in nature (Borcard et al., 2011). The I statistic is given by Anselin (1995):

$$I = (n-1)\frac{x_i - \bar{X}}{\sum\limits_{i=1}^{n}(x_i - \bar{X})^2}\sum_{j=1}^{n} w_{ij}(x_j - \bar{X}) \tag{1}$$

where x_i is the value of the observation i, is the mean of the variable, w_{ij} is the spatial weight ($1/distance^2$) between observations i and j, and n is the number of stations sampled. We used `dnearneigh` function of the `spdep` package to identify neighbours of region points by Euclidean distance between 0 and 75 km. Because we could not determine whether patterns of over-dispersion should be associated with high or low levels of environmental heterogeneity, the few stations (less than 4%) with negative I values were removed from subsequent statistical analyses. We did not find substantial differences between results for I calculated using all the data pooled at the biogeographic level (spatio-temporal I) and I calculated for each sampling year separately (spatial I). Consequently, we view I as a measure of spatial heterogeneity in local environmental conditions across space (Appendix S2, Fig. 1, Eq. 1).

Numerical simulations

We developed a random placement model of community assembly to determine the heterogeneity–species richness relationship in the absence of explicit habitat selection mechanisms. The model has two main components: (1) environmental heterogeneity and (2) species richness, each being simulated independently of the other on a two-dimensional surface (Fig. 2). This approach has been successfully used in various ecological studies aiming to highlight the effect of landscape structures on different aspects of animal biodiversity (Campos et al., 2013; McGill, 2011).

The first model component simulates the spatial patterns of environmental conditions (Fig. 2A). Environmental spatial patterns can be modeled as a fractional Brownian function. The spectral density $S(f)$ of a two-dimensional surface follows a power spectrum $S(f) \propto 1/f^{\beta}$ (Keitt, 2000), where f is frequency and $\beta = 1 + 2H$. The Hurst exponent (H) controls the degree of auto-correlation in environmental conditions; a large $H(H \longrightarrow 1)$ results in relatively homogeneous spatial patterns, whereas a lower $H(H \longrightarrow 0)$ produces more heterogeneous patterns. To generate the environmental spatial patterns in our simulations, we used the Matlab function `noiseonf`, which uses the inverse Fourier transformation of a power spectrum with a predetermined Hurst exponent (Kovesi, 2000). This procedure generates 'neutral' landscapes (e.g., With, 1997; Keitt, 2000) that share several statistical properties with environmental patterns observed in nature. The Hurst exponent of the simulated surface was parameterized using the linear slope of the log-log semi-variogram (Gallant et al., 1994) computed on the scores of the first axis of the PCA of environmental conditions, yielding values of $H \approx 0.4$ in both biogeographical provinces.

The second component (Fig. 2B) of our model simulates the random placement of species with different distribution ranges. We based our random placement model of community assembly on two premises (McGill and Collins, 2003; McGill, 2010): (1) the centroid of each species range is determined by sampling from a uniform distribution over the surface and (2) the range size of species is distributed according to a power distribution. McGill and Collins (2003) reported that implementing either a log-normal or a power distribution did not affect the results of random placement model. Each of our simulation runs proceeded as described in algorithm 1. Local species richness is then calculated by summing the overlap of different species ranges. On the basis of the observed regional distributions of the sampled species (Appendix S2, Fig. 3), we used the following parameters to implement the random placement model: $G = 1000$, $rmin = 10$ km and $rmax = 1000$ km.

Algorithm 1: Random placement of species (component 1, Fig. 2A)

1. Generate a surface of size $G \times G$.
2. Randomly pick the distribution range r of a new species from a power function $f(r) = r^{-a}$ where $r_{min} \leq r \leq r_{max}$ (Appendix S2, Fig. 2).
3. Choose the species centroid randomly from a uniform distribution over the surface.
4. Repeat previous steps until the surface is completely covered by species ranges (ranges are allowed to overlap).

To represent the range of each species on the surface, we used ellipses with major axis length r and minor axis length sampled from a uniform in the interval $[r/4, r/2]$ as described in Proulx et al. (2014). To simulate an anisotropic spatial process, we placed the elliptical ranges with their major axis oriented either horizontally (with probability = 0.75) or vertically (with probability = 0.25). This decision was motivated by the fact that species ranges in both biogeographical provinces are preferentially oriented along rivers and coastlines that broadly conform to the proposed alignment. Finally, to determine the parameter α empirically, we calculated the range of all fish species in each biogeographical province (Appendix S2, Fig. 3) and estimated the power coefficient of the frequency using the log-ratio formula (Eq. 5 in Newman, 2005). We obtained values of $\alpha = 1.214$ for the Virginian province and $\alpha = 1.189$ for the Louisiana province, and therefore used a value of 1.2 in our simulations. Using different combinations of ellipse shape ratio and orientation, we found that the species richness was robust to these changes. Most importantly, varying the shape ratio and orientation of ellipse (species range) did not affect the general direction and relative effect size of the simulated environmental heterogeneity-species richness relationship. We generated the two model components on grids of 1000 x 1000 cells (Fig. 2A and 2B). A total of 10 000 simulations where performed according to algorithm 2. It is to be noted that the model does not aim to approximate the absolute number of species at each location. Consequently, we used relative changes in species richness (Δ_S) to compare modeled and observed results.

