

## Thesis Report and Evaluation

**Panel member:** Dr. Jesús Gómez-Zurita, Institute of Evolutionary Biology (CSIC-Univ. Pompeu Fabra), Barcelona, Spain.

**Author of the PhD thesis:** Mr. Matthias Seidel, Charles University, Prague, Czech Republic.

**Title of the PhD thesis:** Evolutionary history, systematics and biogeography of Southern Hemisphere hydrophilid beetles (Coleoptera).

### REPORT

The main body of this thesis, reflecting the involvement and outcome of the research carried out by the candidate for his PhD project includes six articles, four of them published between 2016 and 2018 in JCR journals, of which the candidate is first author in two, and two more presented as draft manuscripts, where the candidate is first author in both. **The topic and the taxonomic and biogeographic scopes of these articles are compact, coherent and they are well represented by the global title of the thesis.** This work thus constitutes a good example of the specialization and expertise of the candidate, which is perfectly described in the first sentence of the thesis abstract: "the research presented in my PhD thesis consists of phylogenetic, biogeographic, taxonomic and ecological research of Southern Hemisphere water scavenger beetles (Coleoptera: Hydrophilidae) with a special emphasis on New Zealand."

But, in particular, the amount of **taxonomic contributions** resulting from this PhD work toward the modernization of the classification of an important lineage of beetles is outstanding. If I followed it correctly, these contributions would at least include: (1) perhaps a new species of *Adolopus*; (2) a new genus allied to *Adolopus* with one species; (3) another new genus also allied to *Adolopus* with two species; (4) the reinstatement of Andotypini; (5) synonymy of Borborophorini with Cylomini; (6) thirteen new species of *Cyloma*; (7) synonymy of *Psephoboragus* with *Cyloma*; (8) synonymy of *Namostygnus* with *Cyloma*; (9) synonymy of *Gitocyloma* with *Cyloma*; (10) three new species of *Cylomissus*; (11) a new genus allied to *Cylomissus* with one species; (12) the new genus *Enigmahydrus*; (13) a new combination to *Relictorygmus*; (14) new species of *Relictorygmus*; (15) new species of *Rygmostralia*; and (16) reinstatement of *Tormissus magnulus*.

The **introduction** offers an informative description of the separation of Gondwana, based on the hypothesis of Sanmartín & Ronquist (2004) and a less detailed account of the geological evolution of New Zealand, introducing the idea of two competing hypothesis of profound biogeographic relevance for terrestrial organisms: (1) the Oligocene drowning of most emerged land and subsequent colonization of the current islands, against (2) the persistence of terrestrial biota since the isolation of the corresponding landmass in the Late Cretaceous. The candidate seems to take sides with the later hypothesis already in the introduction admitting the existence of relict lineages and accepting the refutation of the drowning hypothesis by studies on chironomids and harvestmen, among others. The idea that both palaeoendemics and recent arrivals to New Zealand may exist appears, however, in the articles and

manuscripts of the thesis, reflecting a broader view than merely accepting the arguments of these pro-Gondwanan studies. Indeed, I put on the table the question whether these studies really refute the drowning hypothesis. Or at least, are these really good examples to refute the models of geologists? In the case of the chironomids (Krosch & Cranston 2013), their credibility intervals actually overlap the expectations under drowning scenarios, even though the authors favor only the conclusions based on their pre-Oligocenic averages of their estimates. Also, dating of Opiliones by Giribet and coauthors is always kind of funny: e.g., Giribet and Boyer (2010) reported ages for New Zealand genera between 55 and 118 Ma, and Giribet et al. (2016) provided ages for the same genera in the range between 135 and 215 Ma.

The **introduction** finishes with a very informative (and nicely illustrated) overview of the Hydrophilidae in New Zealand and the Cylominae in the southern hemisphere, which is very useful for the reader not familiarized with the taxonomy and diversity of these groups to get notions which will be useful to follow the results of the research. The account of species diversity in New Zealand mentions a good number of hydrophiliid species in New Zealand that have ranges outside of the islands, typically in Australia and/or New Caledonia, but also Palaeartic origin. In every case, the occurrence in New Zealand is interpreted as an introduction. Considering that at the core of the thesis there is a discussion about dispersal versus vicariant biogeographic models, understanding these introductions may not be trivial at all. What mechanisms are considered to explain these introductions? Could these mechanisms have operated episodically in the past to add more plausibility to dispersalist models?

