

## Posudek magisterské diplomové práce

Bc. Tomáš Dvořák (2018). Phylogeography and population structure of two loach species (Teleostei: Nemacheilidae) in Southeast Asia

### **Reviewer summary:**

I applaud the author on a very interesting, well written and well researched thesis on much more than what is in the title (Phylogeography and population structure). The thesis will surely serve as a foundation of a well cited research on integrative systematics of the two studied groups and beyond. The only problematic part I found is the biogeographic analyses where a wrong method was used. The author however paid no attention to the wrong results and correctly identified almost all biogeographic events even though they were not recovered by the analyses. A more analytical approach could have also been taken in the examination of the morphological data.

### **Review point by point:**

Language:

The language is generally comprehensible, but is often awkward and grammatically incorrect, but not to a stronger degree than other foreign language theses I have seen.

### **Introduction:**

Well structured, well written, informative, with all the relevant information in particular concerning the interesting and famous faunal breaks along the Thai-Malay peninsula. One thing I feel is not considered in the Introduction is the difference between eustatic and isostatic sea-level changes, and especially consideration of reconstructed past paleogeography (see my comment to p.8, 1.1). The Introduction is written as if the landscape topography was static and with sea-level changes the only changes to have occurred.

The Introductory sections on the studied fish groups are again well written.

### **Methods:**

Methodology is standard and well performed except for the biogeographic analyses and morphological analyses (see Results section for the latter).

Some comments and questions follow.

### ***Dating.***

While the separation of the Indian and Indochinese river basins appears reasonable as a calibration point, it is still only a hypothesis that needs testing, especially its relevance at the species level diversification. The use of this calibration point is questionable geologically but even more questionable biogeographically given that clades, genera and even species of the studied group cross most geomorphological boundaries in SE Asia (see also Q1). I would thus have welcomed more than one calibration point, especially calibration points at various time frames of the past. The biogeographical literature is full of obvious geological calibration points that have been questioned, have been found much more complex, prolonged, or having occurred repeatedly throughout the geological history. The cited study used for the calibration is not a comprehensive palaeogeographical summary of the geology of the focal area.

**Q1:** It is not clearly described which node(s) in the phylogeny were calibrated, what clade with what geography they represent, along which geographical line was the calibration drawn (the southern elongations of the Himalayas are many, which was chosen ?; etc.).

**Q2:** The offset of 0.5 My (rather small for an event 26 My old) for the calibration point was taken obtained how?

**Q3:** Why did you use the Yule process of speciation for your prior?

***Biogeographic analyses:***

**Q4:** I am not a familiar user of BBM analysis. Can you explain how it differs from DIVA type of analyses? Why was BBM analysis preferred over the much more often used analyses also implemented in RASP?

**Results:**

***Dating:***

Despite my comments regarding dating the divergences between in lineages in each species-group are substantial and suggest that the calibration is not very far off. The divergence rate would be about 0.7 % per My for cytb, a reasonable and comparable result.

***Discrepancies between mtDNA and nDNA data:***

I am happy to see that the author discusses these discrepancies. I have seen too many theses where the students just don't care and combine completely conflicting data partitions.

**Q5:** How would the author biologically interpret the conflicts between the mtDNA and nDNA partitions in these particular cases?

There are some interpretations in the Discussion such as „Most likely, at some time in the past female specimens of *P. zonalternans* from the Sittaung basin entered the Irrawaddy basin and bred with the Irrawaddy population. After numerous backcrossing with the Irrawaddy population the nuclear DNA of the Sittaung lineage was replaced completely, but the maternally inherited mitochondria were being still present (mitochondrial introgression).“

Or „In both cases of secondary contact, the lineages interbred to the extent that the nuclear genetic markers of one lineage got lost from our dataset.“

**Q6:** Are these correct interpretations? Are these the only interpretations? Have you found heteroplasmy in your mtDNA sequence data or a similar pattern in your nuclear sequences?