Algorithm 2: Global simulation procedure

1. Generate an environmental grid (component 1, Fig. 2A).
2. Generate a species placement grid (component 2, Fig. 2B).
3. Randomly subsample 400 grid cells (roughly corresponding to the total number of sampling stations in each biogeographic province, Appendix S2, Fig. 4).
4. Calculate the local Moran's I at each subsampled cell on the environmental grid following the procedure described in the *Environmental heterogeneity* section (Equation 1, Appendix S2, Fig. 3).
5. Pair each local I value to its associated species richness value on the environmental and the species placement grid, respectively.
6. Fit a negative binomial regression between the paired values of local Moran's I and species richness (Fig. 2E).
7. Calculate the relative increase in species richness (Δ_S) predicted by the regression curve.

In each of the biogeographic provinces surveyed, approximately 5% of the stations yielded species richness values of zero. These zeros may partly arise from a 'veil effect' (Preston, 1948), and so reflect insufficient sampling effort rather than true absences. Truncation of samples at the veil may induce a spurious negative relationship between richness and predictor variables (Fig. 2E). To represent this effect in the simulated data, we set three veil lines at percentiles 0%, 5% and 15% and excluded species richness values below these thresholds (Fig. 4).

Statistical analyses

We used regression analyses to examine the relationships between species richness and the scores from the first PCA axis of environmental variables. To determine whether environmental heterogeneity had an influence on species diversity for both observed and simulated data, negative binomial regressions were fitted to the points above the veil effect threshold using the `glm.nb` function of the `MASS` package in R (version 3.0.1). We also checked for the presence of spatial autocorrelation in the model residuals.

RESULTS

Fish species richness was not correlated with any of the first three principal components from the analysis of environmental variables (Table 1; Fig. 3A, 3C), or with any of the individual environmental variables (results not shown). However, species richness was related to environmental heterogeneity (Fig. 3B and 3D). For both biogeographic provinces, the negative binomial regressions showed that species richness was greater in more heterogeneous environments (Fig. 3B and 3D). In the Virginian province (Fig. 3B), the mean species richness increased from 4.1 in homogeneous environments to 6.4 in heterogeneous environments, representing a gain of 2.3 $\pm$ 0.11 (95% confidence limits) species which correspond to 56% relative increase. A similar pattern was found for the Louisiana province (Fig. 3D) where mean species richness increased from 3.6 in homogeneous environments to 8.5 in heterogeneous environments, representing a gain of 4.9 $\pm$ 0.16 (95% confidence limits) species which correspond to 136% relative increase. We did not find spatial autocorrelation in the model residuals.

Averaging the results of 10 000 model simulations, the mean species richness relative increase (Δ_S) were of 3.25%, 5.28% and 6.66% for the 0%, 5% and 15% veil effects, respectively (Fig. 4). The probabilities of observing Δ_S greater or equal to 56% (Virginia province) due to a sampling effect for different veils (0%, 5%, 15%) were of 4.68%, 3.7% and 2.12%, respectively (Table 2). Considering a Δ_S of 136% threshold (Louisiana province), these probabilities dropped to 0.05%, 0.01% and 0% (Table 2).

DISCUSSION

Many factors, including habitat heterogeneity, have been reported to affect the diversity of aquatic communities (Field et al., 2009). However, it is likely that the set of factors influencing species richness differs across spatial and temporal scales (Fausch et al., 2002). Moreover, the heterogeneity of the habitat has been identified as a key factor maintaining the animal biodiversity in aquatic environments (Levin et al., 2010). This work combines data from extensive surveys and simulations to demonstrate a positive influence of environmental heterogeneity (*sensu stricto*) on the species richness of fish communities at scales that fish perceive and respond to in their local context. Furthermore, the observed effect of heterogeneity on species richness was substantially greater (Fig. 3) than that generated by the simulations based on a random community assembly model, so it seems unlikely that the observed relationship arose solely as a byproduct of veil or sampling effects.

Environmental variables

Results from Field et al. (2009) and Guégan et al. (1998) suggest that climatic and primary productivity variables have a major influence on species richness at both regional and continental scales. Studies conducted at small grain indicate that environmental variables influence the species presence–absence and abundance structure in local fish communities in both space and time (Menge and Olson, 1990; Rodríguez and Lewis, 1997; Thiel et al., 1995). In contrast to these findings, we did not observe any direct effect of individual environmental variables (Table 1), including salinity, chlorophyll-*a* concentration, and water temperature, on the species richness of local fish communities in either the Virginian (Fig. 3A) or Louisianan (Fig. 3C) biogeographic provinces.

Our simulation framework assumed no relationship between fish species richness and environmental conditions at the site of capture; an assumption supported by empirical data in the present study. Another major assumption of random placement models is that the probability of finding a fish species at a particular site is independent of other species. Such ecological independence between co-occurring species has been shown to accurately reproduce a number of community patterns (McGill, 2010, 2011). For example, a recent study of shrubland plant communities reported that only 7 to 19% of all species pairs showed strong and consistent spatial associations, leading the authors to conclude that ecological processes are leaving no discernible spatial signature (Perry et al., 2014). In contrary, our results suggest

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.567v1 | CC-BY 4.0 Open Access | rec: 30 Oct 2014, publ: 30 Oct 2014

that coastal fish communities may show this signature, as fish species richness was not associated locally with the magnitude of environmental variables, but rather with their spatial heterogeneity.

