Research

The 'rotiferologist' effect and other global correlates of species richness in monogonont rotifers

Diego Fontaneto, A. Márcia Barbosa, Hendrik Segers and Marco Pautasso

D. Fontaneto (diego.fontaneto@gmail.com), Imperial College London, Div. of Biology, Silwood Park Campus, SL5 7PY Ascot, Berkshire, UK. DF also at: Swedish Museum of Natural History, Dept of Invertebrate Zoology, Box 50007, SE-10405 Stockholm, Sweden. – A. M. Barbosa, 'Rui Nabeiro' Biodiversity Chair, CIBIO, Univ. of Évora, PT- 7004-516 Évora, Portugal and Imperial College London, Div. of Biology, Silwood Park Campus, SL5 7PY Ascot, Berkshire, UK. – H. Segers, Belgian Biodiversity Platform, Freshwater Laboratory, Royal Belgian Inst. of Natural Sciences, Vautierstraat 29, BE-1000 Brussels, Belgium. – M. Pautasso, Imperial College London, Div. of Biology, Silwood Park Campus, SL5 7PY Ascot, Berkshire, UK.

Global biodiversity patterns are often driven by different environmental variables at different scales. However, it is still controversial whether there are general trends, whether similar processes are responsible for similar patterns, and/or whether confounding effects such as sampling bias can produce misleading results. Our aim is twofold: 1) assessing the global correlates of diversity in a group of microscopic animals little analysed so far, and 2) inferring the influence of sampling intensity on biodiversity analyses. As a case study, we choose rotifers, because of their high potential for dispersal across the globe. We assembled and analysed a new worldwide dataset of records of monogonont rotifers, a group of microscopic aquatic animals, from 1960 to 1992. Using spatially explicit models, we assessed whether the diversity patterns conformed to those commonly obtained for larger organisms, and whether they still held true after controlling for sampling intensity, variations in area, and spatial structure in the data. Our results are in part analogous to those commonly obtained for macroorganisms (habitat heterogeneity and precipitation emerge as the main global correlates), but show some divergence (potential absence of a latitudinal gradient and of a large-scale correlation with human population). Moreover, the effect of sampling effort is remarkable, accounting for >50% of the variability; this strong effect may mask other patterns such as latitudinal gradients. Our study points out that sampling bias should be carefully considered when drawing conclusions from large-scale analyses, and calls for further faunistic work on microorganisms in all regions of the world to better understand the generality of the processes driving global patterns in biodiversity.

Species richness is known to vary in space and to be correlated with environmental, geographical and historical properties of the spatial units under analysis. Understanding the causes and consequences of such spatial variation in species richness is a major research focus of biogeography, macroecology and conservation (Willig et al. 2003). Modelling analyses and between-group comparisons are commonly performed to infer mechanisms in global patterns of distribution (Gotelli et al. 2009); this can be done for taxa with reliable information on their global distribution, such as vertebrates, plants and some insects (Field et al. 2009). Unfortunately, the distribution patterns of most other organisms are still undocumented, thus limiting the generality of the evidence currently available to understand the processes potentially explaining the patterns.

Microscopic organisms can provide further, independent evidence for or against the global correlates of species richness that are considered important for better-studied groups such as plants, vertebrates and insects. These variables include environmental productivity, habitat and topographic heterogeneity, human influences and historical stability (Schuldt and Assman 2009). Biodiversity patterns of microorganisms

may differ from those of macroscopic organisms because of their high abundances, the existence of resting stages, and their capacity for frequent and long-distance dispersal (Fenchel and Finlay 2004, Fontaneto et al. 2006, Fontaneto 2011). These special features of microorganisms already prompted the question of whether commonly reported patterns, such as the decrease in species richness with increasing latitude, are common to all organisms (Hillebrand 2004, Maraun et al. 2007).

One problem with the analysis of patterns in microorganisms is that their diversity can be difficult to assess; if reliable data on the distribution of all existing species are not available, then biogeographical hypotheses cannot be tested globally. Indeed, macroecological analyses of protists or prokaryotes have so far been difficult or inconsistent because of a lack of reliable data (Dolan 2005). One way to overcome such a problem could be to perform DNA analyses of environmental samples (Fuhrman et al. 2008), but these analyses are often still unfeasible globally. Another approach would be to use a different model taxon, which 1) matches the microbial biological features of small size, high abundances, presence of resting stages, frequent and long-distance dispersal, but also 2) shows a much lower

overall diversity than protists and prokaryotes, and for which 3) faunistic data are already available. Rotifers fulfil all these requirements as a model organism for ecogeographical studies of microscopic diversity.

Rotifers are a group of common and widespread microscopic animals, usually much shorter than 1 mm, living in any wet habitat in all continents. Monogonont rotifers are the richest group of rotifers, with about 1500 species globally (Segers 2007). Despite such relatively low species diversity worldwide, monogonont rotifers can be very diverse locally, with >100 species in any water body at any latitude (Dumont and Segers 1996, Segers and De Smet 2008). They are short-lived and fast-reproducing animals, and produce dormant stages called resting eggs, which can desiccate, persist in sediments, and act as propagules for dispersal (Ricci 2001).

There are two main aims in the present analysis. First, we wished to assess the consistency of the influence of the correlates of global richness described so far for other organisms. For this purpose, we assembled and analysed a global dataset of records of monogonont rotifers. We tested the generality of the previously known patterns by analysing the relationships between rotifer species richness and a series of environmental variables at three different spatial resolutions. Second, we tried to assess the influence of sampling intensity in biasing the results. Whereas correlates of diversity are currently analysed in an increasing number of taxa, the effect of sampling bias is not often considered, and may produce erroneous results (Walther et al. 1995, Barbosa et al. 2010a, Boakes et al. 2010).

