PhD thesis: **Drivers of karyotype evolution in Lepidoptera**
Candidate: **Irena Provazníková**
Review: **Diogo C Cabral de Mello**

General comments

The PhD thesis presented by Irenka Provazníková was focused on the understanding of karyotype/genome evolution on Lepidoptera. It was mainly focused on the understanding of multigene family organization and evolution and origin of neo-sex chromosomes. The data presented is quite interesting and clearly is a good advance on the field, not restricted only to Lepidoptera, but, also at some point, in insects as a whole (and animals as a whole). It is reflected on the journal in which the manuscripts were published, or are in via of acceptance.

The four chapters are well organized and they are connected (chapters I and II; and chapters III and IV), by both biological questions and methods used. The methods used are cutting-edge in the field of Cytogenomics, allowing detailed description of the data, giving interesting insights about the biological question addressed. My only general criticism is that, in my opinion, the title does not describe clearly the content of the work, which was much more focused on the evolution of multigene families and sex chromosomes. In this way, I suggest a more focused title, reflecting the work content, something like: "drivers of multigene families and neo-sex chromosome evolution on Lepidoptera."

The presented work also clearly shows the ability of the student to work in collaboration and capacity to go ahead in deeper details for specific biological questions, using multiple technical approaches. Her involvement on the generation of the data is clearly evident by the authorship position on the four works, as she is the first author of two chapters and shares first authorship in the other two publications. From a technical point of view, it is clear that along her PhD studies Irenka developed skills to work with the cutting-edge methods on Cytogenomics, including computational analysis, chromosomal analysis and molecular analysis. It means that she will have more probability to think of multi-approach/complementary ways to answer biological questions.

Finally, the participation of the student as collaborator in other works out of her thesis, also deserves attention, because it indicates a good capacity to collaborate on side projects, which is important for advance in the career.

Introduction/Synthesis and perspectives

The introduction is well written, giving to the reader the general necessary information to understand the manuscripts. The only lacking information in more detail is about structure and evolution of multigene families. The student included classical references and modern ones, recovering the most part of the history and advances in the field on Lepidoptera chromosome evolution.

The final part of the thesis made a good general integrated discussion about the data and presented clearly future ways to advance on the studies of multigene families and sex chromosome evolution on Lepidoptera.

Chapter I

In this chapter Irenka and colleagues tested some available and routinely used chromosomal markers on other insects on Lepidoptera chromosomes. The work was well designed and a clear advance about the reliability of chromosome markers to study Lepidoptera karyotype evolution is presented. Moreover, it is clearly discussed the patterns of chromosomal (number and location of clusters) evolution of these markers on Lepidoptera. Besides the FISH analysis, it was also studied in more detail the number of copies of the studied repeats and their genomic organization, based on wet lab techniques. The study only lacks deeper genomic information about the organization of the studied multigene families. It is probably related to the period in which the study was developed, probably before the availability of some sequenced genomes using long reads, that could allow to advance on the topic. I think that future studies using the available sequenced genomes could bring deeper insights about multigene family evolution in Lepidoptera. It was well done in chapter 2, but only for major rDNA.

Questions:

1. Why did you not check the abundance of the studied markers using sequenced genomes?
2. Why did you include the species G. pellucidus as an external group? Why did you not select a different species? Or why did you not include more species as an external group on the analysis?
3. You estimated the number of copies for U snDNA genes and H3 histone. Why did you not include the 18S rDNA for a comparative analysis? It could give insights about the number of copies that give a signal on "regular" FISH experiments. What about other histone genes, did you try to check the number of copies?
4. Analysis of genomes sequenced by long reads could give ideas about organization of these repeats.
5. Opinion about the recent publication (Dutrillaux et al., 2023). According to these authors:

- "Our improved conditions for basic chromosome preparations, here applied to 6 butterfly species belonging to families Nymphalidae and Pieridae challenges the holocentricity of their chromosomes: in spite of the absence of primary constrictions, sister chromatids are recurrently held together at definite positions during mitotic metaphase, which makes possible to establish karyotypes composed of acrocentric and submetacentric chromosomes.";
- "It allowed us to observe not only atypical monocentric chromosomes in mitotic cells, but also figures hardly compatible with the occurrence of an inverted meiosis";
- "Consequently, centromeres of butterfly chromosomes are often located in or at close contact with telomeric heterochromatin, whose composition differs from that of centromeric heterochromatin [DeBaryshe and Pardue, 2011] and whose instability may

be responsible of their numerous rearrangements. In other words, the large numerical variations of chromosomes observed in butterflies may be better explained by their telomere/centromere instability than by the facilitated transmission of rearranged chromosomes through inverted meiosis, whose occurrence is challenged here."
Does it make sense for you?
6. In the discussion the authors say: "Our data show that the multiplication of the major rDNA cluster occurs in multiple lepidopteran families and via different mechanisms, without any clear evolutionary pattern" Which types of mechanisms?
7. Why did you not try to recover the intergenic spacers for 5S rDNA and U snDNAs?
8. Why did you use in some cases male cells and in other female cells for FISH? I know that the obtaining of female pachytene is more challenging in comparison to males, but using male cells you miss the opportunity to check the presence of clusters on the W chromosome. How could you face this challenge to check the possible occurrence of clusters on the W chromosome using other molecular/genomic techniques?
9. What is your general hypothesis. Does the karyotypic evolution reflects on the evolution of multigene families or/and vice-versa?