***Biogeographic analyses:***

**Looking at the results of the BBM analysis it is immediately obvious that the BBM analysis is not optimizing biogeographic events.**

**I have thus have redone the biogeographic reconstructions of both studied groups (without branch-lengths obviously) using DIVA optimization (by hand), S-DIVA optimization (in RASP), BayArea optimization (in RASP) and the BBM analysis (also in RASP).**

**The biogeographic reconstructions from all my analyses except BBM have basically one and the same result which is an almost completely vicariant history of the groups, instead of completely dispersalist history as in the thesis and in my results of the BBM analysis. Same data, the opposite result.**

**The reason for this is that the BBM analysis is a character-optimization analysis, not an event-optimization (or area-optimization) biogeographic analysis. It is thus not useful for biogeographic analyses, but for character-transformation analyses (optimization of**

characters states). The BBM analysis is actually in RASP included in the second batch of analyses together with other character-optimization analyses, while all proper biogeographical analyses are in the first batch. While the manual talks about area optimization, this is clearly wrong.

I attach figures of the redone analyses to this review (Figs 1 and 2).

Despite the erroneous results of the BBM analysis an immense credit has to go the author because unlike most students these days he instead of relying on the wrong results correctly identified (I assume by „hand“, i.e. force of reasoning) almost all the biogeographic events and thus almost correctly interprets the biogeography of the group in the accompanying annotations of the phylograms. The interpretation annotations on the phylograms thus absolutely do not correspond with the results of the BBM analysis.

The text interpreting the biogeographic history of the groups is a jumble of the BBM results and authors „intuition“ and thus contains many mixed interpretations. I hope that with help of my input the author will be able to correctly interpret the biogeographic history of the groups. I would like to hear this in the Q&A section of the defense.

**Comment:** Would have been nice to use corresponding colors in Figs. 4 and 2, 6, and in 4 and 5 (and in all derived Figs), as it is, it is hard to follow. This also results in difficulty of comparing text of results with Fig. 4.

#### ***Morphological analyses:***

This part of the thesis could have been done in a more analytical way. The descriptive part in tabular form as done in the thesis is a necessary first step.

The allometric parameters could have been studied using simple growth curves.

The potential for separation of putative species could have been done using multivariate analysis and the degree of separation could have been done using the accompanying statistics.

Another option would have been to optimize the characters on the molecular phylogenies, either externally or in a direct optimization with the molecular data sets.

All these approaches are standard and routinely used in the field of systematic biology.

On the other hand it has to be noted that the morphological part is an extra and most theses I have seen at this degree do not combine both molecular and morphological analyses, but instead focus on one or the other. So clearly the thesis contains much more than is usual.

#### **Discussion:**

The Discussion is well written and in a clear way summarizes the results of the thesis.

I am not making any comments to the reconstructed histories of the studied species groups do to the flawed biogeographic analyses and the jumbled interpretation of them.

**Q7: What is your preferred species concept? Or which species concept you think is appropriate for the type of data, analyses and species groups that you study?** In the Discussion you write for example: „It demonstrates that Irrawaddy and Sittaung lineage on the one hand and the Sittaung and North Thailand lineage on the other hand are not reproductively

isolated when coming in contact.“ This looks like you work within the boundaries of the biological species concept.

Minor comments:

p.II, *Abstract: around 250 specimens of P. zonalternans ... were investigated using ... geologic data.*

p8, 1.1: *Southeast Asia is formed by a continental part and an island part.* This topic is not totally relevant to the thesis which only deals with the „continental part“ following this division. But the division itself is rather irrelevant, since 9/10 of the last two million years and all duration of the time frame of evolution of the studied group the „island part“ was not an island part, but a continuation of the SE Asian peninsula (all the way to Borneo) in terms of subaerial land exposure. Sumatra and Java were part of this peninsula and part a separate western island chain of volcanoes in the older epochs during the Miocene and Pliocene. This is also evident in the fact that the present-day islands are not units of endemism for freshwater fishes and other freshwater organisms – the FF faunas are shared by former „inter-island“ river drainages.

p10 and elsewhere: Nemacheilidae vs. Nemachelidae.

Table 8. The difference between the mtDNA and nDNA divergences basically fits the 4N rule, so no reason for surprise.

**Reviewer conclusion:**

**I consider the strong points of the thesis to outweigh the weak ones and thus consider the thesis as a very good one and recommend it for defense.**

V Českých Budějovicích  
6. 9. 2018  
doc. Mgr. Oldřich Říčan, Ph.D.