**Chapter 1** consists of a revision of the work by McKenna et al. (2015) on the timeline for the evolution of Coleoptera. This article points out some methodological problems with this work relative to the rigour and robustness of analyses, but also on some analytical decisions particularly in the choice of fossils for calibration. The role of the candidate in this work was mainly related to reviewing the literature for fossil selection and justification for the dating analysis. Indeed, the authors propose a new set of calibration points and similar analytical conditions (enforcing the tree topology versus constraining some nodes to be monophyletic, as in the original study) to revise the dates of Coleoptera and of the main lineages in the Order and reveal that most ages should be pushed back in time around 70 to 90 Ma. This implies an origin of Coleoptera [317-349 Ma] during the Carboniferous [299-359 Ma], which may not be necessarily contentious, but generates several problems. Among them, I highlight two, of which the second is also given some consideration by the authors: (1) there appears to be nearly always a huge gap, typically and rather consistently around 100 Ma, between the origin of lineages and the occurrence of reliable fossils in these lineages; and (2) there is a clear dissociation between the origin of Phytophaga [227-252 Ma, in the Triassic], which have a strong association to angiosperms, with the rise [in the Jurassic: 140-180 Ma] and dominance [in the Cretaceous: 72-145 Ma] of this group of plants. I am particularly interested in this last clade, and it certainly shows a strange topology based on previous studies and taxonomic knowledge, including unexpected relationships (Vesperidae deeply nested within a cerambycoid clade) and unexpected time frameworks ([CER+ORS+MEG] much younger than Chrysomelidae).

All in all, these issues (and there may be others which the specialist in each group may identify), makes me entertain the possibility that the results of the study could be flawed too. For example, I have identified what I consider problems relative to the calibration and the implementation of the molecular clock, choice of fossils and definition of calibration points, which I hope I can discuss in detail with the candidate during the defense. And, of course, my main worry about these potential problems is not only about the relevance of this particular

article, which is already published and has been assessed by peers, but the fact that they may have an important repercussion on the next chapters of the thesis, some of which are unpublished. I think the candidate and his co-authors may benefit from some discussion on the topic and hopefully the conversation that we will initiate during the defense can help in that direction.

**Chapter 2** is a molecular phylogenetic study of the subfamily Cylominae of the Hydrophilidae, where the systematics of the group is refined, and paying particular attention to resulting biogeographic patterns, with the main conclusion that New Zealand lineages are palaeoendemic, rejecting the drowning hypothesis for New Zealand, but recognizing a number of long distance dispersal events across the southern hemisphere. This work looks very much as the core of original research that the candidate did for his PhD thesis, which is reflected in his position as first author of the article and his involvement in every stage of the study, chiefly in the joint writing with his supervisor.

The study builds on the knowledge of historical changes of the meaning, limits and composition of the subfamily, as well as on previous hypotheses of relationships based on phylogenetic analyses of morphological characters and preliminary molecular phylogenetic hypotheses, and provides with a robust sample (23/24 genera, 60/84 species) to investigate these previous propositions and refine the systematics of the group. This is one of the unpublished articles, and it is one of two chapters that can benefit the most from discussion in the defense and elsewhere, because it shows some room for improvements (see some of the comments below). In particular, the results of this study may be affected by analytical decisions mostly on how the molecular clock was implemented to obtain divergence dates, a most relevant piece of information resulting from this study, and perhaps on the design of biogeographic analyses, as will be discussed with the candidate. Other structural and presentation aspects that may benefit the presentation of the work could possibly include:

1 - Half of the DNA fragments used for phylogenetic analysis are from rRNA genes, which are prone to alignment issues. Yet, the strategy used for multiple sequence alignment (Muscle/ClustalW) is possibly amongst the less robust among available possibilities. What motivated this choice? Was there any test of the impact of different alignment strategies on the retrieved topologies? The same alignment issues also apply to the results of Chapter 5, and maybe the impact in the latter may be even higher, considering that in this case 3 out of 4 markers are fragments of length-variable rRNA genes.

