

## Review of Article

Title: Geodiversity primarily shapes large-scale limnology and aquatic species distribution in the northern Neotropics

Author(s): Laura Anahí Macario-González et al.

MS No.: bg-2021-298

### General comments

This is an interesting study and this manuscript eventually will be worthy of publication. The authors are to be commended for their effort because this is a rather large data set from a poorly-studied part of the world.

Because the authors are evidently not native speakers of English, there are considerable numbers of grammatical mistakes and poorly-phrased passages, most of which are detailed below with suggestions for improvement.

In several places the text is over-generalized, for example Line 35 which refers to “aquatic biological composition” – in fact, this study is only about ostracods.

One area that requires clarification is the enumeration of ostracods (Line 204). Apparently the investigators counted ostracods, but the details of counting aren't presented. How many per sample? It's also not clear what they mean by “identified three adults” passage. Analysis seems to be based on presence/absence, but if count data are available for each sample, then ordination can be based on square-root-transformed percentage data.

A more serious shortcoming is data handling. Apparently the investigators did not test each variable for normality. Skewed data should be transformed to produce a more normal distribution prior to ordination. For a clear methodology I recommend the most-recent edition of Tabachnick and Fidell's book *Using multivariate statistics*.

The authors should be careful in discussing diversity, which is not the same as richness. Also, such comparisons are difficult to interpret because collection size varies among the studies that they cite. Using ‘alpha diversity’ and ‘beta diversity’ would help.

I suspect that the authors are correct, that geology (local bedrock, karst vs volcanic) is the main influence on ostracod distribution. It's not clear why they use the term ‘geodiversity’ when bedrock geology alone seems to be the main driver. Data on ‘geodiversity’, i.e. local *diversity* of geology (geology, geomorphology, hydrology) aren't presented for each of the lakes.

Another driver seems to be understated, namely precipitation rates. Region YG is low rainfall, Region GSHN is generally higher rainfall. This correlates with karst/low elevation vs volcanic/highlands, but isn't rejected as a main driver. In order to conclude firmly that it's geology rather than rainfall, the authors should add another variable to their data set (i.e. mean annual rainfall at each lake).

### Specific comments

- Line 1 the title is misleading – it indicates that shaping limnology is the primary result of geodiversity. Also, the article is about ostracods, not other aquatic species. Suggest “Geodiversity is a major factor in limnology and ostracod distribution in the northern Neotropics”.
- Line 34 seems unsupported. I see no evidence of linear or continuous influence (if present, perhaps a figure could show this).
- Line 34 better words are available for the last sentence: “lake characteristics” not “limnological structure”, “water chemistry” not “geochemical properties”, and “ostracodes” not “aquatic biological composition”.

- Line 120 not sure what is meant here. Maybe convert  $\text{mm yr}^{-1}$  to  $\text{mm month}^{-1}$ , because the sentence is about the increased rainfall in the wet season, not in an entire year.
- Line 152 Were there multiple samples from the littoral zone of each lake? from the deepest part? Perhaps can indicate “Up to X samples were collected from littoral zones and up to Y samples from the deepest part”.
- Line 153 can you give an approximate depth here, e.g. 30-60 cm ?
- Line 183 “elevation” not “altitude” throughout the article. “Elevation” refers to distance between sea level and the surface of the Earth; “altitude” need not be at ground level. So lakes have elevations, aircraft have altitudes.
- Line 204 I’m confused by the “counting”. Later sentences indicate that only three adults were identified – did you identify three to determine species, then count until you had a larger number? Or did you only “count” the three that you identified? I do realize the difficulty in identifying species – would it be appropriate to identify them only to genus?
- Line 206 specify how ostracodes were isolated from the sediment. Direct ‘picking’ under a dissecting microscope?
- Line 250 Figure 1. In YG I see 39 localities but there are 45 lakes listed for YG. CAN-2, ENC-3, QUE-16, and KAN-28 seem to be missing from the map, and apparently two others. Also TEK-57 is shown on the YG map but is listed among GSHN in the cluster chart.
- Line 465 the cation list is first, the anion list is second. The phrasing here suggests that these anions are less common than these cations, but that indicates a charge imbalance. Ions should be listed by importance in a single list (e.g. Ca, Cl, Mg, Na... etc).
- Line 474 is the volcano “*dormant*” since 1994, or are its crater lakes “extinct” (whatever that means) as is suggested by the phrasing)? Please clarify. The phrase “extinct by the intensification of volcanic activity” makes no sense, because “extinction” cannot be caused by intensification of volcanic activity.
- Line 512 and elsewhere: use “richness” and “diversity” carefully, because they are not identical. “Richness” is the number of species, but “diversity” also factors in relative abundance. In addition, be aware of the difference between alpha diversity (e.g. in one lake) and beta diversity (e.g. in lakes of one region). In Line 515 you’re clearly writing about alpha diversity, but Line 511 seems to be about beta diversity. The two are not equivalent.
- Line 505 and later: these data reflect, in part, collector effort (for example, Hartmann probably collected many samples from around L. Nicaragua). Thus it is hard to reach definite conclusions when comparing diversities.
- Line 530 it’s not clear what “interannually... 38 species” means. Because 38 is less than 44, there seems to be no new information in this phrase.