Environmental heterogeneity

Environmental heterogeneity influences many ecological processes such as fluxes of organisms, material and energy among riverscape elements (Pickett and Cadenasso, 1995). Our results demonstrate that fish species richness responded positively to increased habitat heterogeneity (Fig. 3B and 3D) in both the Virginian and Louisianan biogeographic provinces. Simulations using a random placement model of community assembly showed that species richness increased only slightly in more heterogeneous environments (Fig. 4). For instance, less than 5% of the 10 000 simulations generated Δ_S greater than the conservative value of 56% observed in the Virgina biogeographic province (Fig. 3, Fig. 4, Table 2). Hence, it is unlikely that the positive relationship observed between environmental heterogeneity and species richness in both biogeographic provinces is the result of a sampling effect (*sensu* McGill, 2011).

Aquatic ecologists often use the term 'heterogeneity' rather loosely to refer to habitat complexity, habitat diversity or environmental variability over time (reviewed in Palmer et al., 2010). For example, at small scales, heterogeneity usually refers to the variability in structural physical properties of the aquatic habitat such as riparian vegetation, channel configuration, artificial riffles and substrate granulometry (Palmer et al., 2010). Conversely, studies conducted at regional or continental scales have used large-grained variables such as percentage of different types of biome or drainage area as a proxy for habitat heterogeneity (Field et al., 2009; Guégan et al., 1998; Oberdorff et al., 2011), possibly reflecting the difficulty of obtaining information at a finer resolution. Consequently, studies conducted at regional or continental scales are likely to capture broad-scale environmental heterogeneity that is coarse relative to the local heterogeneity to which individual fish respond, particularly for species having ranges smaller than the study grain size (O'Neill et al., 1986; Turner et al., 1989; Wiens, 1989).

Conclusions

Over the last century, coastal ecosystems have become increasingly impacted by anthropogenic pressures (Lotze et al., 2006), including many human–driven activities that reduce the temporal and spatial heterogeneity of coastal habitats. For example, commercial fish trawlers are known to reduce the spatial heterogeneity of the sea floor structure (Helfman, 2007). Similarly, the temporal variability of water flows in many of the world's largest rivers are regulated by dams (Nilsson et al., 2005). This reduced variability in runoffs has been shown to increase the homogeneity of water channels, as well as to degrade fish habitats (see Moyle and Mount, 2007 and references therein). The current study shows that, independently of the environmental conditions prevailing locally, more homogeneous habitats can support fewer fish species. Hence, restoring or actively protecting areas of high habitat heterogeneity appears of great importance for slowing actual trends of decreasing biodiversity in coastal ecosystems.

ACKNOWLEDGMENTS

We thank the U.S. Environmental Protection Agency's Environmental Monitoring and Assessment Program (EMAP) for freely providing their data. Although the data described in this article have been funded wholly or in part by the U.S. Environmental Protection Agency through its EMAP Estuaries Program, it has not been subjected to Agency review, and therefore does not necessarily reflect the views of the Agency and no official endorsement should be inferred. K. Roach made helpful comments on an earlier version of the manuscript. P. Massicotte was supported by a postdoctoral fellowship from the Centre de Recherche sur les Interactions Bassins Versants-Écosystèmes aquatiques (RIVE), University of Québec at Trois-Rivières.

REFERENCES

Anselin, L. (1995). Local indicators of spatial association—LISA. *Geographical analysis*, 27(2):93–115.

Bivand, R., with contributions by Micah Altman, Anselin, L., Assunção, R., Berke, O., Bernat, A., Blanchet, G., Blankmeyer, E., Carvalho, M., Christensen, B., Chun, Y., Dormann, C., Dray, S., Halbersma, R., Krainski, E., Legendre, P., Lewin-Koh, N., Li, H., Ma, J., Millo, G., Mueller, W., Ono, H., Peres-Neto, P., Piras, G., Reder, M., Tiefelsdorf, M., and Yu., D. (2013). *spdep: Spatial dependence: weighting schemes, statistics and models*.

Borcard, D., Gillet, F., and Legendre, P. (2011). *Numerical ecology with R*. Springer.

Buhl-Mortensen, L., Vanreusel, A., Gooday, A. J., Levin, L. a., Priede, I. G., Buhl-Mortensen, P. l., Gheerardyn, H., King, N. J., and Raes, M. (2010). Biological structures as a source of habitat heterogeneity and biodiversity on the deep ocean margins. *Marine Ecology*, 31(1):21–50.

Campos, P. R., Rosas, A., de Oliveira, V. M., and Gomes, M. A. (2013). Effect of Landscape Structure on Species Diversity. *PloS one*, 8(6):e66495.

Fausch, K. D., Torgersen, C. E., Baxter, C. V., and Li, H. W. (2002). Landscapes to Riverscapes: Bridging the Gap between Research and Conservation of Stream Fishes. *BioScience*, 52(6):483–498.

