Reviewer Report

Title: **Comparison of whole-genome assemblies of European river lamprey (*Lampetra fluviatilis*) and brook lamprey (*Lampetra planeri*)**

Version: Version 3 Date: February 20, 2025

Reviewer Comments to Author:

The authors of the manuscript entitled "Comparison of whole-genome assemblies of European river lamprey (Lampetra fluviatilis) and brook lamprey (Lampetra planeri)" provide haplotype-resolved chromosome-level reference genome assemblies of two sister species of lampreys. Genomes presented in this work showed high standards in terms of contiguity (chromosome level) and completeness. Overall, the manuscript is well explained and the data generated will be very useful, not only for researchers interested in this group of fish, but any interested in comparative genomics.

However, the framing of the study appears somewhat incongruent with the methodology employed. The authors mention that the main objective is to shed light on the species versus ecotype discussion, but it remains unclear whether whole genome assembly is the proper methodology for this. Given that only one individual per species was sequenced, it is unclear whether the data are sufficiently informative to robustly assess species boundaries. As they mention in the introduction, additional population-level sampling and comparative genomic analyses (e.g., population structure, admixture, or selection scans) could provide more direct insights into this question.

Therefore, a reframing of the primary focus of the study is suggested to emphasize the sequencing and comparative analysis of these genomes. The investigation into whether these taxa constitute separate species or ecotypes could then be presented as a secondary objective, to be explored further with more targeted methodologies.

Furthermore, while the introduction extensively discusses the ecological aspects of these taxa, the discussion section does not seem to fully integrate these considerations when interpreting the genomic findings. This manuscript would benefit greatly from some expansion in the discussion.

Particular issues

Additional minor comments have been attached that could help improve the manuscript:

Abstract

Line 25. Replace "… and the brook lamprey" with "… and two of the brook lamprey"

Introduction

Line 43. How about the large, rapidly diversifying lakes such as those in Africa and Nicaragua?

Lines 47 – 49. The authors mention that there are numerous examples of challenges in determining species in purely marine habitats, maybe they could provide some specific examples and their respective references.

Lines 86 – 87. Is it really the main objective?

Methods

Line 100. Not the same tissues, why?

Line 105. Blood was not mentioned before, right?

Line 128. “M SMRT cells”. Above mentioned 8M (line 125), what is it? Are they the same?

Line 137. “Arima Genome-Wide Hi-C+ Kit” Why not the same library protocol?

Line 148. Smudgeplot. Why isn't it included in Table 1?

Line 163. Bedtools version?

Line 173. More details?

Line 179. These assessment tools.

Line 183. Table 1. Maybe this table would be better as supplementary material.

Lines 238 – 239. These codes are not the same as in the rest of the text, tables 2, 3 or supp. table 1. It’s understood that that kcLamPlan1.1.hap1 is kcLamPlan1.2.hap1, and the same for the other three. Please clarify this.

Results

Line 249 – 250. Better PacBio instead of Pacific Biosciences

Line 264 and 266. “… of 1073 Mb (Figure 1) and 963 Mb (Table 2)” “...of 1049 Mb (Figure 1) and 960 Mb”. The genome assembly sizes differ by approximately 100 Mb between haplotypes in both species. The authors could mention something about the possible causes of these differences.

Line 269 – 270. Replace “… and 12.9 Mb in pseudo-haplotype one and two, respectively” with: “in both one and two pseudo-haplotypes.”

Line 273. Code kcLamFluc1 appears instead of kcLamFluv1. Please correct the c.

Table 2. Number of scaffolds. Missing number of chromosomes or pseudochromosomes.

Table 2. Organelles (MT). And maybe they could do something about this if they already have it sequenced, right?

Line 335. “… their support” They all look similar, right?

Line 346. “… sea lamprey” The scientific names should be used, as in the rest of the manuscript, to maintain consistency.

Line 347. Replace rrook with brook.

Line 354. What about comparing different haplotypes (hap1 and hap2) of the same individual for both *L.fluviatilis* and *L.planeri*?

Discussion

General: In the introduction, the ecological aspects of these species are extensively discussed; however, this topic is scarcely addressed in the discussion section. The discussion appears somewhat unbalanced.

Line 399. “… most of the gene trees (40.1%)”. Referring to 40% as 'most' misleading. Could the authors clarify this?

Lines 412 – 413. "… suggests that the two species rather is a species complex representing two ecotypes". Can this conclusion be suggested based on chromosome-level genome assemblies from two individuals?

Supplementary Material

Supplementary Figure 5. Lines 42 to 44. Please indicate the correspondence of the codes in the figure: kcLamPlan1 (RS), kcLamFluv2 (BS), kcLamFluv1 (RB), and kPerMar (SL).

- **Title and abstract**
    - Does the title clearly reflect the content of the article? [X] Yes, [ ] No (please explain), [ ] I don't know
    - Does the abstract present the main findings of the study? [X] Yes, [ ] No (please explain), [ ] I don't know
- **Introduction**
    - Are the research questions/hypotheses/predictions clearly presented? [X] Yes, [ ] No (please explain), [ ] I don't know
    - Does the introduction build on relevant research in the field? [X] Yes, [ ] No (please explain), [ ] I don't know
- **Materials and methods**
    - Are the methods and analyses sufficiently detailed to allow replication by other researchers? [X] Yes, [ ] No (please explain), [ ] I don't know
    - Are the methods and statistical analyses appropriate and well described? [ ] Yes, [ ] No (please explain), [X] I don't know.

      As mentioned before, it is unclear whether the methods align with the proposed hypothesis.
- **Results**
    - In the case of negative results, is there a statistical power analysis (or an adequate Bayesian analysis or equivalence testing)? [ ] Yes, [ ] No (please explain), [X] I don't know
    - Are the results described and interpreted correctly? [ ] Yes, [ ] No (please explain), [X] I don't know

      Could the authors "suggest" that the two species rather is a species complex representing two ecotypes" based on the methodology and the results?
- **Discussion**
    - Have the authors appropriately emphasized the strengths and limitations of their study/theory/methods/argument? [ ] Yes, [ ] No (please explain), [X] I don't know
    - Are the conclusions adequately supported by the results (without overstating the implications of the findings)? [ ] Yes, [ ] No (please explain), [X] I don't know

As mentioned above, the alignment between the main objective, the methods, and the discussion seems somewhat misaligned. The manuscript would benefit greatly from some expansion in the discussion.