2 - Here and in other chapters, the phylogenetic approach of choice is nearly always and in exclusivity Bayesian? I am curious about this methodological loyalty, when complementary approaches like maximum likelihood exist and they are more intuitive and straightforward in terms of interpreting support or how data interact toward a certain topology.

3 - The description of methods relies too heavily on the supplementary material, they are not always comprehensive, and should benefit from a clearer and ordered explanation with better justifications:

- For example, previous studies are cited, but it is not immediately obvious how much new information is contributed in this study (I believe it is a lot and this should be highlighted).

- Also, when it comes to calibration points (and there are some issues which may affect them, as will be discussed, hopefully), differences with previous analyses are mentioned but not justified. Moreover, it is impossible to recognize which specific nodes in the phylogeny apply to which constraints, unless one scans the whole tree and is familiar with the taxonomy of the group. Highlighting graphically these nodes in a phylogeny would be really useful.
- About the actual dating analyses, we are not given any information about topological constraints, if there are any, or about clock models implemented. The latter are particularly relevant, since the unconstrained tree seems to reflect high rate heterogeneity, and if constant clocks were applied, the results might show important biases. In fact, only when reading the results (first sentence of 'Divergence dating' section) one learns that there must have been a comparison of relaxed and strict clock models and that the former were selected; however, this test is not described in the methods and the results are not shown.
- In the description of biogeographic analyses, specific technical details of the analysis (tested ranges and dispersal rates) are presented before one knows which particular analysis will be performed and why these details were important.

4 - The results start by describing phylogenetic patterns that are not obvious in the accompanying figures: the idea of "main clades of Cylominae" appears before describing or justifying it; the idea that clades are not limited geographically is not obvious unless one knows the distributions of taxa; the correspondence to Hansen's (1991) classification is not obvious unless one is familiarized with the taxonomy of the group. Some of the arguments leading to these claims can be found later in the text and crosschecking with the figure showing the phylogeny, but until then, the description is difficult to follow without assistance. All of these minor problems, as with the fossil placement mentioned above, can be easily solved by adding relevant information to Figure 1 (e.g., color-coded branches using the same colors as in Figure 2 for distributions; naming the "main clades"; labeling or coding Hansen's tribes; etc.).

5 - The chronogram of biogeographic splits in figure 3 shows a split of Australia and South America that is younger (30 Ma) than proposed for standard biogeographic models as shown in the introduction and described in the methods of the article (35-52 Ma).

**Chapter 3** of the thesis is a published taxonomic paper dealing with nomenclature and establishing Cylominae as valid name for the focal group of interest of this PhD project. It is a revisionary work likely resulting from the original and necessary immersion of the candidate in the taxonomy of the group of study. It is a useful study and a nice example of obtaining academic and curricular benefit from the time devoted to study a research problem. From the point of view of structure of the thesis, I think it would have been perhaps better to list this nomenclatural article as the second chapter of the thesis, because it provides useful taxonomic background to interpret information used, for example, in current Chapter 2.

**Chapter 4** is a published article detailing a taxonomic study of two of the genera of Cylominae which were previously considered one genus distributed in Chile and South Africa: *Cylorygmus* (Neotropical) and *Relictorygmus* (Afrotropical), the latter described in the course of the chapter. In this article, the candidate is the first author and has been responsible for all the relevant stages of the research, except for the description of immature stages, which has resulted from collaboration with a Japanese colleague, specialist in this type of studies.

This collaborative aspect of the thesis is considered very constructive and formative and it is evaluated as very positive. The study is rigorous and the descriptions of both adult and larval stages of the species involved are exquisite. The zeal of the authors to save their molecular phylogenetic results for other articles generates, however, an odd situation where one of the most relevant conclusions of the study, as relevant as to be used as the eye-catcher in the title ('breaking a disjunct distribution'), is only contained in a sentence in the discussion stating that "a preliminary analysis of molecular data suggests that *Cylorygmus* and *Relictorygmus* are not closely related" (p. 400 of the article). Otherwise, other statements that try to link purely taxonomic findings with biogeographic considerations are inexact:

"Our study demonstrates that the trans-Atlantic disjunct distribution of *Cylorygmus* was based on inaccurate taxonomic treatment and did not reflect the real evolutionary history of these beetles" (abstract).