Field, R., Hawkins, B. A., Cornell, H. V., Currie, D. J., Diniz-Filho, J. A. F., Guégan, J.-F., Kaufman, D. M., Kerr, J. T., Mittelbach, G. G., Oberdorff, T., O'Brien, E. M., and Turner, J. R. G. (2009). Spatial species-richness gradients across scales: a meta-analysis. *Journal of Biogeography*, 36(1):132–147.

Gallant, J., Moore, I., Hutchinson, M., and Gessler, P. (1994). Estimating fractal dimension of profiles: a comparison of methods. *Mathematical Geology*, 26(4):455–481.

Guégan, J.-F., Lek, S., and Oberdorff, T. (1998). Energy availability and habitat heterogeneity predict global riverine fish diversity. *Nature*, 391(January):382–384.

Hale, S., Hughes, M., and Strobel, C. (2002). Coastal ecological data from the Virginian Biogeographic Province, 1990-1993. *Ecology*, 83(10):2942–2942.

Helfman, G. S. (2007). *Fish Conservation: A Guide to Understanding and Restoring Global Aquatic Biodiversity and Fishery Resources*. Island Press.

Kadmon, R. and Allouche, O. (2007). Integrating the effects of area, isolation, and habitat heterogeneity on species diversity: a unification of island biogeography and niche theory. *The American Naturalist*, 170(3):443–454.

Keitt, T. (2000). Spectral representation of neutral landscapes. *Landscape Ecology*, 15:479–493.

Kovalenko, K. E., Thomaz, S. M., and Warfe, D. M. (2011). Habitat complexity: approaches and future directions. *Hydrobiologia*, 685(1):1–17.

Kovesi, P. D. (2000). MATLAB and Octave Functions for Computer Vision and Image Processing.

Levin, L., Sibuet, M., and Gooday, A. (2010). The roles of habitat heterogeneity in generating and maintaining biodiversity on continental margins: an introduction. *Marine Ecology*, 31(1):1–5.

Lotze, H., Lenihan, H., and Bourque, B. (2006). Depletion, degradation, and recovery potential of estuaries and coastal seas. *Science*, 312(5781):1806–1809.

MacArthur, R. and MacArthur, J. (1961). On Bird Species Diversity. *Ecology*, 42(3):594–598.

MacArthur, R. H. and Wilson, E. O. (1967). *The theory of island biogeography*. Princeton University Press, Princeton, N.J.

Mandrak, N. E. (1995). Biogeographic patterns of fish species richness in Ontario lakes in relation to historical and environmental factors. *Canadian Journal of Fisheries and Aquatic Sciences*, 52(7):1462–1474.

McGill, B. and Collins, C. (2003). A unified theory for macroecology based on spatial patterns of abundance. *Evolutionary Ecology Research*, 5:469–492.

McGill, B. J. (2010). Towards a unification of unified theories of biodiversity. *Ecology letters*, 13(5):627–42.

McGill, B. J. (2011). Linking biodiversity patterns by autocorrelated random sampling. *American journal of botany*, 98(3):481–502.

Mellin, C., Parrott, L., Andréfouët, S., Bradshaw, C. J. a., MacNeil, M. A., and Caley, M. J. (2012). Multi-scale marine biodiversity patterns inferred efficiently from habitat image processing. *Ecological Applications*, 22(3):792–803.

Menge, B. A. and Olson, A. M. (1990). Role of scale and environmental factors in regulation of community structure. *Trends in Ecology & Evolution*, 5(2):52–57.

Moyle, P. B. and Mount, J. F. (2007). Homogenous rivers, homogenous faunas. *Proceedings of the*

National Academy of Sciences of the United States of America, 104(14):5711–5712.

Newman, M. (2005). Power laws, Pareto distributions and Zipf's law. *Contemporary Physics*, 46(5):323–351.

Nilsson, C., Reidy, C., Dynesius, M., and Revenga, C. (2005). Fragmentation and flow regulation of the world's large river systems. *Science*, 308(5720):405–408.

Oberdorff, T., Tedesco, P. a., Hugueny, B., Leprieur, F., Beauchard, O., Brosse, S., and Dürr, H. H. (2011). Global and Regional Patterns in Riverine Fish Species Richness: A Review. *International Journal of Ecology*, 2011:1–12.

O'Neill, R. V., Deangelis, D. L., Waide, J. B., and Allen, G. E. (1986). *A Hierarchical Concept of Ecosystems.* Princeton University Press.

Palmer, M. A., Menninger, H. L., and Bernhardt, E. (2010). River restoration, habitat heterogeneity and biodiversity: a failure of theory or practice? *Freshwater Biology*, 55(Suppl. 1):205–222.

Perry, G. L. W., Miller, B. P., Enright, N. J., and Lamont, B. B. (2014). Stochastic geometry best explains spatial associations among species pairs and plant functional types in species-rich shrublands. *Oikos*, 123(1):99–110.

Pickett, S. T. and Cadenasso, M. L. (1995). Landscape ecology: spatial heterogeneity in ecological systems. *Science*, 269(5222):331–334.

Preston, F. (1948). The commonness, and rarity, of species. *Ecology*, 29(3):254–283.

Proulx, R., Roca, I. T., Cuadra, F. S., Seiferling, I., and Wirth, C. (2014). A novel photographic approach for monitoring the structural heterogeneity and diversity of grassland ecosystems. *Journal of Plant Ecology*, pages 1–8.

