

PhD Thesis Acceptance Report
Research Discipline Council of Biological Sciences
Jagiellonian University in Kraków

Candidate's name and surname: Evgenii Baiakhmetov

PhD Thesis Title: Inferring hybridisation and introgression processes within *Stipa* (Poaceae)

Thesis Supervisor: prof. dr hab. Marcin Nobis, Jagiellonian University (Kraków)

Co-supervisor: dr Polina D. Gudkova, Tomsk State University (Tomsk, Russia)

Reviewer: dr hab. Michał Ronikier, W. Szafer Institute of Botany, Polish Academy of Sciences

THESIS EVALUATION

1. Scientific merit of the thesis

a. Originality of the research:

The PhD thesis of Mr. E. Baiakhmetov contributes to the long-term research on the taxonomically intricate genus *Stipa* (Poaceae) and addresses the problem of interspecific hybridization, considered a major key for understanding this group, for the first time with extensive use of molecular tools. The articles included in the dissertation undoubtedly bring a significant input into the knowledge on the taxonomy and phylogeny of the studied genus, and particularly on revisiting the role of hybridization and introgression in shaping its taxonomic diversity (hitherto, with one exception, assessed based on morphological traits only). Based on combined molecular and morphological analyses, the candidate was able to retrace hybridization events in populations of several species characterized by smaller or greater phylogenetic distance and locally occurring in sympatry. He revealed several hybrids and formally defined a new nothotaxon but also assessed the incidence of reticulation events, and analysed correlations between genetically revealed hybridization/introgression status and morphological settings of individuals, hence actual expression of hybridization in morphological traits – a starting point to re-evaluate earlier, morphology-based investigations on *Stipa* hybrids. The important major conclusion of the study is that hybridization is confirmed as an important process in *Stipa* but a combined morphological and molecular approach is indispensable for its appropriate evaluation and for drawing pertinent taxonomic inferences. Besides the main axis of the work focused on the above, it also provides a valuable contribution to – poorly recognized – Central Asian biodiversity, and – as a robust case study – also to the general knowledge on hybridization as a driver of evolution and population genetic processes. Last but not least, the articles provide an important methodological asset for future research into the *Stipa* taxonomy and evolution (see also comments on particular papers below). Overall, I consider the research as scientific merit of the thesis as significant. It presents a limited novelty at a general level but high novelty for the studies of the genus *Stipa*. Although all papers were published in a very recent time perspective (2020–2021), the first of them (Article 1) has already been cited 7 times by external authors (Research Gate, checked 06.06.2022), which shows the interest of the international botanical community in this research.

b. Scientific merit of the chapters / articles:

The thesis includes three articles published in high-rank (Q1 in *Journal Citation Reports*) open access, interdisciplinary scientific journals. All of them are based on comprehensive data and provide scientifically sound outputs; below I summarise their main scientific merits.

Article 1 presents a case study based on several *Stipa* spp. co-occurring in a spot of the Central Asian mountains. Its main merits include: (i) a comprehensive morphological and molecular evidence for a hybrid origin of one taxon/morphotype observed in this area – a novel contribution to taxonomy of *Stipa*; (ii) demonstrating hybridization among a pair of phylogenetically relatively distant species; (iii) testing the possibility of using the NGS-based genotyping approach DArTseq in the studied genus.

Article 2 presents outputs of a long-read PacBio sequencing of *Stipa capillata*. Its main merits include: (i) building a major methodological background for future research by providing an annotated draft genome of *Stipa capillata*, including arrays of protein-coding genes, organellar genomes, characteristics of transposable elements and identification of potential SSR markers; (ii) a phylogenetic reconstruction (based on new and on previously published data), with divergence time estimation for the genus *Stipa* based on molecular clock was presented, an important aspect although somehow loosely related to the presented aim of the article.