Methods

Species data

We compiled a database of >45 000 published records of identified rotifer species in the world from 1960 to 1992, reported in 1871 papers listed in De Ridder and Segers (1997) and in the references cited by De Ridder (1986, 1991, 1994) for the same time period. These reviews gathered all the published taxonomic, faunistic and ecological studies reporting monogonont rotifers. Records older than 1960 were not considered because of potential taxonomic problems with unstable nomenclature in the past, whereas records after 1992 have not yet been gathered in any formal review. Rotifer species described before 1960 and never found again and rotifer species described after 1992 are not included in our dataset. Thus, the number of listed species is lower than the 1500 species currently known. The electronic database produced from this literature review is provided as Supplementary material (Appendix 1), and is also made freely available on the website of the Freshwater Animal Diversity Assessment (FADA, Balian et al. 2008: < http:// fada.biodiversity.be/group/show/4 >). This is the first time that a database of worldwide distribution records is available for any group of microorganisms, and we hope that the database will be useful for further research.

Geographical units

Many studies have stressed the importance of investigating diversity gradients over a range of spatial scales, as patterns may vary with scale (Field et al. 2009). We defined geographical units for this study following the Biodiversity Information Standards (TDWG, formerly Taxonomic Database Working Group: < www.tdwg.org/ >). These units are defined at four nested levels of spatial resolution: TDWG level 1 is at the coarser scale and includes 9 continental areas (Africa, Antarctica, Asia-temperate, Asia-tropical, Australasia, Europe, Pacific, Northern America, and Southern America); level 2 includes 52 sub-continental regions with an average area of 2 825 000 km^2; level 3 includes 369 units with an average of 398 000 km^2, mostly corresponding to countries or large intra-national regions; level 4 has 609 units, including states for the largest countries, and represents the most detailed scale, with an average area of 241 000 km^2 per unit (Brummit 2001; Fig. 1). Due to the nature of the published records on which we based our study, no georeferenced data points or interpolations into grid cells could be obtained.

We followed the TDWG standards for defining the geographical units because they can be identified from our bibliographic search even in cases where the exact locality of origin of a species record is missing; moreover, digitised georeferenced maps of these geographical units, suitable for use in a Geographic Information System (GIS), are available (e.g. from the Royal Botanic Garden, Kew: < www.rbgkew. org.uk/gis/tdwg >), and there is increasing use of these geographical subdivisions for ecological analyses (Davis et al. 2009, Jones et al. 2009, Kisel et al. 2011).

Variables

Species richness for each geographical unit was used as the response variable. Explanatory variables considered in the models (Supplementary material Appendix 2) were area of each geographic unit, average latitude (calculated with the GIS), average altitude, altitude range (a proxy for topographic heterogeneity), average annual precipitation, precipitation seasonality (coefficient of variation of precipitation), average annual temperature, annual temperature range (Hijmans et al. 2005), habitat diversity (number of different land cover categories among water, evergreen needle-leaf forest, evergreen broadleaf forest, deciduous needle-leaf forest, deciduous broadleaf forest, mixed forest, woodland, wooded grassland, closed shrubland, open shrubland, grassland, cropland, bare ground, and urban/built land: Fritz et al. 2009), total urbanised area (Global Land Cover Facility: Hansen et al. 1998), and human population size (Oak Ridge National Laboratory 2008). Some of these variables, e.g. temperature, habitat diversity and urbanised area, are measures of terrestrial variability; nevertheless, they may also influence the continental aquatic habitats where rotifers live (Johnson and Host 2010, Cole et al. 2011). We also included the number of rotifer records and the number of published papers for each geographical unit as a measure of sampling intensity.

Variables were imported into GRASS (GRASS Development Team 2009) through the Quantum GIS (Quantum GIS Development Team 2009) interface. Their values per geographic unit (Supplementary material Appendix 2) were obtained using the r.statistics module of GRASS, except for area (obtained with the v.to.db module) and latitude (obtained with r.mapcalc followed by r.statistics).
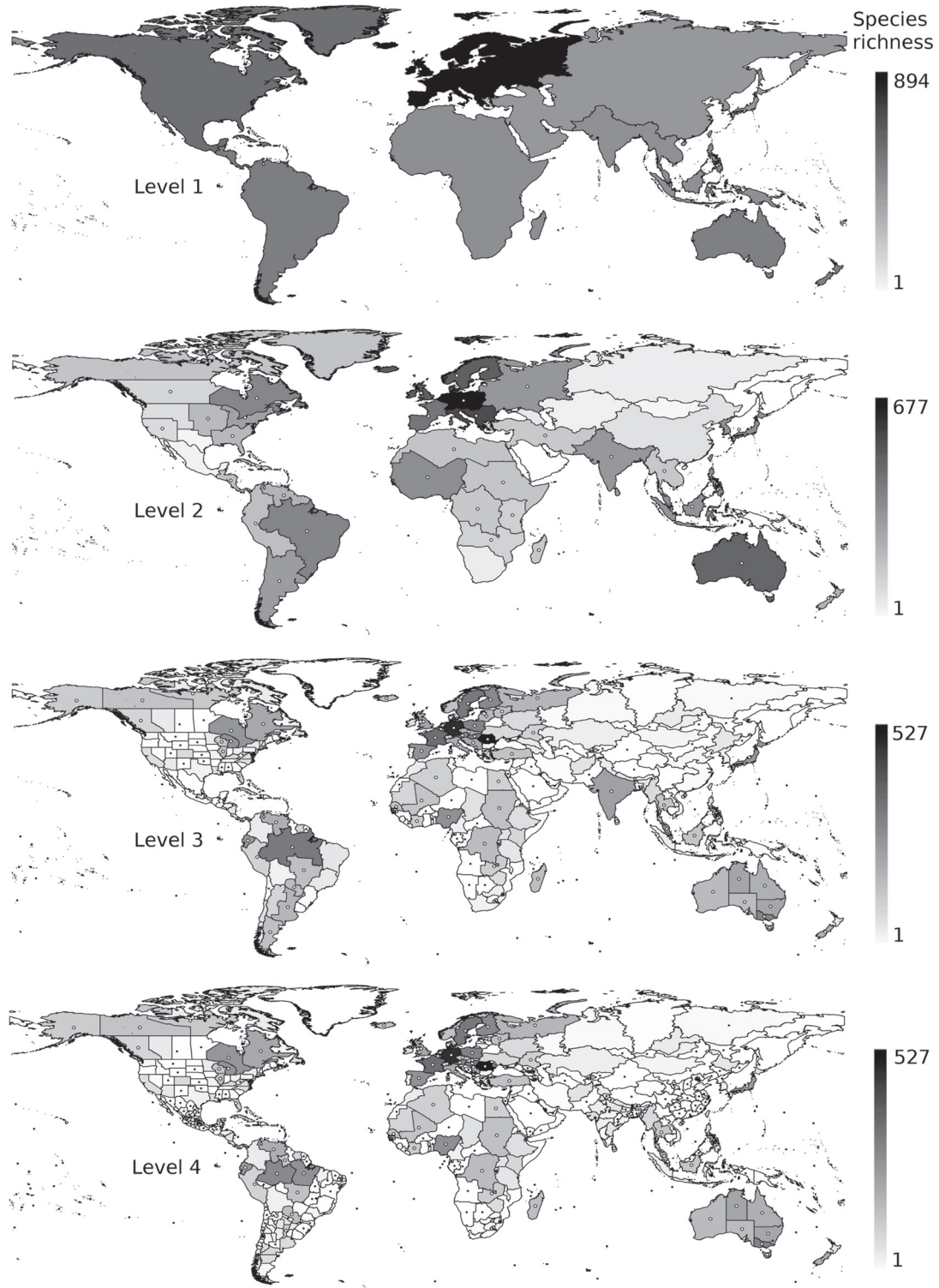
Figure 1. Global species richness of monogonont rotifers recorded at the four Biodiversity Information Standards (TDWG) resolution scale levels. Black dots mark geographical units with no rotifer records; these units were not included in the analyses; white circles mark the units with at least 100 recorded species and 5 papers, included in the analyses of the better sampled data sets. Level 1 was not analysed.