Ricciardi, A. and Rasmussen, J. B. (1999). Extinction Rates of North American Freshwater Fauna. *Conservation Biology*, 13(5):1220–1222.

Rodríguez, M. A. and Lewis, W. M. (1997). Structure of fish assemblages along environmental gradients in floodplain lakes of the Orinoco River. *Ecological Monographs*, 67(1):109–128.

Seiferling, I., Proulx, R., and Wirth, C. (2014). Disentangling the environmental-heterogeneity–species-diversity relationship along a gradient of human footprint. *Ecology*, 95(8):2084–2095.

Smokorowski, K. and Pratt, T. (2007). Effect of a change in physical structure and cover on fish and fish habitat in freshwater ecosystems-a review and meta-analysis. *Environmental Reviews*, 15:15–41.

Thiel, R., Sepulveda, A., Kafemann, R., and Nellen, W. (1995). Environmental factors as forces structuring the fish community of the Elbe Estuary. *Journal of Fish Biology*, 46(1):47–69.

Tisseuil, C., Cornu, J.-F., Beauchard, O., Brosse, S., Darwall, W., Holland, R., Hugueny, B., Tedesco, P. a., and Oberdorff, T. (2012). Global diversity patterns and cross-taxa convergence in freshwater systems. *The Journal of animal ecology*, 82(ii):365–376.

Turner, M. G., O'Neill, R. V., Gardner, R. H., and Milne, B. T. (1989). Effects of changing spatial scale on the analysis of landscape pattern. *Landscape Ecology*, 3(3-4):153–162.

Wiens, J. A. J. (1989). Spatial scaling in ecology. *Functional ecology*, 3(4):385–397.

With, K. A. (1997). The Application of Neutral Landscape Models in Conservation Biology. Aplicacion de Modelos de Paisaje Neutros en la Biologia de la Conservacion. *Conservation Biology*, 11(5):1069–1080.

Worm, B., Barbier, E. B., Beaumont, N., Duffy, J. E., Folke, C., Halpern, B. S., Jackson, J. B. C., Lotze, H. K., Micheli, F., Palumbi, S. R., Sala, E., Selkoe, K. a., Stachowicz, J. J., and Watson, R. (2006). Impacts of biodiversity loss on ocean ecosystem services. *Science*, 314(5800):787–90.

Yeager, L. A., Layman, C. A., and Allgeier, J. E. (2011). Effects of habitat heterogeneity at multiple spatial scales on fish community assembly. *Oecologia*, 167(1):157–68.

Figure 1 (on next page)

Main figures of the article

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.567v1 | CC-BY 4.0 Open Access | rec: 30 Oct 2014, publ: 30 Oct 2014

Using empirical and simulated data to study the influence of environmental heterogeneity on fish species richness in two biogeographic provinces

Figures

Philippe Massicotte[1], Raphaël Proulx[1], Gilbert Cabana[1], and Marco A. Rodríguez[1]

[1]Centre de Recherche sur les Interactions Bassins Versants- Écosystèmes aquatiques (RIVE). Université du Québec à Trois-Rivières, Trois-Rivières, Canada

