Institute of Animal Physiology and Genetics CAS, v. v. i.
Rumburská 89, 277 21 Liběchov, Czech Republic
URL http://www.iapg.cas.cz
Alexandr Sember: E-mail: sember@iapg.cas.cz; tel.: **+420 315 639 575**

Opponent's review of Ph.D. Thesis

Thesis title: **Evolutionary dynamics of chromosomal rearrangements within the family Pipidae**

Thesis author: **Mgr. Nicola Reinaldo Fornaini**

Supervisor: **Ing. Martin Knytl, Ph.D.**

The presented Ph.D. thesis aimed to investigate the subgenome evolution along with chromosome rearrangements and repetitive DNA patterns during ongoing re-diploidization in polyploid representatives of the anuran family Pipidae, namely *Xenopus*, *Silurana*, and *Hymenochirus*. The thesis encompasses three published papers (all of them original research articles) in respected journals (Q1, Q2), where the candidate is once a first author. I would like to acknowledge that the candidate achieved and completed very interesting scientific outputs within the frame of 4–4.5 years.

Thesis objectives and outputs

The thesis is conciped as a compilation of thematically related papers (but I will come to one specific issue later). The main objective of the thesis was to assess the chromosomal dynamics in allopolyploid genomes in selected pipid frog species using a suite of molecular cytogenetic methods. In the first publication (Knytl et al. 2023) the authors showed, using the chromosome painting probes prepared from diploid *Xenopus tropicalis*, that they stain more intensely one (more conserved) of the two subgenomes of tetraploid *X. calcaratus*, providing thus satisfactory evidence of its allopolyploid origin and demonstrating differential evolution of subgenomes compared to previous study on *X. mellotropicalis*. The authors also found no evidence for a chromosomal translocation that occured in the ancestor of *X. mellotropicalis*. The candidate further recently analysed genomic data on *X. borealis* which suggest that particularly the coding regions may remain functional in both subgenomes in *X. borealis*, providing thus intriguing contrast to subgenome divergence data as revealed on cytogenetic level in *X. calcaratus*. The second paper (Fornaini et al. 2023) provided insights into dynamics of several repetitive gene clusters (rDNA, snDNA, hisDNA) in post-polyploid genome in five *Xenopus* species and including comparison with diploid *X. tropicalis*. The last publication (Gvoždík et al. 2024) deals with African pipids and the key finding is that while captive *Hymenochirus* sp. is diploid, wild *H. boettgeri* is allotetraploid and hence likely represents a separate taxonomic unit apart from *H. boettgeri.*

Thesis formal structure and quality

While the candidate achieved indeed valuable scientific outputs, the thesis itself, unfortunately, gives an impression it was written in a hurry, without enough time devoted to proper delineation of thesis structure and the quality of writting. I can only speculate whether there was sufficient communication between the student and his supervisor but something was definitely not going well with the thesis preparation. The structure and organization of the thesis is atypical and confusing. After Abstract, Literature overview and Material and Methods, we proceed directly to Discussion, followed by Conclusions. Only then, within the segment named „Accompanying sections“, we may find a list of papers (with some subtle formatting errors) along with the specified contribution of the candidate. And here comes another confusion. This section (as well as the Contents section and inserted papers) gives an impression that the work builds on five papers. However, only three of them are related to frogs, while the remaining two are devoted to *Carassius* fish. Bearing in mind the topic of the thesis and lack of any *Carassius*-related passages in Introduction and

Discussion, I pressume the thesis is, in fact, composed of three papers. This amount of published work still fulfills, to my opinion, the doctoral study program requirements. However, the two *Carassius*-related papers, while they also show the candidate's productivity (and perhaps it should be acknowledged the candidate did not try to forcefully connect them to the thesis topic) would be placed apart as „Publications not related to the thesis“. Futhrermore, summary of the papers has been, for some reason, provided as independent supplement, not inserted to the thesis main body. Only later I found that the same text provided as „Summary of results“ is served in the thesis as „Conclusions“. To my opinion, the correct list of publications (along with the contribution of the author) and their brief summary, followed by an array of inserted publications (now placed at the very end of the thesis) should be placed after Material and Methods section, to facilitate smooth understanding of the subsequent Discussion. This might help the reader to be well oriented in the work done, its scope and main questions investigated before the Discussion and Conclusions parts start. It would be also helpful if the author keeps the marking of the papers (e.g. Paper I-III) interlinked between the sections so that the reader always knows which thesis-related paper is now being discussed. Especially as Discussion starts with the sentence "In the first Paper". But we know nothing about what the first paper is. If I look to the appropriate section, the first paper is about *Carassius* cytogenetics.