Article 3, finally, relying on two earlier papers, provides the first comprehensive analysis of hybridization patterns in five *Stipa* species based on a significant sampling from several hybrid zones in Central Asia. Its main merits include: (i) a detailed analysis of past hybridization/introgression events across the studied populations using genome-wide molecular markers; (ii) an attempt to infer potential incidence and dynamics of hybridisation (and backcrossing) based on studied populations and showing traces of reticulation events in ca. 20% of samples; (iii) analysis of hybridisation influence on expression of morphological traits, (iv) revealing a weak relevance of morphology to infer hybridization (assignment of both hybrids and pure taxa) in the genus due to low number of conservative features, hence the need to entirely re-evaluate traditional concepts of taxa and nothotaxa; (v) showing that presence of frequent hybridisation is correlated with a relatively recent (mid/late Quaternary) temporal frame of evolution of the studied taxa.

2. Substantial merit of the thesis

General introduction provides a brief but adequate background to the research topic including a historical account of taxonomy of *Stipa*, potential role of hybridization and state of research on the genus. It is useful and well planned except its first part (see critical remarks below), and shows a good orientation in the taxonomy of the studied genus and earlier research although no larger context of this research (going beyond this specific subject) is outlined. Objectives of the thesis, presented in a synthetic way in the introductory part and then within particular articles, are clear, simple and directly related to advancement of the knowledge on taxonomy and phylogeny of the genus *Stipa*. This gives a straightforward and adequate framework for the thesis.

The work is based on a very good sampling for a study based on Central Asia. Although only partly a direct merit of the candidate, it gives an excellent basis for the research. A wide array of methods and data analysis approaches are also a strength of the thesis, which

reflect an excellent scientific toolbox elaborated by the PhD candidate. As far as I could retrieve from Materials & Methods chapters of the articles, wet-lab implementation of library preparation for Illumina and PacBio was mostly subcontracted from commercial companies and the doctoral student was more involved in comprehensive treatment of resulting data rather than in the lab itself.

The studies are based on combined morphological and molecular data to confront expression of morphological traits with genetic structure being the key aspect of the work. The molecular analyses introduce state-of-the-art methods with genome-wide sequencing markers, and drafting the whole genome based on long-read PacBio sequencing. Data are explored using clustering analyses, population-level analyses, time-calibrated phylogenies, and population genetic methods. Results are adequately discussed in particular chapters and provide rather straightforward responses to the aims formulated in the research. The dissertation is closed by a 'Final conclusions and perspectives' chapter which is more focused on perspectives, which leaves somehow a gap regarding a synthetic discussion of the whole thesis (see also critical remarks).

All the articles included in the thesis are first-authored by E. Baiakhmetov who is also the corresponding author in all cases. No separate author statements were provided but descriptions of author contributions included in published articles clearly show the important role of the PhD candidate in conceiving the studies, and his leading role in performing data analyses and writing manuscripts.

3. Layout and register

The thesis begins with one-page English and Polish abstracts, a general introduction (5 pages), a brief presentation of overall objectives of the thesis, sources of data and methodological framework, which mostly refers to descriptions of Materials and Methods in the papers. The main part consists of three published articles, preceded by a half-page introductions which present the key aspects of each of them, and accompanied by a rich supplementary documentation. The set of papers is coherent and well related to the title and overall aim of the thesis, except that Article 2 and its aim only indirectly concerns hybridization and introgression; however, it provides important data for future studies of these processes. The dissertation is closed by a 'Final conclusions and perspectives' chapter (6 pages) followed by a comprehensive list of references for the unpublished chapters of the dissertation. The structure of the thesis is clear and appropriate and it has been prepared very carefully. I only have minor remarks regarding the content of the unpublished chapters (see critical notes below). As the thesis consists of published papers, graphical aspects (figures) underwent editorial check and typesetting in journals and they are relevant and of good quality (a correction of figures has been published for Article 3). Language of the thesis is generally good with only minor grammar and syntax errors noted in both published and unpublished chapters.

4. Critical notes

The core part of the thesis consists of articles published in high-rank international journals, hence they underwent a review process and editorial procedures. They are based on solid data and have been well prepared and documented. The entire thesis including the unpublished chapters has also been carefully prepared. I only have a few minor critical notes and a few questions to its content and methodological aspects.