October 20, 2014

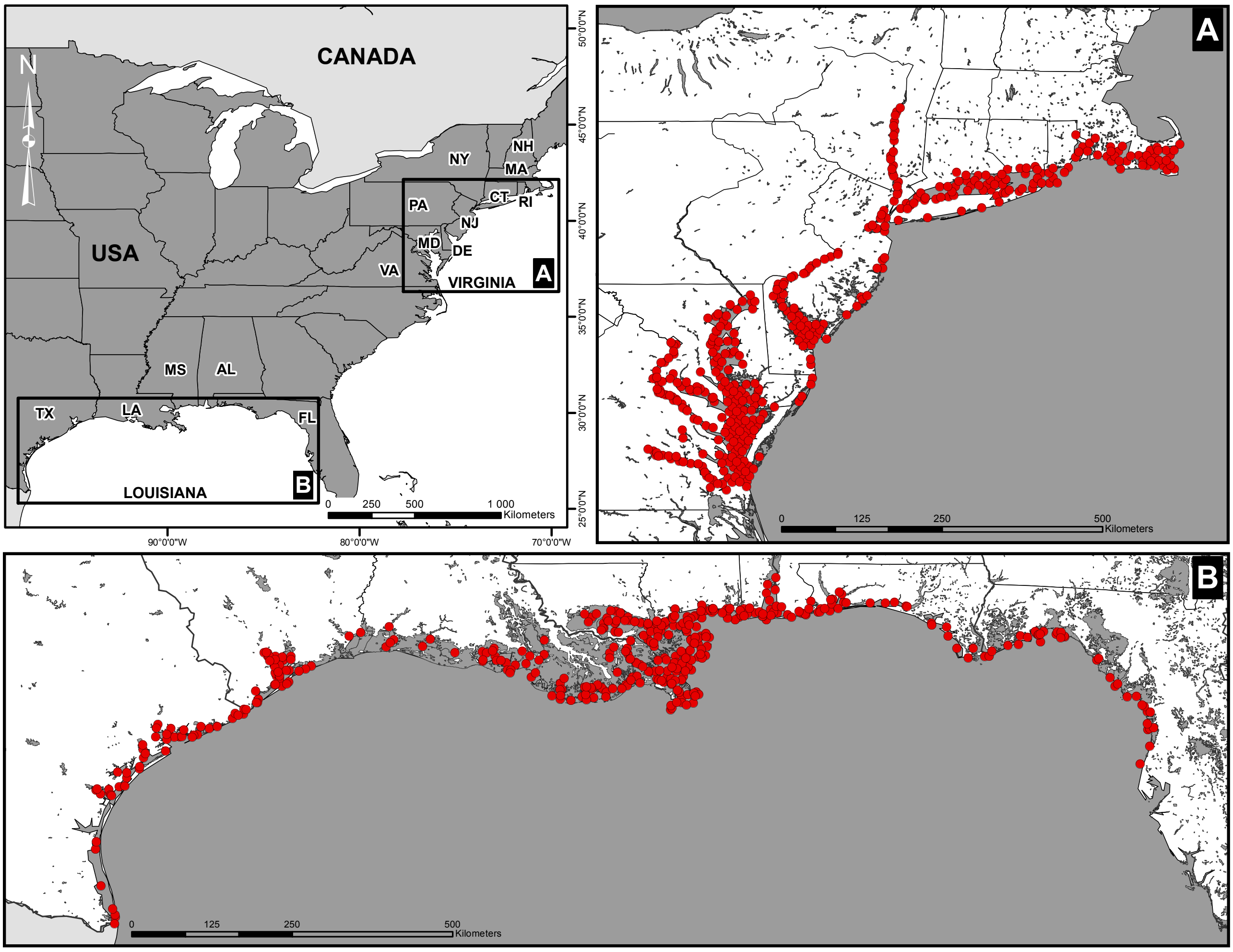

Figure 1: Spatial distribution of sampling sites for (**A**) Virginia and (**B**) Louisiana biogeographic provinces. Surveys were conducted by the U.S. Environmental Protection Agency's Environmental Monitoring and Assessment Program (EMAP) between 1990 and 1994.