Furthermore, many parts of the thesis are uncarefully formulated, with long sentence constructions and/or with unclear meaning or uninteliglible wording, lack of proper explanations and sometimes missing relevant information and references, which altogether makes the text hard to digest. Regarding English, sometimes important words such as verb is missing in the sentence, or singlular vs. plural forms are being used incorrectly. Various other mistakes also occur repeatedly (see below). At least typographical errors are not so frequent.

I will list only some examples of what is otherwise repeatedly occuring in the text:

Mistakes in grammar: The level of susceptibility for polyploidy it’s different according to the species (page 12). This it’s just an hypothesis (page 36); …but a previous studies (page 15).

Awkward wording: transposon transposition (page 36)

Examples of uninteliglible sentence:
page 39, Conclusions: Alternatively, if two independent allotetraploidization events occurred, probably the translocation occurred either diploid ancestor of the common one of *X. mellotropicalis* and *X. epitropicalis*, the most recent probably.
Page 15: Pipa has been hypothesized to occur chromosome fission (Mezzasalma et al., 2015).

Examples of extremely long sentence constructions, hard to follow, sometimes with unclear meaning:
page 33: Unfortunately we were not able to localize 5S on *H. boettgeri* because the first attempts didn’t show us any result and, unfortunately, the continuous cycle of staining and destaining on the slides, led to chromosomal structure damaging, making the detection of any signal impossible, fortunately on *Hymenochirus* sp. we didn't register this problem.
page 37: We first tried to start the cell line using organs but this attempt failed, resulting in very frequent contamination or a low growth rate, so we started to use the hand limbs from tadpoles, with a higher success rate, but the procedure is long and needs high precision and often the final concentration of cell suspension is low, so we had to make sure we had a high initial amount of cells, but it took even more time, leading to potentially new contaminations.

The main text contains 17 figures, nine of which are placed in Discussion and directly taken from the publications in question. I think this was unnecessary since the publications are also attached. Remaining figures mostly complement the text appropriately e.g. by showing the phylogenetic relationships and diagnostic morphological features of studied frogs and related species. I appreciate the phylogeny of the

genus *Xenopus* in Fig. 7, however, it would be useful to improve the scheme by adding the column with known chromosome characteristics (at least 2n) to make the figure more informative for the context of the thesis. I, on the other hand, find unhappy the placement and content of Figure 6. Not only it is unnecessary to narrow the processes of polyploidization and re-diploidization to plants, but also the content of both pictures A+B is not explained either in the figure legend, nor it adheres to the main text. Which brings me to the structure of Introduction. The candidate chose to first introduce the model frog species, with many details of their applications which is interesting but so far we do not know what will be important for the thesis objectives. Then the author jumps to the concept of polyploidy. Later he describes the Pipidae family and returns back to detailed description of frog species. I find this structure unhappy as the more general evolutionary concepts and questions, along with the identified knowledge gaps, should be introduced first and then the text should focus to the model organism and why it is suitable for study of the chosen scientific questions. The information flow in certain sections also deserves caution. In chapter 1.3 about polyploidization (starts on page 12) only very later we repeatedly receive the information what is auto- and allopolyploidy (e.g. end of pages 13 and 14). This is something that should be delineated in the first paragraph. Then it would make more sense to explain mechanisms that facilitate polyploidy emergence and subsequent polyploid genome evolution. The problematics of unreduced gametes along with various mechanisms leading to them (not properly delineated by the candidate) are, for example, neatly summarized in Mason and Pires 2015 (doi: 10.1016/j.tig.2014.09.011). Several key concepts and mechanisms following the re-diploidization process are not adequately introduced such as genome downsizing (e.g. Wang et al. 2021; doi:10.1111/tpj.15363) and subgenome dominance (Bird et al. 2018; doi: 10.1111/nph.15256). To my opinion, these concepts are important for the interpretation of the achieved (as well as future) results. The polyploidization section also sometimes contains huge segments of text with a dense information content but without any (or just one/few) supporting reference(s), which applies especially to last two paragraphs of this section (page 14).