"A detailed examination of the morphology of the Chilean and African species and the discovery of the third new species revealed that the Chilean and South African species are not closely related. This justifies the placement of the African species into a new genus, *Relictorygmus* gen. nov., breaking up an incorrectly assumed disjunct distribution of *Cylorygmus*" (conclusions).

I think that the split of *Cylorygmus* in two genera based on the taxonomic analysis presented in the work is sound and rigorous. However, if *Cylorygmus* and *Relictorygmus* had been sister groups, and this information is not available in the taxonomic treatment or examination of morphology presented in the work, it would still support the same biogeographic scenario and essentially disjunct distribution of an evolutionary lineage. It is important to bear in mind that genera are, after all, artificial entities. Thus, what challenges the scenario is not the taxonomic treatment, including the recognition of a new genus, but the phylogenetic relationships, which are not the object of the study or substantiated with any hard data or analysis, but a mere reference to preliminary data.

**Chapter 5** is an unpublished study of diversity, distribution and evolution of another Cylominae genus (*Saphydrus*), allowing for the recognition of an overlooked evolutionary lineage thanks to the study of unknown larvae, which is interpreted as a new genus of Cylominae, sister to *Saphydrus*, the genus *Enigmahydrus*. This work is based on rigorous taxonomic analysis of adults and larvae of the species of interest, but it goes one step beyond in the analysis of the diversity of these groups to try to offer an explanation on the reasons for the scarcity of these taxa in collections. Research for this work was done in a truly collaborative manner where different authors contributed their main expertise towards a cohesive study, with a main involvement of the candidate in molecular phylogenetic analyses as well as taxonomic study of adult beetles.

The main objection that I identified in the case of this study is that the conclusions on changes in the patterns of occurrence are not based on any kind of formal analysis, but on informal speculations based on the description of trends deduced from face-value, more or less circumstantial observations. In our trade, this kind of anecdotal data is usually exploited to elaborate hypotheses rather than drawing conclusions. Could climate change in the past 40-50 years explain shifts in species ranges to areas that have not been surveyed regularly, so that what was common in an area in the 70s is now common but undetected elsewhere, thus challenging notions of conservation status based on circumstantial data? [A little confusing detail in the manuscript: *altitude* generally refers to the distance between an object (e.g., an

insect flying) and the surface above which it is placed, whereas the term *elevation*, defined as the distance from any point of the surface relative to the average sea level, would be the correct choice in the descriptions presented in the text.]

The split of *Saphydrus/Enigmahydrus* would be, according to the results in Chapter 2, compatible with a separation during the Oligocene. Especially considering that the ages given in Chapter 2 may be overestimated (if one accept my arguments for biases in the molecular clock analyses), would they still qualify as palaeoendemics, as proposed in the manuscript?

Finally, **Chapter 6** describes larval features of *Rygmodes* (Cylominae) species and their ecology both as adults and larvae, reporting an unusual life-style for adults consisting of generalist pollen-feeding. Once again, this is a truly collaborative study with different authors combining their different areas of expertise to merge taxonomy with evolution and ecology to describe an intriguing, very original system. In this case, the candidate (second author) contributed molecular data for the molecular taxonomy of *Rygmodes*. Among the results of the study, a morphological analysis of pollen from the guts of adult beetles revealed a trophic association of these species with as many as 14 plant families. Were molecular tests considered as a possibility to investigate these associations?

After reading the six chapters of the thesis, all of them coauthored articles or manuscripts, I really missed a **general discussion** that elaborated on the reflections of the candidate, with his own words, and not coached as part of a team, on the topics that kept him busy over the past few years, and his own opinion on the progress that his work may have contributed. The thesis touched several general topics, including molecular and 'classical' taxonomy; phylogenetic systematics and molecular clock theory and practice; biogeography and global patterns of distribution; life histories and ecology; and, even some degree of functional analysis when trying to predict trophic behaviors of larvae, for example. At the very least, if a general discussion for all these topics were too ambitious, it would have been neat to have a bulleted list of major **conclusions** and achievements from the PhD project.