Statistical analyses

We ran various models to study the effect of the explanatory variables on rotifer diversity. After checking for multi-collinearity between the explanatory variables, we excluded the number of papers, which was highly correlated with the number of records (Pearson's $r > 0.8$ for all spatial levels). Other variables were also correlated, but always with $r < 0.8$, and we kept them in the following analyses. Latitude was included as an absolute value and also as a quadratic

component to detect possible hump-shaped responses. Because the explanatory variables were measured on different scales and with different units, we standardised them to optimise model fit, so that each variable had a mean of zero and standard deviation of one (Zuur et al. 2009). We used the logarithmic instead of the raw values for species richness and for the number of records to normalise the distribution of residuals and thus to conform to the assumptions of the statistical analyses. Although 70 rotifer species have been reported in Antarctica, this continent was excluded from the analyses because reliable climate variables were not available for this region (Hijmans et al. 2005).

We implemented linear models (LM), starting with a full model including all variables. We then performed formal model selection using a backward stepwise procedure based on Akaike's information criterion (AIC), aiming for model parsimony, i.e. the trade-off between accuracy and complexity. The best model has the minimum AIC value and contains the most information with the lowest number of variables. Thus, at each step we excluded from the model the variable that produced the highest decrease in AIC value. The final model retained the variables whose deletion would have caused an increase in AIC values (Zuur et al. 2007). We always checked that this final model was indeed the minimal adequate model: by computing analysis of variance tables for fitted model objects, we tested that the final model was not significantly worse than the full model as a description of the data (Crawley 2007).

For each model, we tested also for the presence of spatial structure. This was achieved by performing generalized least squares (GLS) models on the same variables, which allow to test the shape of various potential spatial correlation patterns, including exponential, Gaussian, linear, rational quadratic, and spherical correlation structure (Zuur et al. 2009). We used AIC to select the best model among the non-spatial and spatial ones, with all the different correlation structures for each set of models. Model fit through AIC indicated that spatial models were not significantly better than the non-spatial ones. Visual inspection of the shape of the variograms also confirmed that significant spatial autocorrelation was not present in any of the models. The results of the GLS models always matched qualitatively and quantitatively the results from the LM. Thus, we show the results of LM, which have no spatial structure and are simpler in their interpretation.

We performed these analyses for the three finest TDWG levels of spatial resolution, from level 2 to level 4. We did not perform statistical analyses for the coarsest TDWG level 1, as fitting models with only 8 data points can be meaningless; nevertheless, we show the recorded species richness for TDWG level 1 in Fig. 1. For each level, we included only geographical units where rotifers have been reported (i.e. recorded species richness > 0). Thus, the analysed sample sizes were 48, 216 and 294 for TDWG level 2 (92% of the units), level 3 (59%) and level 4 (48%), respectively (Table 1). As a test for differences between continents, and for differential effects in specific geographic areas, we performed additional separate analyses for each continental area at TDWG levels 3 and 4, using the same procedures outlined for the global analysis. We implemented all models first without including the number of records, and then including it as a measure of sampling bias, to highlight potential misinterpretations of the results.

Interactions between variables often result in overlapping effects even if variables are not correlated (for example number of records and precipitation), so that part of the variation explained by a variable in a multivariate model is shared with other variables in the model (Borcard et al. 1992, Real et al. 2003, Ribas et al. 2007). We thus performed variance partitioning to distinguish how much of the variance explained by a variable was due exclusively to that variable, and how much was shared with other variables. The regression of species richness on one variable provides the total amount of variance explained by that variable, including the amount shared with other variables. The amount of variance explained exclusively by one variable is the total variance explained by the multivariate model minus the total variance explained by the other variable(s) (Real et al. 2003).

The geographical units with the lowest number of records may be too under-sampled and, thus, limit the capacity of our models to detect any relevant effect. Analysing only a subset of better-sampled regions might help in detecting the environmental effects on rotifer species richness. Thus, we created reduced datasets to check the consistency of the results of the analyses performed on the complete datasets, by eliminating the geographical units whose number of records was too low. We chose a threshold of 100 species because this is the number of rotifer species that should be present in any water body at any latitude (Dumont and Segers 1996, Segers and De Smet 2008). We repeated the analyses (LM and GLS with variogram observation) on these reduced datasets for all TDWG levels. Our expectation was that the percentage of variance explained by environmental variables would be higher, with a lower influence of sampling bias.