The author cited 124 literature resources. I found some discrepancies between the bibliography and in-text citations. There is Höbel and Fellows (without year) and Miranda-Ribeiro 1937 which are not listed in the Reference list. Also some papers (such as e.g. Irisarri et al. 2011 and Bredeson et al. 2024) are listed as more publications (e.g. Bredeson et al. 2024 as a, b, c) but they mostly have the same doi number and, in the text, there is no distinction between a, b, or c. Conversely, Sokol 1969b is not present in Reference list. The Reference list also contains, on the first page, two randomly inserted doi numbers which belong to papers (Nenni et al. 2019, Zhou and Gui 2017) being also cited in a standard way more downstream.

Here I have several specific remarks, suggestions and questions. I highlight in bold (text segments or page + number) a portion of them which I would like to ask to be preferentially adressed during the thesis defense:

Page 8: Unclear statement: I do not understand in the context of *X. tropicalis* and *X. laevis* genomes what the author means by „fully sequenced“. What level of genome assembly has been achieved?

Page 9: I would not emphasize the phylogenetic distance of frogs to human. There are numerous other important models phylogenetically closer to humans.

Page 12: I do not agree with this (also not very happily formulated) statement: „Polyploids have a relatively high frequency in flowering plants, this can be also found in some species of frogs and fishes. However, in higher vertebrates, as amphibians, reptiles, birds, and mammals, animals that are characterized by ….. a more complex structure compared to lower vertebrates like fish, they don’t seem to have an high tolerance to polyploidy very well.“

There are many examples of natural polyploidy in amphibians (both sexual and asexual) and in parthenogenetic reptiles. Moreover, I do not understand, upon which reasoning/based on which resource

the author considers amphibians and reptiles as higher vertebrates. I find it misleading and incorrect. Also my opinion is that it is better to categorize vertebrates to cold-blooded and warm-blooded.

Page 13: This statement not only contradicts the one I mention above but also it is problematic in other way:“ Polyploidy is common in plants, fish, and frogs, indicating clear fitness advantages.“ Tolerance to polyploidy does not necessarily mean ultimate fitness advantage. There is still ongoing debate on this matter (Barker et al. 2016; https://www.jstor.org/stable/newphytologist.210.2.391; Laurent et al. 2017; doi: 10.1371/journal.pone.0176384; Barker et al. 2024; doi: 10.1002/ajb2.16395).

Page 13: For polyploidy linked with asexual reproduction, I miss very important relevant sources such as e.g. Lamatsch and Stöck 2009 (doi:10.1007/978-90-481-2770-2_19), Stöck et al. 2021 (doi:10.1098/rstb.2020.0103).

Page 14: „Unlike, for example, teleost fishes where whole genome duplication became an established evolutionary process for the entire group, polyploidy in amphibians has arisen independently on multiple occasions within different families (Schmid et al., 2015).“

- This is incorrect interpretation. In teleost and also non-teleost fishes, there were multiple additional polyploidization events at various taxonomic levels. Not only teleost-specific whole-genome duplication at the base of teleost lineage.

