

## ***Interactive comment on “Constant diversification rates of endemic gastropods in ancient Lake Ohrid: ecosystem resilience likely buffers environmental fluctuations” by K. Föller et al.***

**A. Weigand (Referee)**

WeigandA@gmx.net

Received and published: 6 October 2015

General comments The paper tackles the hypothesis that the largest endemic organism group of Lake Ohrid - namely non-pyrgulinid Hydrobiidae (Gastropoda: Caenogastropoda) - underwent a constant rate of diversification during their *in silico* lake evolution. This is particularly interesting, since insights into potential diversification shifts for those gastropods may further be used to obtain a deeper understanding about the presence and impact of catastrophic events and/or environmental + climatic changes during the evolution of Lake Ohrid and its resilience potential in general. I very much like the addressed research question and in principle also the methodological pipeline

C6244

used, however I assume that several of the authors arguments are not / not fully supported by the data they present. I raise seven points for discussion:

Specific comments 1) In the discussion section, the authors state that "[t]he findings suggest that the non-pyrgulinid Hydrobiidae form a monophyletic, speciose and endemic clade and thus, by definition, represent a species flock". However, and since Bayesian posterior probability values have a strong tendency towards high values (see e.g. Cummings et al. 2003 - Comparing bootstrap and posterior probability values in the four-taxon case), a support value of 0.87 can not be seen as a decisive support nor as an argument for the monophyly of this organism group. Thus, hypothesis 1 must be questioned. Furthermore, and because this hypothesis is the authors self-defined requirement to perform all subsequent diversification analyses, the monophyly of this group has to be unambiguously demonstrated first. I would suggest to additionally implement Maximum Likelihood analyses complementing the BI data and perhaps an even more dedicated partitioning scheme with different partitions based on the three CO1 codon positions and different partitions for the stem and loop regions of the 16S rDNA fragment. A nuclear marker would be great, but I see that this will be hard to achieve for all the specimens.

2) Again in the discussion, the authors argue that "[o]ur LTT-plot (Fig. 4a) shows that several lineages have already existed when the lake reached deep-water conditions or even before it came into existence (see Trajanovski et al., 2010 for a potentially similar pattern in the Dina leach flock). Therefore, the term “cradle” may not only be used for monophyletic species evolving within the lake (intra-acustrine speciation) but also for a group of monophyletic species that started to evolve within the (palaeo-)basin." Important here and addressed are the two oldest nodes of the non-pyrgulid Hydrobiidae (> 2mya) and their four corresponding lineages. However, both assumed-to-be monophyletic groupings within the non-pyrgulinid Hydrobiidae (i.e. the resulting 4 lineages) have no posterior probability support at all (= below 0.5 according to the authors), must be collapsed and hence can not be regarded as being monophyletic.

C6245

The argument of several lineages existing prior to lake Ohrid formation thus does not hold true. Same is true for the monophyly of the mentioned species groups.

3) Based on the inferred constant rate of diversification the authors conclude that "[the] initial working hypothesis – rate homogeneity – cannot be rejected. [...] If we assume that the rate of diversification in the non-pyrgulinid Hydrobiidae from Lake Ohrid is constant, linking environmental/climatic fluctuations to changes in tempo of diversification becomes impossible (see specific goal iv)." The diversification rate is defined as the net sum of speciation rate - extinction rate. Rate homogeneity thereby refers to the following (see Ricklefs 2007 - Estimating diversification rates from phylogenetic information): "The simplest diversification process presupposes that rates of speciation and extinction are the same for all lineages and do not vary over time. This is the assumption of rate homogeneity." Thus, rate homogeneity is mostly used as a model and can not be inferred from a constant rate of diversification as it refers to constant rates of speciation and extinction. Ricklefs further states that "different combinations of speciation and extinction rates can produce the same expected clade size". This means, that different combinations in speciation and extinction rates may lead to similar inferred net diversification rates. As an example: potential environmental/climatic fluctuations (as proposed by some studies for Lake Ohrid) may have lead to extinction events and subsequent adaptive radiations within relative short periods of time. This is known for other taxa including hydrobiids from other regions of the world. The outcome would be a temporally increased rate of extinction and subsequent increased rate of speciation. However, and for the observed 0.1 my intervals, the "phylogenetic window" may be too broad to see those changes in extinction and speciation rates leading to a similar net diversification rate (i.e. birth and death of lineages) as expected under rate homogeneity. Thus, rate homogeneity as defined as constant speciation and extinction rates can not be inferred nor supported by the data presented as only a constant rate of diversification is observed. Finally Ricklefs (2007) states that "it is unlikely that rate homogeneity can be unambiguously supported for any clade." However, the inference drawn by the authors that "Lake Ohrid never experienced catastrophic environmental

C6246

events that resulted in the extinction of all or most of its endemic taxa and thus caused a "reset" of diversification processes" can be partially drawn as old lineages can be observed. However, lineages may have even survived those events. If this point is still dealt with in the discussion, it should be re-formulated according to what is supported by the data. I would be really careful in drawing the final conclusion that a high ecosystem reliance can be supported by the "rate homogeneity" of the investigated gastropod taxon. The opposite may be true: If a significant rate shift is observed and can be temporarily linked to a given environmental event, this may be seen as support for the influence of this event on the diversification process. Related to this discussion, see also point 4.

4) The authors refer to the species coverage, an important factor when calculating diversification rates based on phylogenetic tree hypotheses, by saying: "As for the sampling size (i.e., 17 out of 27 nominal species studied), the high diversity of evolutionary lineages found in our phylogenetic analyses indicates that our sampling design likely recovered most major evolutionary lineages within this cryptic group." It is correct, that there is a high probability that the authors cover all older /major lineages. However, at the same time, many more recent species may have been missed. This seems very plausible since a high amount of cryptic species is observed in Lake Ohrid, e.g. see *Pseudohoratia ohridana*. If more recent nodes accumulate, net diversification rate may show a shift in more recent times, e.g. during glacial cycles (Lindhorst et al. 2015). The oldest node referring to the taxon *P. ohridana* is even at 1 my of age. This taxonomic coverage problematic and its influence on the diversification rate estimates has to be addressed more thoroughly in the discussion. However, distinguishing such a pattern from a "normal" pull-of-the-present effect due to speciation only processes will get challenging.

Some further methodological aspects:

5) The authors state that "we did not test for substitutional saturations as both genes have been suggested to be not saturated within the family Hydrobiidae (Wilke et al.,

C6247

2001, 2013)". However substitutional saturation is dataset- and sampling-dependant and a general statement can not be made here. Please test your dataset for substitutional saturation as this is easily done but at the same time may heavily bias the reconstruction of the phylogenetic tree hypothesis by lowering the phylogenetic information content of the data.

6) Briefly explain the methodological procedure of the diversification rate analyses performed in TreePar in more detail, as it is used as an additional argument besides the LTT-plots. The analytical difference should become more clear.

7) In the discussion, the authors argue that "the TreePar analysis used does account for incomplete sampling and we did infer two single rate shifts in the present study." More precisely, a single rate shift each in two independent trees has been identified. Or do you have the assumption of two rate shifts during the evolution of non-pyrgulinid Hydrobiidae in Lake Ohrid? Would it be possible to show even more trees. I am not familiar with the standard procedure, but showing ten trees only seems rather too less.

---

Interactive comment on Biogeosciences Discuss., 12, 14271, 2015.