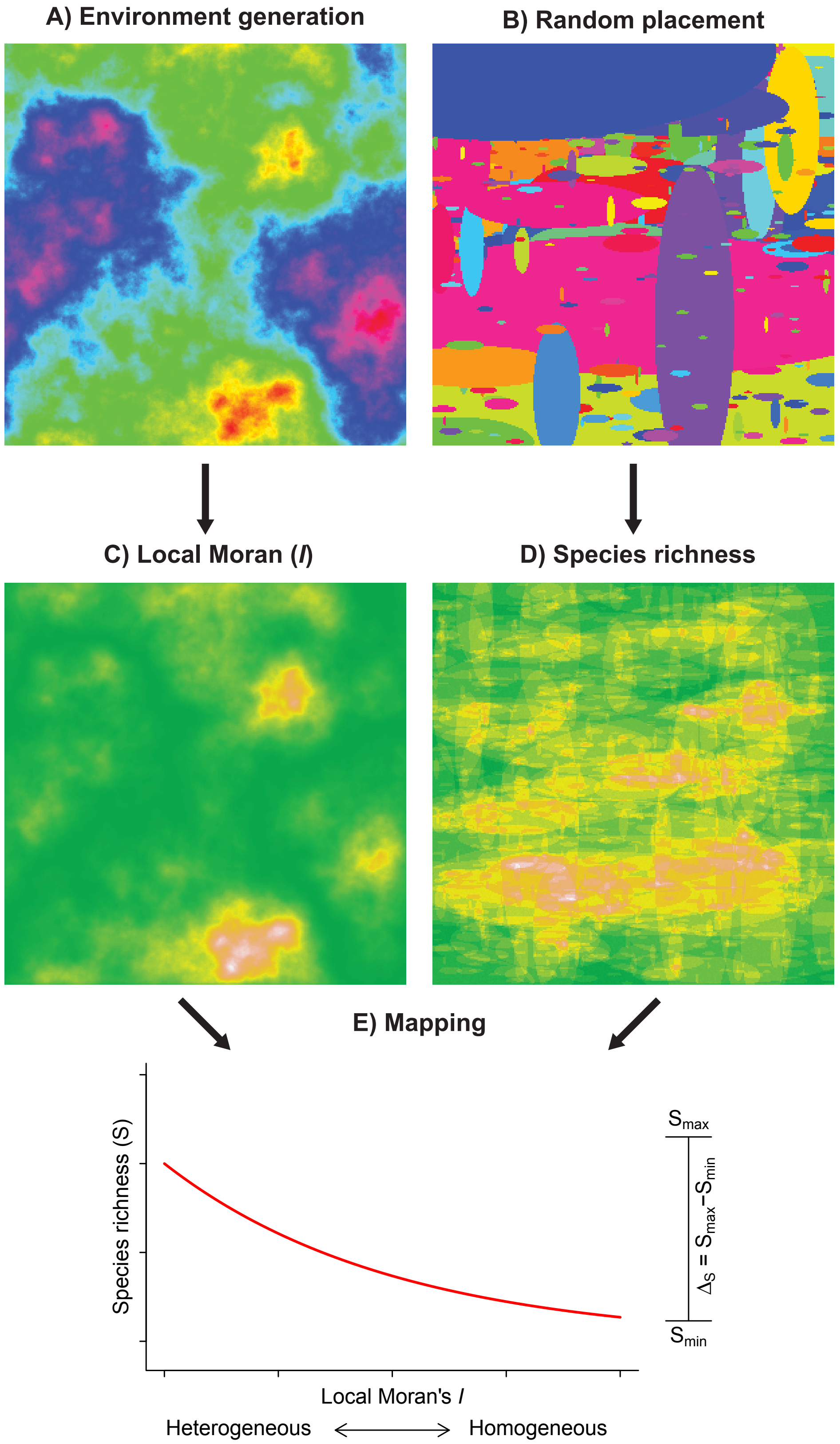

Figure 2: Framework of the random placement model of community assembly used to determine the relationship between fish species richness (S) and habitat heterogeneity in absence of any particular habitat selection mechanisms. Both environmental scores (**A**) and the regional distribution range of species (**B**) were generated independently and parameterized using observed data. Habitat heterogeneity (**C**) and species richness (**D**), the two resulting model components, were superimposed such that each Moran's I value on the grid was associated to a value of species richness (**E**). S_{min} and S_{max} represent the range spanned by a fitted GLM negative binomial regression (red curve). To simulate possible artifacts due to unsampled fish (false 0), we added a veil effect threshold (dashed horizontal red line) to the data generated by the model. A total of 10 000 simulation have been produced.

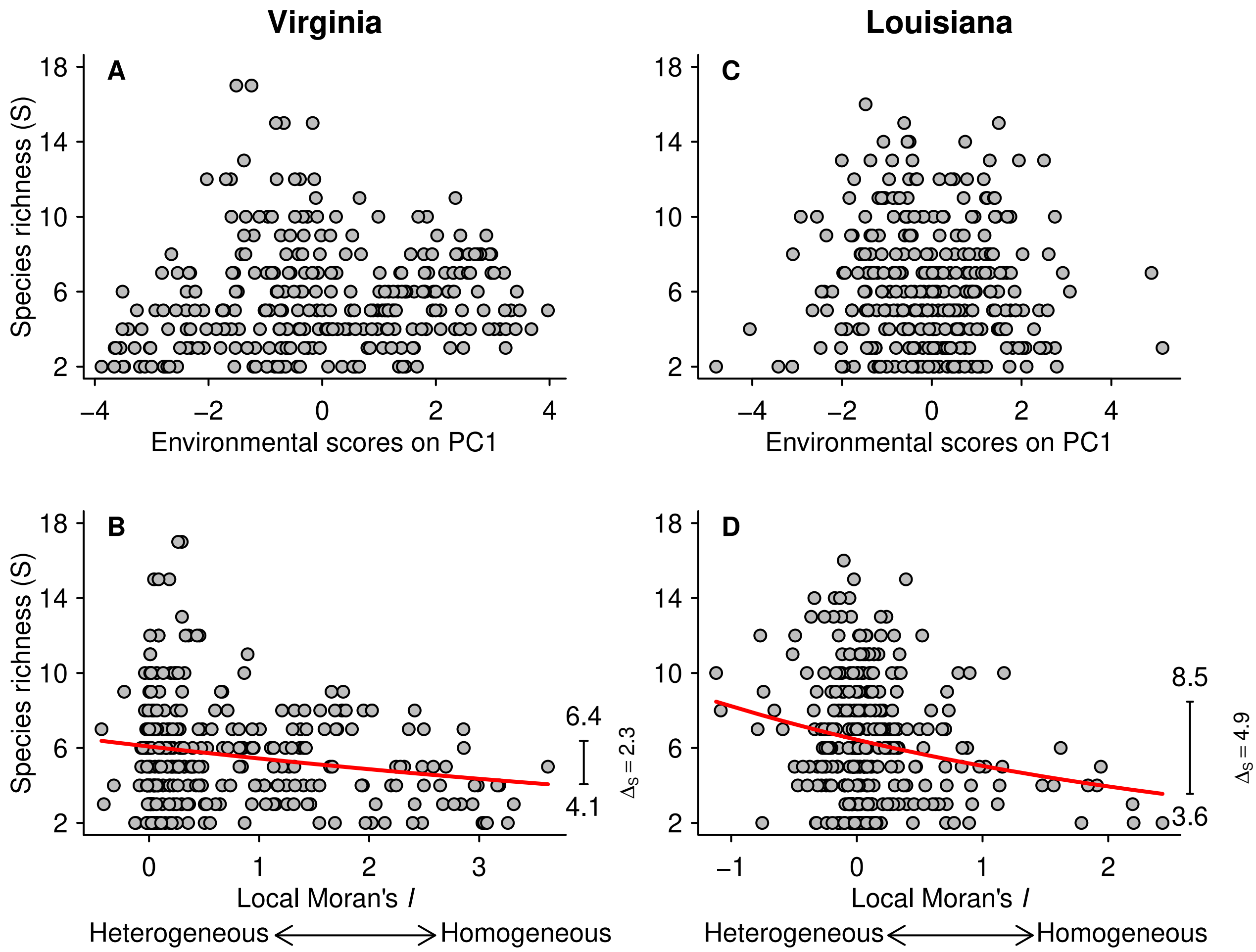

Figure 3: Relationships between species richness (S) and PCA scores for the first axis (panels **A** and **C**) and local Moran's I (panels **B** and **D**) for the Virginia and Louisiana biogeographic provinces. The red lines represent the fitted GLM negative binomial regressions between local Moran's I and S (Virginian $p < 0.001$, Louisianian $p < 0.001$). The right-margin insets in panels **B** and **D** show the amplitude of species richness (Δ_S) described by the regression curves.