Regarding the Discussion part:
The relevant paper which is being discussed in chapter 3.1 (i.e. Knytl et al. 2023) is not cited on the entire page 20. In this context, it is difficult to understand several statements including that the authors used "different combination of cytogenetic techniques". It is further not clear from the text whether the study Knytl et al. 2023 used chromosome paints from all *X. tropicalis* chromosomes and whether all of these probes generated brighter signals on one subgenome of *X. calcaratus* compared to the second one. Figure 9, which might be of great help in this regard is not properly explained in its legend. **Furthermore, in this section (page 20) I do not understand how cytogenetics may explicitly show that two genomes differ in gene expression. This is a misleading statement. Similar statement is provided in subsection 3.5 (page 38): „ … and let us confirm our hypothesis that both subgenomes were expressed, as seen with ZOO-FISH“.** Yes, we can see the subgenome divergence at chromosomal scale but this type of data does not provide any direct evidence of gene expression patterns. I, nevertheless, appreciate that the candidate and his team achieved such a nice result to distinguish subgenomes by differental intensity of painting probes. Both provided explanations for this observations are possible but I adhere more to a different rate of subgenome differentiation than to the quality of WCP probe with more dissected chromosome copies. The related freshly analysed genomic data on *X. borealis* presented in the subsection 3.5 (last part of Discussion) further suggest that while there might have been a rapid turnover of repeat content on the chromosomes of one subgenome in tetraploid *Xenopus*, particularly the coding regions may remain functional in both subgenomes. This would be very interesting to discuss in the context of subgenome dominance phenomenon, perhaps in the future work. **Still in this section, on the page 23, I do not agree that the only explanation for highly divergent subgenomes in *X. calcaratus* is that different polyploidization at different timepoint has occured. Re-diploidization of genomes might have highly different dynamics and outcomes in different allopolyploid species, so, we cannot fully discard that still a single polyploidization might have occured here until proven otherwise by more detailed analyses.**

Part 3.2 is also linked to the paper by Knytl et al. (2023) but again we cannot extract this information explicitly from the text (paper not cited here). The clue is only in the legends of Figs 10 and 11. This part is about gene mapping via TSA-FISH and shows that a specific chromosome translocation found in *X. mellotropicalis* is not present in *X. calcaratus*, which is an interesting finding. **On the page 25 I do not understand how the author imagines to reveal deletion on chromosomes by WCP**. Maybe a really large deletion would be possible but the same linkage group may undergo size contraction or expansion just via

repetitive DNA dynamics and so the size difference of particular homeologous chromosome in different species cannot be ultimately taken as proxy to infer deletion. I further do not understand the sudden placement of general paragraph about nucleolar organizer regions, placed without any context. This part belongs rather to Introduction. Here, I do not agree that NORs typically appear as a single pair (along with the obsolete reference Schmid et al. 1987). There has been indeed a huge load of research on rDNA dynamics since 1987 showing highly diverse rDNA patterns, summarized e.g. in Animal rDNA Database (Sochorová et al. 2018; doi:10.1007/s00412-017-0651-8). **In the context of allopolyploidization, a phenomenon of nucleolar dominance would be also relevant to be mentioned**. Strikingly, Figure 12 is not from Fornaini et al. 2023 as is mentioned in the figure legend but from Knytl et al. 2023.

Subsection 3.3 is linked to paper Fornaini et al. 2023 (again, clues only in the legends of Figure 13 and 14, because the announced "second paper" is, according to the publication list, Knytl et al. 2023). Compared to the other subsections, the title here is very general, reads like "methods" and does not introduce any scientific issue to be solved. In this section, I do not agree that short read sequences present challenges in repeat quantification. Bioinformatic pipelines such as RepeatExplorer are able to quantify repeats and characterize them from short reads efficiently. Moreover, this way the repeat qualtification is also not "costly" as the candidate states more downstream. I also do not understand why for such a broad statement that repetitive DNA mapping has been used in various animal groups, the supporting references are few case studies instead of robust reviews. Tandem repeats of genes coding snRNA should be properly termed snDNA which the candidate suddenly starts using only from page 36 onwards. **I miss in the text the information about what are the general trends of snDNA and histone genes evolution/dynamics and comparison how reliable are these tandemly repeated multigene families (along with rDNA) as cytogenetic markers (i.e. whether they display low/high variability in site numbers which reflects vs. does not much reflect the trajectory of karyotype rearrangements etc.).** Then it is hard to understand the arguments further in Discussion. On the page 30 it is particularly misleading to highlight that one insect lineage has multiple sites of histon H3 cluster because this is certainly rather exception than a standard pattern found in animals (see, e.g. Provazníková et al. 2021; doi:/10.1038/s41598-021-91665-7). The candidate also sometimes makes no clear distinction between repeat copy number (not analysable by FISH) and physical number of loci on chromosomes detected by FISH. Change in the number and position of snDNA or other repeat cluster may be facilitated not only by direct transposition but also by ectopic recombination events (the latter with the help of various repeats, not necessarily only TEs). **I find it also too exaggerated to state "Our research uncovered a fascinating genomic connection between different repetitive elements and rRNA, potentially providing insights into genome organization and evolution."** First, the observations of co-localization of various repeats are widespread. Second, telomeric regions are generally more vulnerable to ectopic recombination, which itself might explain the observed patterns and it has been widely reported/discussed in many studies previously.