To further reduce the effect of sampling bias, we obtained estimates of species richness from the reduced datasets using the Chao1 estimator from incidence data (Colwell and Coddington 1994). The datasets were thus further reduced by excluding units with < 5 published papers, in order to have reliable Chao1 estimates. We then performed all the models with Chao1 species richness estimates as the response variable.

To obtain a graphical representation of sampling bias, we drew cumulative curves describing the relationship between the total number of rotifer species and the number of recorded species per paper for each of the richest geographical units at TDWG levels 3 and 4 (corresponding to France, Germany, Hungary, Poland, Romania, and Sweden). Moreover, we extrapolated the theoretical total number of species for each of these units using the Chao1 estimator on species incidence (Colwell and Coddington 1994), and calculated the percentage of observed vs expected species.

All LM were performed with the statistic software R 2.10.0 (R Development Core Team 2009), GLS and variograms with package nlme 3.1-96 (Pinheiro et al. 2009), Chao1 estimates and cumulative curves with package vegan 1.15-4 (Oksanen et al. 2009).

Results

The species database

The compiled database contained 46 549 records of 1329 monogonont rotifer species (Supplementary material

Table 1. Results of the final linear models with all significant variables (not including the number of records as a measure of sampling bias), and percentage of variance explained by each variable. Between brackets the number of units analysed in each dataset.

	variable	estimate	standard error	p value	% variance
TDWG2 (48)	(intercept)	4.63	0.21	<0.0001	
	habitat diversity	0.82	0.21	0.0003	25
TDWG3 (216)	(intercept)	3.47	0.17	<0.0001	
	area	0.21	0.09	0.027	4
	altitude	−0.22	0.09	0.016	2
	habitat diversity	0.35	0.10	0.0006	3
	latitude	0.39	0.17	0.026	1
	temperature range	−0.20	0.09	0.034	2
TDWG4 (294)	(intercept)	2.94	0.22	<0.0001	
	area	0.29	0.08	0.0003	6
	altitude	−0.31	0.07	0.0001	3
	habitat diversity	0.25	0.08	0.003	4
	human population	0.23	0.08	0.004	1
	temperature	−0.30	0.14	0.03	1
	temperature seasonality	−0.44	0.12	0.0005	3
	latitude	1.50	0.52	0.004	1
	latitude (square term)	−0.67	0.26	0.01	2

Appendix 1). Species described before 1960 and never found again or described after 1992 are not included in this database. Some of the records could not be unambiguously located in a particular TDWG unit; therefore, slightly reduced datasets were used in the analyses, with 46 136 records for level 1 (99.1%), 45 668 for level 2 (98.1%), 43 869 for level 3 (94.2%), and 43 527 for level 4 (93.5%).

The richest geographical units were mostly in Europe at all TDWG levels (Fig. 1). Nevertheless, even in Europe, countries such as Bulgaria, Denmark, Greece, and Portugal had <60 recorded species. For all the richest geographical units on TDWG level 4, cumulative curves did not reach an asymptote (Fig. 2), suggesting that new species can still be found even in these areas. The highest expected number of rotifer species (Chao1 estimator) was obtained for Romania, with 612 species (from 527 observed species). The number of observed species in the richest geographical units ranged from 61% of the expected Chao1 number for Sweden to 87% for Germany.

For many areas in the world, a very low number of species or even no species at all were recorded (Fig. 1): at TDWG level 2, 5 units (10% of the ones with rotifer data) had <10 species; at TDWG level 3, 46 units (21%); at TDWG level 4, 79 (27%).
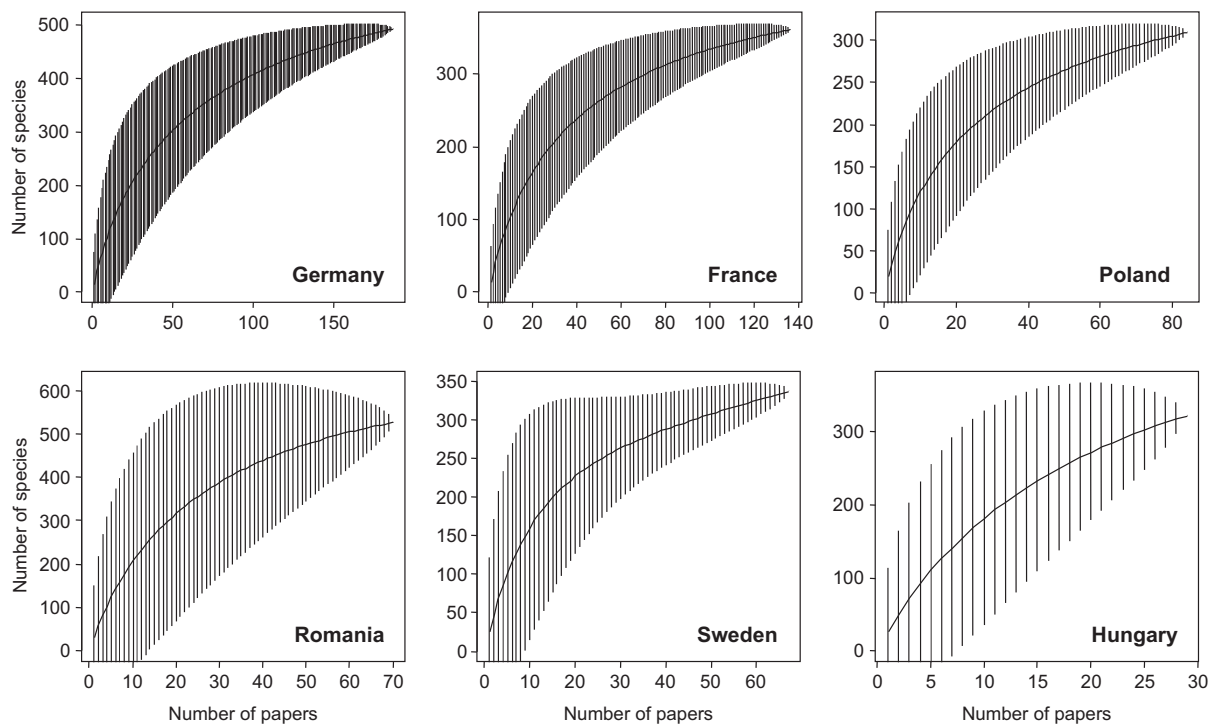

Figure 2. Cumulative curves and 95% confidence intervals of species richness of monogonont rotifers in the six richest geographical units at TDWG level 4.