## Technical corrections

- Line 20 “ecosystem” not “ecosystems”
- Line 23 insert comma after “limnology”; throughout the paper, in a list of three or more words, insert a comma before “and” (e.g. “geodiversity, limnology, and aquatic species associations”)
- Line 26 “karst” not “Karst”
- Line 27 change “mid-elevations” to “mid-elevation” and place between “volcanic” and “terrains”
- Line 28 insert “to” after “attesting”
- Line 28 change “identifies” to “identified”
- Line 28 change “ionic” to “chemistry”
- Line 29 remove comma after “formation”
- Line 31 add “s” to “association”
- Line 31 change “are” to “were”
- Line 32 change “reveals” to “revealed”

- Line 34 insert comma after “continuous” (as was done at Line 23)
- Line 35 insert comma after “properties” (as was done at Line 23)
- Line 35 change “Lakes” to “lakes”
- Line 40 insert “its” between “At” and “interaction”
- Line 40 change sentence from “contribute.... ecosystems” to “influence nutrient delivery to ecosystems”
- Line 45 change “is” to “results from”
- Line 45 insert comma after “composition”
- Line 48 change “Tropics” to “tropics”
- Line 54 change “central-southern” to “south-central”
- Line 56 delete comma after “types” and insert “along with” between “types” and “high”
- Line 63 replace “Panama” with “the Panamanian”
- Line 65 replace “Amazonas” with “Amazon”
- Line 66 delete “by dispersion”
- Line 68 replace “to” with “with”
- Line 71 change “In terrestrial taxa.... patterns” to “Therefore, for terrestrial biota, geodiversity and climate change have been the main influences on distributions”
- Line 72 change “on the basis of” to “following models of”
- Line 72 change “biomes” to “biome”
- Line 79 change “on” to “to”
- Line 80 change “source” to “sources”
- Line 82 change “dispose” to “disposal”
- Line 84 delete comma after “therefore”
- line 84 insert “places” between “suitable” and “to”
- Line 86 change “landscapes” to “landscape”
- Line 88 change “tracked back based on...facilitate” to “seen in fossil and subfossil communities, which improves”
- Line 93 change “well-suited” to “appropriate”
- Line 94 again, maybe some other word than “traits”
- Line 95 delete comma after “microcrustaceans”, and move “in recent environments” to follow “distributed”
- Line 97 change “with distributions ranging from” to “with individual species distributed in”
- Line 101 change “To” to “to”
- Line 102 change “(water and sediment... climate)” to “(geology and climate, as reflected by water and sediment chemistry)”
- Line 110 change “vascular endemic” to “endemic vascular”
- Line 112 need a year after “Partnership Fund”
- Line 112 change “On this region... America” to “In addition, species have arrived from South America and from more distant North America”
- Line 120 put negative sign in the superscript: “mm yr<sup>-1</sup>” not “mm yr<sup>-1</sup>”
- Line 121 be consistent – use superscripts OR slashes throughout (e.g. mm day<sup>-1</sup> OR mm/day)
- Line 126 delete “Whereas” and change “in” to “In”
- Line 127 change “include” to “can be categorized as”
- Line 129 change “(subterranean) rivers” to “rivers (both surface and subterranean)”
- Line 129 insert “as well as” between “rivers,” and “permanent”
- Line 133 change “data was” to “data were”
- Line 140 change title to “Field sampling”
- Line 142 insert “of” after “Peninsula”
- Line 145 change “the” to “a”
- Line 146 delete “echosounder”
- Line 147 change “the” to “a” and delete “navigator”
- Line 148 change “Water samples... Na<sup>+</sup>”) to “Sample for water chemistry”.