Chapter 3.4 refers to paper by Gvoždík et al. 2024 dealing with cytogenetic analysis of one male of *Hymenochirus boettgeri* from natural population and six individuals (3 males, 3 females) of *Hymenochirus* sp. from aquarium strain. The main outcome is that while captive *Hymenochirus* sp. is diploid, wild *H. boettgeri* was found to be allotetraploid and hence likely represents a separate taxonomic unit. In the thesis the author does not specify that the observed B chromosome was present in both sexes of *Hymenochirus* sp. leaving thus a space for speculations. The statement on page 34: "It's known that cells and embryos of polyploid species are smaller than diploid ones (Miller et al., 2023)." is a mystification. The cited study by Miller et al. (2023) explicitly states: "Despite its large genome, *X. longipes* eggs are slightly smaller than those of *X. laevis* at 1.1 mm in diameter." And it is meant as a specific situation, not a standard. Lastly, the ancestral karyotype cannot be expressed as "primordial karyotype" (page 36).

In Conclusions, I do not understand why „the findings underscore the slow nature of genome evolution in these species“? I thought that there are data showing rather rapid evolution of one subgenome of *X. calcaratus* (?) Finally, it is very problematic to talk about the pace of genome evolution without any information about the timing of the allopolyploidization event(s) (at least the text seems not to provide such information).

Questions:

1) Since the publications have passed a demanding review process, it is not my place to subject them to further critical analysis. But I would like to ask what was the most discussed or problematic point (if any) during the reviewing process of these four papers?

2) Regarding the statement on the page 17:“ Some hybrids retain the same ploidy level as their parental species,..“ Under which conditions this usually happens? I mean generally, not only in clawed frogs.

3) I have a question to statement on the page 26: "In the *Xenopus* subgenus at least eight independent episodes of allopolyploidization occurred, with each subgenera experiencing at least one allotetraploidization event." This is extremely important information which is somehow burried in the text, moreover without any supporting reference. Also the meaning is a bit ambiguous. I would like to ask for better explanation of this statement along with supporting reference(s). I also cannot find in the text any information about estimated timing of these allopolyploidizations which is also very relevant when post-polyploidization genome dynamics is being traced.

4) I would like to ask the author how (upon which reasoning) he came to the conclusion (page 31) that "Our study highlighted that, in addition to translocation, inversion, deletion, and degeneration, the reduction and expansion of tandem repeat copy numbers play a crucial role in the evolutionary processes following allopolyploidization."? Isn't the site number change of repetitive DNA rather a consequence of the re-diploidization processes?

5) I would like to ask the author if he may briefly describe the general mechanisms shaping the genome during the re-diploidization process.

6) On the page 36 a do not understand this statement: „Consequently, we would expect a diploid-haploid cycle and crossing-over in tetraploid *H. boettgeri*.“ Could the author explain more clearly what does it mean?

To conclude, the candidate achieved interesting and scientifically sound results in the field of genome/karyotype evolution of the post-polyploid genomes of the anuran family Pipidae and demonstrated the ability to conduct creative scientific work. From this viewpoint there is no doubt that a sufficient amount of work was done to achieve a scientific degree Ph.D. However, regarding the quality of the thesis, the candidate did not convince me that he understands well the scientific background and concepts linked to his research topic and that he achieved sufficient level of cultivated thinking to be able to organize the text and write clearly about the targeted scientific problems and results interpretation, with a proper structure and information flow of the text. Therefore, it will be not an easy task for a committee but let's see how the candidate will perform during the thesis defense.

In Liběchov, 7.1.2025 Mgr. Alexandr Sember, Ph.D.