Correlates of species richness

Without considering sampling bias, habitat diversity was positively related to species richness at TDWG level 2, explaining 25% of its variance. Many other variables appeared related to species richness at the finer scales of TDGW levels 3 and 4 (Table 1). However, they explained only a small proportion of the variance in species richness, totalling <21% (Table 1). When the number of records was included as a covariate in the models (Table 2), it stood out as the most important variable, explaining > 50% of the variance at all spatial scales. Habitat diversity was retained in the final models for TDWG level 2 and 4, even when including the number of records, which nevertheless reduced the variance explained by habitat diversity alone from 25 to 14% in TDWG level 2 and from 4 to 1% in TDWG level 4. This was because part of the variance explained by habitat diversity was shared with the number of records. Analysing TDWG levels 3 and 4 in different continents separately the number of records was the only significant variable, explaining between 55 and 85% of the variance in species richness (results not shown).

The number of records was consistently the most important variable influencing our knowledge of species richness of monogonont rotifers. This was unlikely driven by the large amount of units with few studies and few species reported: analysing only reduced datasets that excluded geographical units with <100 species, the number of records was still the only variable always retained (Table 3).

When trying to remove the effect of sampling bias by analysing Chao1 estimates of species richness for the best sampled geographical units, the number of records was retained for TDWG level 2, for which sampling bias was not removed. For TDWG levels 3 and 4, Chao1 estimators may be considered more reliable, as the number of records was not retained in these models; precipitation was the only variable retained in the final models, although it was not significant (Table 4).

Discussion

The most important variable influencing species richness of rotifers is an artefact: sampling intensity is obviously not driving species richness, but it is indeed conditioning our knowledge of species richness and distribution patterns. This corroborates previous suggestions pointing out that the distribution of rotifers seems to reflect the distribution of rotifer scientists more than that of rotifers themselves (Dumont 1983, Segers and De Smet 2008), in what we could call the 'rotiferologist' effect. Thus, even for a group of organisms like rotifers, which are not as diverse as prokaryotes or protists and for which many records (nearly 45 000) are now available, results from global analyses can be misleading if sampling bias is not considered. Faunistic data on rotifers are still far from being comprehensive. Nevertheless, with only a few hundred species, taxonomic knowledge and completeness should not be a limiting factor in enumerating species diversity. We cannot test whether taxonomic knowledge can be related to the number of papers, although we can suggest that it is not, given that the same European taxonomists worked also abroad, in order to find new species to describe in understudied areas. Sampling bias, measured by the number of published records per geographical unit, is thus a strong correlate of the apparent spatial patterns of rotifer diversity. This kind of bias is known to affect patterns of species richness observed also for larger organisms, at least birds, helminth parasites and ground beetles (Boulinier et al. 1998, Ribas et al. 2007, Barbosa et al. 2010a), even if with less influence. It is surprising that this kind of bias is rarely considered in such studies, as it can give distorted views of the spatial patterns and processes in biodiversity, even for really well-known groups of birds such as the Galliformes (Boakes et al. 2010).

The strong relationships between species richness and sampling intensity in rotifers could be also due to the fact that the data have been recorded within arbitrary political boundaries (TDWG levels), which are also the sampling unit for our analyses. It may be that these strong relationships are obscuring effects of other variables, which are not measured on the same sampling unit structure but rather are averaged across those units, meaning that their predictive power is likely weakened in this averaging. Unfortunately, no interpolation into grid cells could be performed with the available rotifer dataset. Nevertheless, two environmental variables, habitat diversity and precipitation, were retained in different models including a measure of sampling bias. Habitat diversity is important at a coarse scale (TDWG level 2) and at the finest scale (TDWG level 4) when including the number of records. Precipitation is significant at the finest scales (TDWG levels 3 and 4) when considering estimates of species richness to try to correct for sampling bias. The latter variable adds no more than 1% of the total explained variance of each model; whereas habitat diversity may explain up to 25% of the total variance at TDWG level 2 (Table 1, 2).

Table 2. Results of the final linear models with all significant variables including number of records as a measure of sampling bias, and percentage of variance explained (between brackets is the total variance explained by each variable, including the percentage shared with the number of records, obtained by variance partitioning).

	variable	estimate	standard error	p value	% variance
TDWG2	(intercept)	5.07	0.15	<0.0001	
	records	1.58	0.21	<0.0001	52
	habitat diversity	0.63	0.14	<0.0001	14 (25)
TDWG3	(intercept)	4.24	0.06	<0.0001	
	records	1.83	0.08	<0.0001	69
TDWG4	(intercept)	3.95	0.04	<0.0001	
	records	2.01	0.08	<0.0001	68
	habitat diversity	0.10	0.04	0.02	1 (4)

Table 3. Results of the final linear models on the reduced datasets (at least 100 species per geographical units) and percentage of variance explained. Between brackets the number of units retained in each reduced dataset.

	variable	estimate	standard error	p value	% variance
TDWG2 (32)	(intercept)	5.57	0.005	<0.0001	
	records	0.81	0.06	<0.0001	81
TDWG3 (67)	(intercept)	5.37	0.02	<0.0001	
	records	0.43	0.03	<0.0001	77
TDWG4 (63)	(intercept)	5.38	0.02	<0.0001	
	records	0.46	0.03	<0.0001	76

Given the many sources of noise in biogeographical analyses, it is not uncommon to find low proportions of variance in species richness explained by environmental variables, even for well-studied macroorganisms such as plants and birds (Pautasso and Gaston 2005, Pautasso and Parmentier 2007, Lilleskov et al. 2008). Thus, these two variables may be considered very important for rotifers, as they stand out notwithstanding the strong confounding effects of sampling bias.