- Line 150 move superscript “3” to be a subscript
- Line 152 change “At” to “In”
- Line 157 recommended to avoid sub-sections here e.g. “Water chemistry analysis:” Instead, start the first sentence of each paragraph with a phrase of its data, e.g. for the first paragraph just delete “Water chemistry analysis”. Often it makes sense to start this sentence with “For” (e.g. “For sediment analysis, “) or “In order to”.
- Line 161 make “3” a subscript
- Line 162 change “balance,” to “balance;”
- Line 164 change “Past” to “PAST” and give version number
- Line 180 change title to “Cluster analysis and PCA of lake groups”
- Line 187 capitalize “Vegan”
- Line 190 change “for each... cluster” to “for each cluster” and delete comma after “cluster,”
- Line 191 insert comma after “variables”
- Line 192 change “arrows” to “vectors”
- Line 196 delete “of the data sets ... altitudinal” to “and to evaluate environmental”
- Line 198 by “feeding data” do you mean “input data” ?
- Line 201 transpose “software” and “Surfer®”
- Line 203 suggest title be “Ostracode analysis”
- Line 211 insert “are” between “and” and “currently”
- Line 212 change “used for ... using the Past” to “calculated using PAST”
- Line 220 change “vegan” to “Vegan”
- Line 227 delete “the software”
- Line 232 change “to model” to “the modelling of”
- Line 234 change “, which are... geodiversity,” to “(which are related to geodiversity)”
- Line 236 change “assumed as” to “taken as”
- Line 237 start the sentence with “We also considered other factors,” and delete “were also taken into account”
- Line 238 “used” not “use”
- Line 241 “R-squared” not “R-square”
- Line 242 delete comma after “data set”
- Line 245 change title to “Lake regions”
- Line 257 change “Lake full names of codes ... Figure 1.” to “Lake names are given in Table 1 and correspond to those in Figure 1.”
- Line 262 change “founded” to “found”
- Line 265 Table 1 could be presented more clearly. Align the column of Site names to the right, not centered. Reduce width of the first column. Use hyphens in column titles where appropriate. Increase width of coordinates so that digits fit on a single line. Make “Origin” a full column on its own (it is not, after all, a coordinate). Use abbreviations where helpful (e.g. T for tectonic, Mx for Mexico) and define these (and others like “Nd”) in the table’s caption. Transpose “Country” and “Site number” columns. It may be necessary to reduce the font by 1 or 2 units to prevent wrapping of cell contents. Please check the use of “<<” as it seems reversed: for example, as presented Site 2 has much less bicarbonate than sulfate.
- Line 265 Table 1 it may be worthwhile to separate these data into two tables, one for physical features (country, lat/long, elevation, surface area) and one for other features (with quantitative data for ions, e.g. Mg and Ca in their own columns). It may help to list the common ostracods in each lake here as well.
- Line 271 use past tense throughout, e.g. “explained” not “explain” and “were” not “are”
- Line 275 use lower-case for element names (e.g. “magnesium” not “Magnesium”) throughout the paper
- Line 283 be consistent with two decimal places in these percentages
- Line 308 suggest title be changed to “The YG lakes (karst terrain)”
- Line 309 rephrase “subgroup (YG1; n=9... central-southern,” as “subgroup in Region YG was

Subgroup YG1 which included nine lakes in south-central”

- Line 315 please provide reference to the age of the lake
- Line 315 again use past tense in most places in this paper.
- Line 319 “north-central” not “central-northern” and make similar changes throughout the paper
- Line 324 change “))” to “)”
- Line 325 put “both YG1” in parentheses
- Line 331 suggest title be changed to “The GSHN lakes (volcanic terrain)”
- Line 346 delete “in the northern Neotropics”
- Line 347 change “its” to “their”
- Line 351 change “to identify” to “the identification of”
- Line 353 change “whereas... value was” to “whereas the maximum value of the Shannon diversity index was”
- Line 355 insert “was” after “average”
- Line 356 delete “based on species occurrence data,”
- Line 359 delete “Guatemalan... American”
- Line 382 suggest “Relating ostracodes to environmental variables”
- Line 390 “Axis 1” not “axis 1”
- Line 398 change “thus,” to “thus were”
- Line 399 change “altitude” to “elevation” here and elsewhere
- Line 407 delete “and composition”
- Line 410 “Figure” not “figure”
- Line 411 “accounts” not “account”
- Line 415 “p” not “*p-value*”
- Line 415 “and the gray path is”, not “and gray path are”
- Line 416 “Numbers” not “Number”
- Line 426 change “GSHN group,” to “GSHN groups”
- Line 432 “peninsula” not “Peninsula”
- Line 433 delete “water ionic” and insert “in the water” after magnesium
- Line 435 delete “in the”
- Line 436 replace “and” with “and/or”
- Line 437 delete “spatial” and “map” and “contents in lake waters” and the comma after “3c)”
- Line 438 replace “seems to be the most” with “may be the more”
- Line 441 insert comma after “region”
- Line 449 “amounts” not “amount”
- Line 450 “subgroup” not “sub-group” – or change all to “sub-group” in the paper
- Line 455 units of “a<sup>-1</sup>” are given as “yr<sup>-1</sup>” elsewhere. Please be consistent.
- Line 455 delete comma after “2011a)” and change “allow” to “allows”
- Line 456 “increases” not “increase”
- Line 460 insert “Many” before “Central American”
- Line 465 insert “often” after “are”
- Line 466 replace comma with colon
- Line 467 delete space in “precipitation -evaporation”
- Line 472 delete comma after “2017)” move “(Armienta... 2017)” to follow “Central America”
- Line 473 change “than” to “as”
- Line 485 insert “to” after “finally”
- Line 478 replace “because of” with “as evidenced by the”
- Line 482 delete space in “ion- specific”, and replace “carbonates-bicarbonates” with “carbonates and bicarbonates”
- Line 484 use “1)”, “2)”, and “3)” as you do elsewhere, rather than “i)”, “ii)”, and “iii)”.
- Line 495 please rephrase this sentence, for example as “At regional and local scales, aquatic systems vary in limnology and sedimentology, and these are strongly influenced by morphology and origin.