Remarks on the content:

- The general introduction provides a good background on earlier research into the taxonomy of *Stipa* and is relevant to the topic except the first page (p. 5), which addresses the general importance of the Poaceae family, distantly related to the thesis. In my opinion, it would be more beneficial for the reader to have, for example, a brief overview of hybridization as an evolutionary mechanism / problem in taxonomy instead.
- The dissertation objectives (p. 10) are generally simple but relevant, however, I find the point no. 2 trivial. Is it really necessary to test whether integrative taxonomy, i.e., analyses based on multiple data sources (instead of, for example, morphology or genetics alone) is a useful approach? I think it is obvious that the more sources of evidence we assemble, the closer we can get to the real evolutionary/taxonomic relationship.
- Having in mind the objectives of the thesis, I miss in the introductory part a brief rationale for the selection of case studies (why these areas, and these taxa were studied).
- In the set of articles presented, Articles 1 and 3 are very coherent and directly related to the title and objectives, while Article 2 is mainly a data paper which features a draft genome sequencing but is not directly related to hybridization in *Stipa*. Then, it is also not entirely coherent internally because it also contains a time-calibrated phylogeny based on selected loci which does not fit the presented aims and remaining content of the paper.
- The chapter 'Final conclusions and perspectives' is more of perspectives than of conclusions. I think a good point-by-point synthesis of the most important achievements would be useful for the reader. However, I also appreciate the focus on the perspectives of future research which is clear in this chapter. It testifies the enthusiasm of candidate for his research subject and rightly shows a larger future context to consider. I also find valid the more general thoughts of the author regarding, for example, the databases.

Remarks on methods:

- Testing the DArTseq genotyping method is presented as one of major objectives of the thesis (p. 10, no. 1). It is a rather uncommon method (beyond crop studies) among genome complexity reduction-based procedures which use next-generation sequencing and I wonder why has this method been selected for testing? Were any preliminary tests of other methods (such as ddRAD, often used for both population genetic and phylogenetic studies) performed beforehand, which directed further analyses toward this approach?
- In general, results are adequately discussed in the articles, including clear conclusions and uncertainties. I found one aspect, which remains unclear to me: in Article 3, outputs of fastStructure and Newhybrids analyses show a significant discrepancy in assigning hybrid/introgression status, which is important taking into account main objectives of the study. What could be the reasons behind these differences?

- In Article 2 a strict molecular clock has been used for time calibration of phylogeny, which implies a strong assumption. Has a model testing been performed which supported selection of this model?
- In the articles, there are a few minor inconsistencies, where specific issues are not coherent with the aims or sufficiently explained. For example: In Article 2, the aims are related to the draft genome assembly, while Results show also a time-calibrated phylogeny (based on a few loci), not directly related with the draft genome. In Article 1 (p. 19) it is not clear on what basis do the authors refer to *Stipa glareosa*, which was not included in the study set?
- Finally, in Article 3 a few individuals display very low levels of admixture (<0.1). I wonder whether measures for controlling cross-contamination were applied in the analyses? I did not find information on that in Materials & Methods and it is an important point to assess data reliability.

5. Final grade

The doctoral dissertation of E. Baiakhmetov brings valuable and novel scientific input into plant taxonomy. It shows that the candidate is able to efficiently prepare scientific publications and developed excellent competences and skills in biosystematic analyses, which include handling and analysing morphometric data, data from genome-wide genotyping and long-read sequencing, with a wide array of methods from clustering and population genetic methods to phylogeny reconstruction with temporal calibration. Taking all this into account, I, hereby, declare that the reviewed PhD thesis by **Evgenii Baiakhmetov** meets the criteria pursuant to art. 187 of Act of 20 July 2018 The Law on Higher Education and Science (Journal of Laws of 2018, item 1668, as amended) and request that the Research Discipline Council of Biological Sciences of the Jagiellonian University in Kraków accepts **Evgenii Baiakhmetov** for further stages of doctoral proceedings in the field of exact and biological sciences, in the discipline of biological sciences

YES/NO

Taking into account the quality of articles and rank of the journals in which they were published, the importance of the contribution to advance our understanding of taxonomy of the group, wide competences gained by the candidate during his PhD, and important research perspectives opened by the thesis outputs, I, hereby, request that the thesis is accepted with distinctions.

YES/NO

June 10th 2022

.....
date



.....
Reviewer's signature