The two variables, habitat diversity and precipitation, cover both groups of correlates of species richness suggested for other taxa, namely habitat heterogeneity and environmental productivity (Field et al. 2009). Interestingly, these variables have different roles in explaining rotifer species richness at different spatial resolution scales, confirming the potential scale-dependence in patterns and determinants of diversity (Crawley and Harral 2001, Pecher et al. 2010). As is common for other organisms (Kerr et al. 2001, Hortal et al. 2009), more species of rotifers are present in areas with a higher diversity of habitats; thus, many rotifer species may be ecologically restricted to waters in specific habitats. Species assemblages of rotifers typical of water bodies with different levels of eutrophication were proposed as a system to monitor environmental degradation, due to their response to changes in habitat quality (Sladecek 1983). Rotifers partition available space and resources even at the very local scale of the single host in epibiont species assemblages (Fontaneto and Ambrosini 2010); thus, the positive relationship between rotifer species richness and terrestrial habitat diversity is biologically meaningful, and supported also at a more local scale (Obertegger et al. 2010). The influence of the heterogeneity of the terrestrial habitat on species richness of aquatic organisms confirms the hypothesis that diversity in the freshwater biota is affected by the terrestrial habitats surrounding the water bodies (Leroux and Loreau 2008, Johnson and Host 2010, Cole et al. 2011).

At the finer scale, precipitation explains a small part of the variation in rotifer species richness, when avoiding sampling bias by using estimates of species richness. Since these animals are aquatic, water availability may be directly linked to their diversity, or may increase the abundance and diversity of aquatic habitats within a given area.

One variable that was surprisingly not retained in any of the models including or correcting for sampling bias is latitude. Latitude is only a description of position, but it often correlates with ecological variables that are biologically relevant and is a commonly used variable in analyses of species' geographical patterns (Hawkins and Diniz-Filho 2004). Latitude was never significant for our rotifer dataset, except when not including sampling bias (Table 1). The latitudinal gradient in species richness is documented as a common and widespread pattern (Gaston 2007); even in rotifers, latitudinal gradients in species richness have been suggested in the few well-studied rotifer taxa for which detailed biogeographical analyses are possible (*Brachionus*: Pejler 1977; Lecanidae: Segers 1996; Trichocercidae: Segers 2003). Also the species composition of rotifer communities is known to change with latitude (Green 1972). On the other hand, hotspots of rotifer diversity and endemicity exist in Australia, China, North America and tropical South America, whereas endemicity appears to be low in Africa and in the Indian subcontinent (Segers and De Smet 2008), without any clear relationship to environmental or latitudinal gradients. The lack of a relationship between species richness and latitude in our study may still be an artefact of sampling intensity. Indeed, absolute latitude and number of records are positively correlated at TDWG level 3 ($r = 0.14$, $p = 0.04$) and at TDWG level 4 ($r = 0.11$, $p = 0.06$); this relationship could have potentially masked a possible latitudinal decrease in true species richness. Nevertheless, there is further evidence that richness of rotifers and of other small soil invertebrates does not increase towards the equator (Maraun et al. 2007, Segers 2008) and that polar areas such as Svalbard islands are at least as rich as temperate ones for bdelloid rotifers (Kaya et al. 2010). Only a more complete dataset might enable us to detect (or rule

Table 4. Results of the final linear models on the reduced datasets (at least 100 species and 5 papers per geographical units) and percentage of variance explained, using Chao1 estimates of species richness as the response variable. Between brackets the number of units retained in each reduced dataset.

	variable	estimate	standard error	p value	% variance
TDWG1 (7)	none				
TDWG2 (32)	(intercept)	437.52	17.30	<0.0001	
	records	178.46	22.12	<0.0001	68
TDWG3 (59)	(intercept)	348.61	19.01	<0.0001	
	precipitation	14.65	19.18	0.45	1
TDWG4 (58)	(intercept)	342.71	19.70	<0.0001	
	precipitation	16.64	19.87	0.41	1

out) the presence of such a gradient. In particular, detailed information on rotifer species richness in (sub)tropical regions would be instrumental to refine our analysis.

Common environmental correlates of species diversity, such as altitude and area, were not significant in our models, even when excluding poorly sampled geographical units from the analysis. No relationship between rotifer species richness and human population or urbanized area was observed either, indicating that records of rotifers and their species richness may not be related to how many people live in the area. This is in contrast with previous studies reporting significant associations between human population and species richness for other taxonomic groups (Luck 2007, Barbosa et al. 2010a, b).

Our analysis of global patterns in species richness of rotifers provides compelling evidence that nearly all potential explanatory variables (except for habitat diversity and precipitation) are wiped out when including sampling intensity in the models. Thus, the main general conclusion is that macroecological analyses should explicitly consider sampling intensity as a source of error. There is evidence that sampling bias may be present in many analyses at different spatial scales, even in well-known groups such as birds (Boulinier et al. 1998, Boakes et al. 2010) and ground beetles (Barbosa et al. 2010a). If not considered, sampling bias may result in artefactual correlations with explanatory variables (Barbosa et al. 2010b). Of course, for large organisms such bias is not necessarily present (Pautasso and McKinney 2007), but this cannot be assumed a priori. Thus, for all organisms, even if not massively undersampled as rotifers, the effect of sampling bias needs to be taken into account, to avoid distorted views on biodiversity patterns (Boakes et al. 2010). Another conclusion is that faunistic data for rotifers and other inconspicuous organisms are still far from being able to document their diversity patterns in a reliable way. Faunistic studies on rotifers are known to be challenging because remarkable biodiversity levels can be found even in a single pond (Dumont and Segers 1996, Segers and Sanoamuoang 2007). Moreover, species richness is known to be underestimated using single sampling or short-term strategies (Walsh et al. 2007); nevertheless, reliable taxonomic knowledge for a few hundred species is accessible. Hence, we call for more long-term studies on rotifers, especially in understudied areas such as northern Asia, southern Africa and the central part of North America (Fig. 1). This first global analysis on non-model organisms suggests new avenues of research to challenge previously known patterns such as latitudinal gradients, biodiversity hotspots, and species-area relationships. Such studies on rotifers and on other microscopic organisms worldwide would be pivotal to assess the generality of the processes driving global patterns in biodiversity.