Because morphology and origin reflect geological factors it is clear that, overall, geology is the primary determinant of lake properties and therefore also of geographic lake groups.”

- Line 500 suggest title “Geology as driver of ostracode distribution”
- Line 501 change “ostracode species” to “ostracodes”
- Line 501 change “systems ... Neotropics” to “systems, and the fact that we found seventy species confirms high richness for this group the northern Neotropics”. Is 70 species really considered a large number of species??
- Line 505 delete “of”
- Line 505 insert “individual” between “in” and “aquatic”
- Line 509 change “as in ancient” to “as ancient”
- Line 512 “equator” not “Ecuador”
- Line 513 change “occurred at” to “during”
- Line 514 replace “Stadials ... scale” with “Stadials, caused most shallow lakes in this region to dry out completely”
- Line 515 suggest to start a new paragraph with “The low species diversity...”
- Line 520 insert comma after “availability”
- Line 520 replace “in biological evolutionary processes of” with “the evolution of”
- Line 521 delete “(as benthic organisms)” and delete “and species adaptation”
- Line 521 “alternative” not “alternatively”
- Line 522 delete comma after “Neotropics” and replace “rely in” with “relies on”
- Line 522 rephrase “human activity intensification in” as “intensified human impacts on lakes”
- Line 523 change “America water ... ecosystem” to “America, the greatest human impacts are frequently water extraction, fishing, fish farming, eutrophication, and waste water disposal”
- Line 525 delete “are known to”
- Line 526 insert “both” between “(“ and “genetic”
- Line 527 change “development” to “investigation”
- Line 529 change “Comparatively... 44 species” to “In South America, the floodplains of the Upper Parana River may support up to 44 species”
- Line 531 change “In Colombian... however,” to “However in Colombia”
- Line 534 change “are” to “is”
- Line 534 change “of other” to “from other” and delete worldwide
- Line 536 change “how ... group” to “and latitudinal diversity gradients”
- Line 538 change “do not... faunas” to “form distinct assemblages”
- Line 538 delete “ly” from “similarly”
- Line 540 rephrase sentence as “Ostracode assemblages in the YG region (OST1 and OST2) are distinct from those of the GHSN region (OST3, OST4, and OST5).”
- Line 543 delete comma after “subregions”
- Line 544 change “over” to “on”
- Line 544 replace “in fact, recovered” to “confirmed”
- Line 544 rephrase sentence as “The indirect effect of conductivity on species composition, however, was significant in SEM and explained 67% of the variation in the data.
- Line 548 delete comma after “conductivity”
- Line 548 inset “The” before “SEM” and change “reveal” to “reveals”
- Line 548 change “altitude” to “elevation” here and elsewhere in the paper.
- Line 549 rephrase as “The SEM models also suggest a reduction of species richness with elevation of individual sites.” (this is alpha diversity)
- Line 553 change “suggest” to “suggests”
- Line 554 delete comma after “variable”
- Line 556 delete “, for example”
- Line 558 change “to” to “with”
- Line 559 change “In Central America .. evident” to “In the GHSN region ostracodes were also distinct

between subregions”

- Line 560 change comma after “5b)” a colon
- Line 561 change “to” to “with”
- Line 564 delete “Future .... groups.”
- Line 567 change “structures” to “characteristics”
- Line 569 change “for” to “over”
- Line 573 change “of the” to “within each”
- Line 573 change “lakes composition” to “limnology”
- Line 575 change “trait” to “traits”
- Line 580 rephrase last sentence perhaps as “Further studies should aim to expand our knowledge of lake and biotic regions by including a greater number of lakes, variables, and taxonomic groups.”

-