Chapter II

In this work Martina, Irenka and colleagues performed an investigation about the genomic organization and evolution of major rDNA using Lepidopteran species as model, joining genomic and chromosomal data. The manuscript is quite interesting and allowed great advancements about the organization/evolution of major rDNA. The selected species were adequate to solve the main questions and this chapter has good connection with the first chapter, as it advanced on questions that arose in this previous work. It took the advantage of short and long genome sequencing, besides the cytogenetic tools. The manuscript is submitted to Genome Biology and Evolution, a prestigious journal on the genomic field.

Chapter III

Here Leonela, Irenka and colleagues tested the hypothesis of occurrence of neo-sex chromosomes on a superfamily of Lepidoptera, i.e. Gelechioidea. The experimental work involved well established FISH analysis for detection of W chromosomes in Lepidoptera and the analysis of sex dosage (number of copies) for some genes, located on chromosomes putatively involved on fusion with sex chromosomes. It means that the experimental design of the work was well performed, allowing to answer the main question. Besides the evidence of one or two sex-autosome fusion, the authors made a general discussion about the incidence of neo-sex system on Lepidoptera that has female heterogamety, in comparison to neo-sex system on groups with male heterogamety. They suggested that the paucity of neo-sex chromosomes is not an intrinsic feature of female heterogamety. According to Google Scholar the manuscript has until now 19 citations. Although the most citations were in works that studied chromosomes of lepidopterans, the manuscript was cited on works with general ideas about sex chromosome evolution, revealing its important contribution to the field.

Questions:

1. What do you think about generalizations based on restricted sampling?
2. Considering that: "It has been shown that chromosome fusions are not random in this insect order since independent fusions observed in distant species involve the same small and repeat rich chromosomes (Van't Hof 2013; Ahola et al. 2014)." Could you consider that the origin of neo-sex chromosomes in Gelechioidea could be independent, just occasionally involving same chromosome? Which type of analysis do you propose to have a more consistent conclusion?
3. On *D. daucella* you considered that the *Pix* gene on the autosomal chromosome is product of a secondary translocation. Is it possible that before the ancestral chromosome fusion that originated the neo-sex chromosomes the autosome involved in the fusion suffered a fission, and only part of the LG7 was fused with the ancestral ZW? Physically where are the genes studied by qPCR for LG7 (*Pix* and *Chit*)?
4. Based on actual technology and sequencing advances, if you have the opportunity to advance in the question presented in this work, what kind of analysis do you consider important to be done?

Chapter IV

Similar to chapter III, this chapter delayed the evolution of sex chromosome system on Lepidoptera, using as model Yponomeuta species and related outgroups. The methods also involved cutting-edge technologies on Cytogenomics, reliable to identify sex-related genes, allowing the inference of sex-chromosome evolution. The analysis points to ancient sex chromosome system origin, not related to the host shift in Yponomeutidae. Although the data is highly interesting and has potential to be published in a good journal, the discussion section is apparently quite preliminary. Moreover, additional data could be obtained to have more strong evidence about the main conclusions and the data could be better described.

Questions:

1. Did you check mitotic chromosomes of females using a telomere probe to be sure that the telomere is not on the neo-W chromosome (as observed for example in *Diatraea saccharalis*)?
2. You stated: "Our results thus show that formation of the multiple sex chromosomes probably did not play a major role in the host shift of European ermine moths as they predate the split of the *Yponomeuta* and *Teinoptila* genera". Please take care with this general conclusion. You know that these genus (at least most species studied) have multiple sex chromosomes, but not necessarily the same chromosomes are involved in the fusion. You have data of qPCR that shows the involvement of LG2 of B. mori on the origin of multiple systems only for three species. I suggest a broader analysis by qPCR in all species to include on the work, checking if the same chromosomes are involved in the fusion. This is also important for the main conclusion of the work: "we hypothesize that sex chromosome turnover in Lepidoptera could be driven by sexual antagonism". If a different chromosome is involved in the fusion, this assumption could be limited for a few species.

3. About the fewer occurrence of neo-sex chromosomes on female heterogamety on vertebrates. There are some examples of birds with neo-sex chromosomes, as for example:
https://www.nature.com/articles/s41467-022-28585-1;
https://www.nature.com/articles/hdy201170;
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6736294/;
https://onlinelibrary.wiley.com/doi/full/10.1111/jeb.14096
Please let me know your opinion if you think that it is a real pattern (fewer occurrence of neo-sex chromosomes on female heterogamety) or it could be biased by the fewer studies using more advanced approaches (that allow more precisely the identification of neo-sex chromosomes) on the vertebrates with female heterogamety.

Final recommendation

Based on the quality of presented data and on the clear advance on the field I ***RECOMMEND*** the defense of the present thesis.

Diogo Cavalcanti Cabral de Mello:05961889432

Assinado de forma digital por Diogo Cavalcanti Cabral de Mello:05961889432
Dados: 2023.02.13 21:39:07 +01'00'

Prof. Dr. Diogo C Cabral-de-Mello