References

Balian, E. V. et al. 2008. The freshwater animal diversity assessment: an overview of the results. – Hydrobiologia 595: 627–637.

Barbosa, A. M. et al. 2010a. Is the human population a large-scale indicator of the species richness of ground beetles? – Anim. Conserv. 13: 432–441.

Barbosa, A. M. et al. 2010b. Positive regional species–people correlations: a sampling artefact or a key issue for sustainable development? – Anim. Conserv. 13: 446–447.

Boakes, E. H. et al. 2010. Distorted views of biodiversity: spatial and temporal bias in species occurrence data. – PLoS Biol. 8: e1000385.

Borcard, D. et al. 1992. Partialling out the spatial component of ecological variation. – Ecology 73: 1045–1055.

Boulinier, T. et al. 1998. Estimating species richness: the importance of heterogeneity in species detectability. – Ecology 79: 1018–1028.

Brummitt, R. K. 2001. World geographical scheme for recording plant distributions, 2nd ed. – Hunt Inst. for Botanical Documentation, Carnegie Mellon Univ., Pittsburgh.

Cole, J. J. et al. 2011. Strong evidence for terrestrial support of zooplankton in small lakes based on stable isotopes of carbon, nitrogen, and hydrogen. – Proc. Natl Acad. Sci. USA doi: 10.1073/pnas.1012807108

Colwell, R. K. and Coddington, J. A. 1994. Estimating terrestrial biodiversity through extrapolation. – Phil. Trans. R. Soc. B 345: 101–118.

Crawley, M. J. 2007. The R book. – Wiley.

Crawley, M. J. and Harral, J. E. 2001. Scale dependence in plant biodiversity. – Science 291: 864–868.

Davis, A. P. et al. 2009. A global assessment of distribution, diversity, endemism, and taxonomic effort in the Rubiaceae. – Ann. Mo. Bot. Gard. 96: 68–78.

De Ridder, M. 1986. Annotated checklist of non-marine Rotifera from African inland waters. – Koninklijk Museum voor Miden Afrika, Tervuren, Zoologische Documentatie 21.

De Ridder, M. 1991. Additions to the 'Annotated checklist of non-marine rotifers from African inland waters'. – Rev. d'Hydrobiol. Trop. 24: 25–46.

De Ridder, M. 1994. Additions II to the 'Annotated checklist of non-marine rotifers from African inland waters'. – Biol. Jaarboek Dodonaea 61: 99–153.

De Ridder, M. and Segers, H. 1997. Rotifera Monogononta in six zoogeographical regions after publications between 1960 and 1992. – Studiedocumenten van het Koninklijk Belgisch Inst. voor Natuurwetenschappen 88.

Dolan, J. R. 2005. An introduction to the biogeography of aquatic microbes. – Aquat. Microb. Ecol. 41: 39–48.

Dumont, H. J. 1983. Biogeography of rotifers. – Hydrobiologia 104: 19–30.

Dumont, H. and Segers, H. 1996. Estimating lacustrine zooplankton species richness and complementarity. – Hydrobiologia 341: 125–132.

Fenchel, T. and Finlay, B. J. 2004. The ubiquity of small species: patterns of local and global diversity. – BioScience 54: 777–784.

Field, R. et al. 2009. Spatial species-richness gradients across scales: a meta-analysis. – J. Biogeogr. 36: 132–147.

Fontaneto, D. (ed.) 2011. Biogeography of microorganisms: is everything small everywhere? – Cambridge Univ. Press.

Fontaneto, D. and Ambrosini, R. 2010. Spatial niche partitioning in epibiont rotifers on the waterlouse *Asellus aquaticus*. – Limnol. Oceanogr. 55: 1327–1337.

Fontaneto, D. et al. 2006. Patterns of diversity in microscopic animals: are they comparable to those in protists or in larger animals? – Global Ecol. Biogeogr. 15: 153–162.

Fritz, S. A. et al. 2009. Geographical variation in predictors of mammalian extinction risk: big is bad, but only in the tropics. – Ecol. Lett. 12: 538–549.

Fuhrman, J. A. et al. 2008. A latitudinal diversity gradient in planktonic marine bacteria. – Proc. Natl Acad. Sci. USA 105: 7774–7778.

Gaston, K. J. 2007. Latitudinal gradient in species richness. – Curr. Biol. 17: R574.

Gotelli, N. J. et al. 2009. Patterns and causes of species richness: a general simulation model for macroecology. – Ecol. Lett. 12: 873–886.

GRASS Development Team 2009. Geographic resources analysis support system (GRASS) software. – Open Source Geospatial Foundation Project, <http://grass.osgeo.org>.

Green, J. J. 1972. Latitudinal variation in associations of planktonic Rotifera. – J. Zool. 167: 31–39.

Hansen, M. et al. 1998. UMD Global Land Cover Classification, 1 km, 1981–1994. – Dept of Geography, Univ. of Maryland, College Park, MD, <www.landcover.org/data/landcover>.

Hawkins, B. A. and Diniz-Filho, J. A. F. 2004. 'Latitude' and geographic patterns in species richness. – Ecography 27: 268–272.

Hijmans, R. J. et al. 2005. Very high resolution interpolated climate surfaces for global land areas. – Int. J. Climatol. 25: 1965–1978.

Hillebrand, H. 2004. On the generality of the latitudinal diversity gradient. – Am. Nat. 163: 192–211.