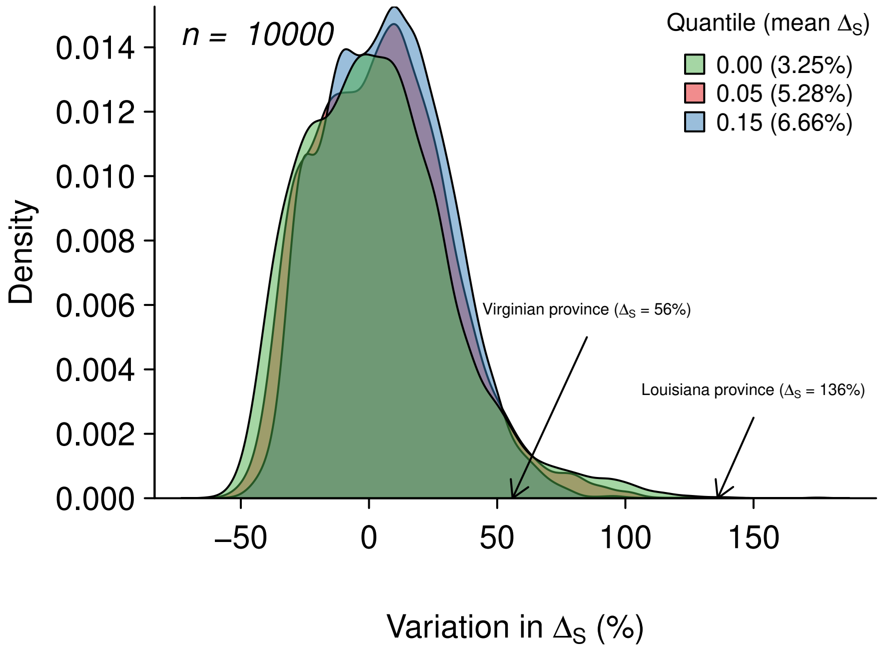

Figure 4: Results of 10000 simulations showing the influence of quantile cut (veil effect) on modeled species richness. The green, red and blue polygons represent the distribution of Δ_S under veil effects of percentiles 0%, 5% and 15%. The numbers in parentheses represent the mean of Δ_S for each veil simulation. The arrows indicate the Δ_S observed in both biogeographic provinces.

Table 1(on next page)

Main tables for the article

Using empirical and simulated data to study the influence of environmental heterogeneity on fish species richness in two biogeographic provinces

Tables

Philippe Massicotte[1], Raphaël Proulx[1], Gilbert Cabana[1], and Marco A. Rodríguez[1]

[1]Centre de Recherche sur les Interactions Bassins Versants-Écosystèmes aquatiques (RIVE). Université du Québec à Trois-Rivières, Trois-Rivières, Canada

October 20, 2014

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.567v1 | CC-BY 4.0 Open Access | rec: 30 Oct 2014, publ: 30 Oct 2014

Table 1: Loadings and summary statistics for environmental variables. The first three principal components generated from environmental variables were retained based on Kaiser's criterion. These components explained 75% of the total environmental variability in both biogeographic provinces.

	Virginian					**Louisianan**				
Variable	Loadings Comp. 1	Loadings Comp. 2	Loadings Comp. 3	Mean	Std. Dev.	Loadings Comp. 1	Loadings Comp. 2	Loadings Comp. 3	Mean	Std. Dev.
Water density (σ_t)	-0.49	0.02	0.12	9.08	8.68					
Dissolved oxygen (mgL^{-1})	-0.10	-0.69	0.03	6.90	1.25	-0.42	0.55	-0.10	6.89	1.33
Fluorescence	0.28	-0.34	0.42	11.82	7.70					
PAR ($mEm^{-2}s^{-1}$)	-0.05	-0.27	-0.85	545.76	464.29	-0.51	-0.41	-0.10	813.25	477.61
pH	-0.28	-0.53	0.16	7.93	0.48	-0.40	0.47	0.41	8.00	0.46
Salinity (ppt)	-0.49	-0.00	0.11	16.18	11.05	-0.06	-0.14	0.84	13.47	10.70
Temperature (°C)	0.39	-0.21	-0.16	25.40	2.46	-0.50	0.02	-0.32	29.77	1.41
Transmissivity (%)	-0.44	0.10	-0.14	53.37	23.19	-0.39	-0.54	0.11	63.97	16.12

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.567v1 | CC-BY 4.0 Open Access | rec: 30 Oct 2014, publ: 30 Oct 2014

Table 2: The probabilities of observing Δ_S greater or equal than 56% (Virginia) or Louisana (136%) due to sampling effect (i.e. random) under different scenarios of veil effects (0%, 5%, 15%). See Methods and Fig. 4 for detailed information.

	Veil at 0%	Veil at 5%	Veil at 15%
Virginia (56%)	4.68	3.70	2.12
Louisiana (136%)	0.05	0.01	0.00