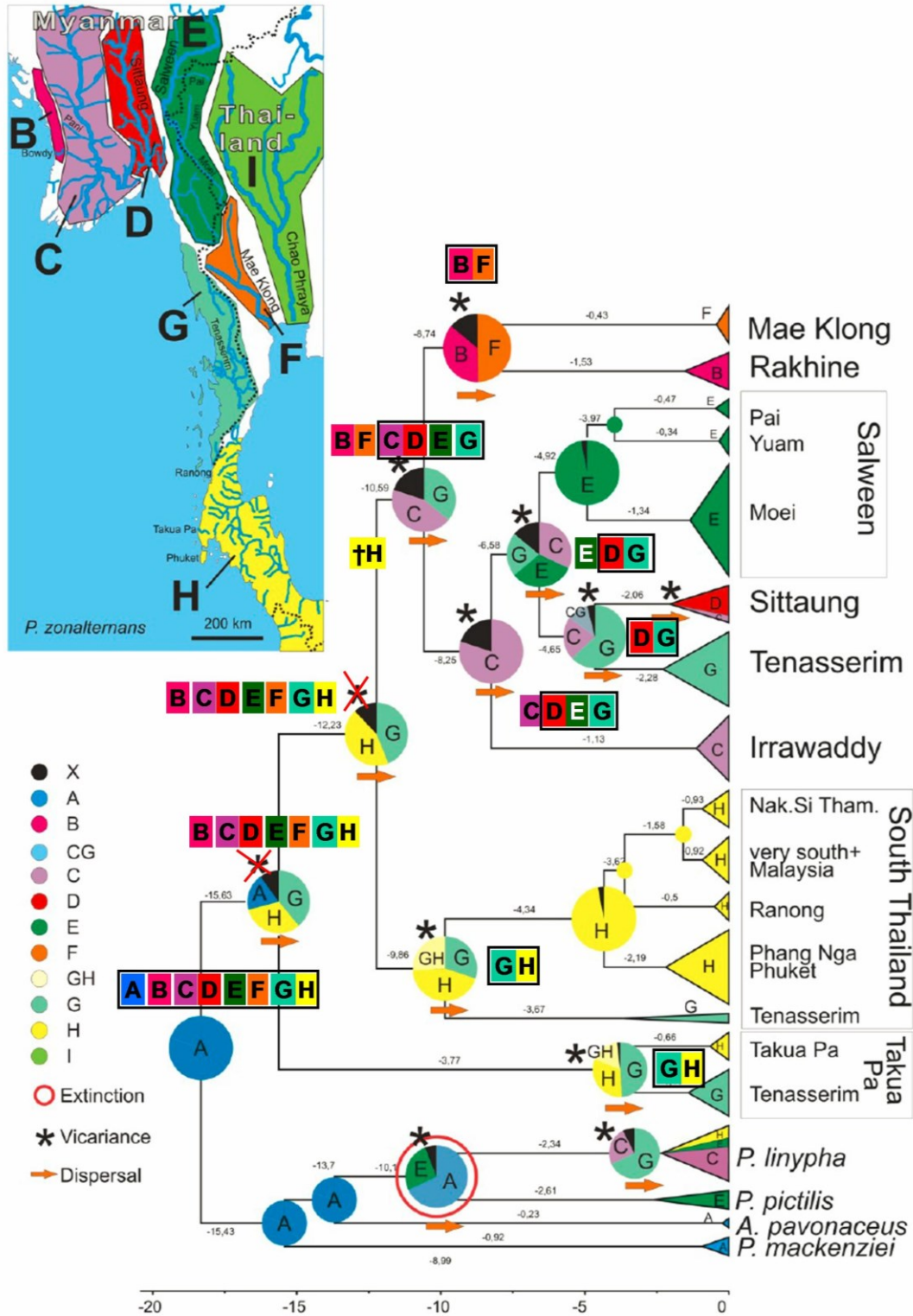


Fig. 6. Ancestral area reconstruction for *P. zonalternans* group performed with Bayesian binary MCMC analyses. Pie charts at the nodes represent the probabilities of each unit area in the ancestral range. X (black) - areas with probability lower than 10%. A (blue, not included in the map) - India/Bangladesh.