Hortal, J. et al. 2009. Island species richness increases with habitat diversity. – Am. Nat. 174: E205–E217.

Johnson, L. B. and Host, G. E. 2010. Recent developments in landscape approaches for the study of aquatic ecosystems. – J. N. Am. Benthol. Soc. 29: 41–66.

Jones, O. R. et al. 2009. Using taxonomic revision data to estimate the geographic and taxonomic distribution of undescribed species richness in the Braconidae (Hymenoptera: Ichneumonoidea). – Insect Conserv. Divers. 2: 204–212.

Kaya, M. et al. 2010. Survey of moss-dwelling bdelloid rotifers from middle Arctic Spitsbergen (Svalbard). – Polar Biol. 33: 833–842.

Kerr, J. T. et al. 2001. Remotely sensed habitat diversity predicts butterfly species richness and community similarity in Canada. – Proc. Natl Acad. Sci. USA 98: 11365–11370.

Kisel, Y. et al. 2011. How diversification rates and diversity limits combine to create large-scale species–area relationships. – Phil. Trans. R. Soc. B, in press.

Leroux, S. J. and Loreau, M. 2008. Subsidy hypothesis and strength of trophic cascades across ecosystems. – Ecol. Lett. 11: 1147–1156.

Lilleskov, E. A. et al. 2008. Divergent biogeography of native and introduced soil macroinvertebrates in North America north of Mexico. – Divers. Distrib. 14: 893–904.

Luck, G. W. 2007. A review of the relationships between human population density and biodiversity. – Biol. Rev. 82: 607–645.

Maraun, M. et al. 2007. Awesome or ordinary? Global diversity patterns of oribatid mites. – Ecography 30: 209–216.

Oak Ridge National Laboratory 2008. LandScan Global Population data. – <www.ornl.gov/sci/landscan/index.html>.

Obertegger, U. et al. 2010. Rotifer species richness along an altitudinal gradient in the Alps. – Global Ecol. Biogeogr. 19: 895–904.

Oksanen, J. et al. 2009. vegan: community ecology package. – R package ver. 1.15-4, <http://CRAN.R-project.org/package=vegan>.

Pautasso, M. and Gaston, K. J. 2005. Resources and global avian assemblage structure in forests. – Ecol. Lett. 8: 282–289.

Pautasso, M. and McKinney, M. L. 2007. The botanist effect revisited: plant species richness, county area and human population size in the US. – Conserv. Biol. 21: 1333–1340.

Pautasso, M. and Parmentier, I. 2007. Are the living collections of the world's botanical gardens following species-richness patterns observed in natural ecosystems? – Bot. Helv. 117: 15–28.

Pecher, C. et al. 2010. Scale-dependence of the correlation between human population and the species richness of stream macroinvertebrates. – Basic Appl. Ecol. 11: 272–280.

Pejler, B. 1977. On the global distribution of the family Brachionidae (Rotatoria). – Arch. Hydrobiol. (Suppl.) 53: 255–306.

Pinheiro, J. et al. 2009. nlme: linear and nonlinear mixed effects models. – R package ver. 3.1-96, <http://cran.r-project.org/web/packages/nlme/nlme.pdf>.

Quantum GIS Development Team 2009. Quantum GIS geographic information system. – Open Source Geospatial Foundation Project, <http://qgis.osgeo.org>.

Real, R. et al. 2003. Relative importance of environment, human activity and spatial situation in determining the distribution of terrestrial mammal diversity in Argentina. – J. Biogeogr. 30: 939–947.

Ribas, A. et al. 2007. Geographical patterns of the species richness of helminth parasites of moles (*Talpa* spp.) in Spain: separating the effect of sampling effort from those of other conditioning factors. – Vie Milieu 57: 45–52.

Ricci, C. 2001. Dormancy patterns in rotifers. – Hydrobiologia 446: 1–11.

Schuldt, A. and Assman, T. 2009. Environmental and historical effects on richness and endemism patterns of carabid beetles in the western Palaearctic. – Ecography 32: 705–714.

Segers, H. 1996. The biogeography of littoral *Lecane* Rotifera. – Hydrobiologia 323: 169–197.

Segers, H. 2003. A biogeographical analysis of rotifers of the genus *Trichocerca* Lamarck. 1801 (Trichocercidae, Monogononta, Rotifera), with notes on taxonomy. – Hydrobiologia 500: 103–114.

Segers, H. 2007. A global checklist of the rotifers (Phylum Rotifera). – Zootaxa 1564: 1–104.

Segers, H. 2008. Global diversity of rotifers (Rotifera) in freshwater. – Hydrobiologia 595: 49–59.

Segers, H. and Sanoamuang, L.-O. 2007. Note on a highly diverse rotifer assemblage (Rotifera: Monogononta) in a Laotian rice paddy and adjacent pond. – Int. Rev. Hydrobiol. 92: 640–646.

Segers, H. and De Smet, W. H. 2008. Diversity and endemism in Rotifera: a review, and *Keratella* Bory de St Vincent. – Biodivers. Conserv. 17: 303–316.

Sladecek, V. 1983. Rotifers as indicators of water quality. – Hydrobiologia 100: 169–201.

Walsh, E. J. et al. 2007. How well do single samples reflect rotifer species diversity? A test based on interannual variation of rotifer communities in Big Bend National Park (Texas, USA). – Hydrobiologia 593: 39–47.

Walther, B. A. et al. 1995. Sampling effort and parasite species richness. – Parasitol. Today 11: 306–310.

Willig, M. R. et al. 2003. Latitudinal gradients of biodiversity: pattern, process, scale, and synthesis. – Annu. Rev. Ecol. Evol. Syst. 34: 273–309.

Zuur, A. F. et al. 2007. Analyzing ecological data. – Springer.

Zuur, A. F. et al. 2009. Mixed effects models and extensions in ecology with R. – Springer.

Supplementary material (Appendix E6850 at <www.oikosoffice.lu.se/appendix>). Appendix 1–2.