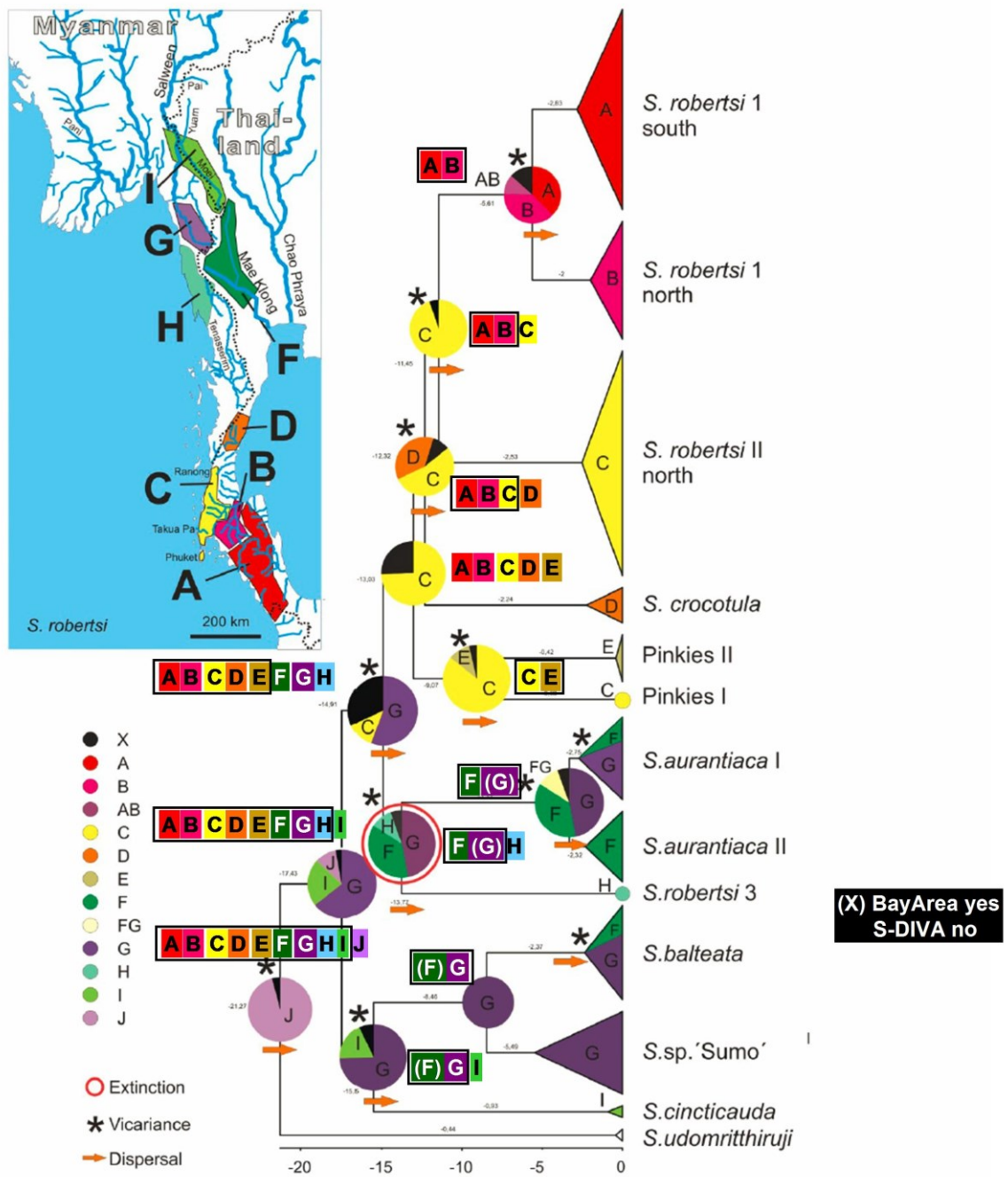


Fig.8. Ancestral area reconstruction for *S. robertsi* group performed with Bayesian binary MCMC analyses. Pie charts at the nodes represent the probabilities of each unit area in the ancestral range. X (black) - areas with probability lower than 10%. J (light violet, not included in the map) - India/Bangladesh.

Fig. 2