

The logo for Peer Community In Genomics features a stylized circular network of nodes and lines, with a central cluster of nodes and lines radiating outwards, set against a background of blue and white dots.

# Peer Community In Genomics

## New chromosome-scale genome assembly for the Dwarf Goby, a River Rhine invader

**Iker Irisarri**  based on peer reviews by **Ruiqi Li** , **Tereza Manousaki** and 1 anonymous reviewer

Alexandra Schoenle, Nadège Guiglielmoni, Tobias Mainz, Carola Greve, Alexander B. Hamadou, Lisa Heermann, Jost Borchering, Ann-Marie Waldvogel (2024) Chromosome level genome reference of the Caucasian dwarf goby *Knipowitschia* cf. *caucasica*, a new alien Gobiidae invading the River Rhine. bioRxiv, ver. 3, peer-reviewed and recommended by Peer Community in Genomics. <https://doi.org/10.1101/2024.04.22.590508>

Submitted: 03 May 2024, Recommended: 22 October 2024

### Cite this recommendation as:

Irisarri, I. (2024) New chromosome-scale genome assembly for the Dwarf Goby, a River Rhine invader. *Peer Community in Genomics*, 100376. [10.24072/pci.genomics.100376](https://doi.org/10.24072/pci.genomics.100376)

Published: 22 October 2024

Copyright: This work is licensed under the Creative Commons Attribution 4.0 International License. To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

---

Since the opening of the Rhine-Main-Danube-Channel, four goby species are known to have invaded the River Rhine. Of these, the most recent and numerous is the Caucasian Dwarf Goby, which has been found in the Rhine since 2019. This study presents a new high-quality genome for this species (*Knipowitschia* cf. *caucasica*) (Schoenle et al. 2024). Currently, chromosome-scale genome assemblies represent a key first step in invasion biology, allowing the reconstruction of a species' invasion history and monitoring its progress, as well as identifying and characterizing candidate genes that control invasiveness (McCartney et al. 2019).

The authors sequenced the nuclear and mitochondrial genomes of this species using state-of-the-art methods including long-read sequencing techniques, scaffolded based on chromatin conformation data, and annotated using both direct transcriptomic and protein homology evidence. Data analyses follow currently established pipelines for genome assembly, scaffolding, annotation, and downstream bioinformatic analyses. The quality of the final genome was thoroughly assessed and conforms to what is expected from other genomes of fishes in the family Gobiidae. This study follows other recent endeavors that generated high-quality genomes to improve our understanding of invasion biology (e.g. Shao et al. 2020 and Kitsoulis et al. 2023). These studies are successfully contributing to increasing the genomic resources for the world's most damaging invasive species, which were not available for even a third of the top 100 invasive species just five years ago (McCarthy et al. 2019). Beyond invasion biology, the Dwarf Goby genome is also an important resource for many other applications, including evolutionary genomic analyses and phylogeography of this species and closely related ones in their native ranges.

### References:

Kitsoulis CV, Papadogiannis V, Kristoffersen JB, Kaitetzidou E, Sterioti A, Tsigenopoulos CS, Manousaki T (2023) Near-chromosome level genome assembly of devil firefish, *Pterois miles*. Peer Community Journal 3:e64. <https://doi.org/10.24072/pcjournal.295>

McCartney MA, Mallez S, Gohl DM (2019) Genome projects in invasion biology. Conservation Genetics 20:1201–1222. <https://doi.org/10.1007/s10592-019-01224-x>

Schoenle A, Guiglielmoni N, Mainz T, Greve C, Hamadou AB, Heermann L, Borchering J, Waldvogel A-M (2024) Chromosome level genome reference of the Caucasian dwarf goby *Knipowitschia* cf. *caucasica*, a new alien Gobiidae invading the River Rhine. bioRxiv, ver. 3 peer-reviewed and recommended by PCI Genomics. <https://doi.org/10.1101/2024.04.22.590508>

Shao F, Ludwig A, Mao Y, Liu N, Peng Z (2020). Chromosome-level genome assembly of the female western mosquitofish (*Gambusia affinis*). GigaScience 9:giaa092. <https://doi.org/10.1093/gigascience/giaa092>

## Reviews

### Evaluation round #2

DOI or URL of the preprint: <https://doi.org/10.1101/2024.04.22.590508>

Version of the preprint: 2

### Authors' reply, 17 October 2024

Thank you for your feedback and for taking the time to review our revised manuscript.

To clarify regarding the BUSCO completeness values: we assessed the completeness of the three assemblies immediately after assembly, prior to any scaffolding or polishing. This is why the Flye assembly presents two different BUSCO completeness values: one reflects the initial assembly, while the other represents the final genome after scaffolding, gap-closing, and polishing. Notably, the Flye assembly exhibited the highest completeness at the initial stage compared to the other assemblers (wtdbg2 and Raven), which is one of the reasons why we selected it for the final analysis. All BUSCO assessments were conducted using the actinopterygii\_odb10 dataset. To enhance clarity, we have added two sentences in both the Materials and Methods and Results sections.

We also appreciate your attention to detail regarding the caption error where we alternated between wtdbg2 and Raven. We have corrected this mistake and the typo in the revised manuscript.

### Decision by Iker Irisarri , posted 16 October 2024, validated 16 October 2024

#### Minor revision

Dear authors,

Many thanks for revising the manuscript.

I think all the Reviewers' concerns have been properly addressed and I will happily recommend your study.

There are just two minor things that I would like you to double-check for correctness.

The BUSCO completeness in Table 1 (90.52%) does not seem to agree with what is reported in the abstract (92.3%) or the snail plot (92.3%). It might seem as if some numbers came from an old analysis made with the vertebrata and not actinopterygii dataset.

Table 1. It seems as if BUSCO completeness was not tested for wtdbg2, and not for raven as stated in the caption. Please also check "based on actinopterygii\_odb10".

Thanks!

## Evaluation round #1

DOI or URL of the preprint: <https://doi.org/10.1101/2024.04.22.590508>

Version of the preprint: 1

### Authors' reply, 02 October 2024

[Download author's reply](#)

[Download tracked changes file](#)

### Decision by [Iker Irisarri](#) , posted 26 June 2024, validated 26 June 2024

#### Minor revision

Dear authors,

Thank you very much for submitting your study to PCI Genomics. As you have seen, the three Reviewers are very supportive of your study and I can only agree with them.

A recurrent suggestion was to add some comparative analyses with close relatives. I think, for example, a synteny analysis could strengthen the credibility of the current assembly.

It seems the genome annotation is not publicly available. Unless there are strong objections (e.g. use in upcoming studies), I would strongly recommend to also publish the genome annotation data.

Also, two minor comments.

In line 22 unpublished data is mentioned. It would be good to at least mention the source.

Please check line 24 makrozoobenthos; Line 154 Junniorprofessorship

Thank you!

Iker

### Reviewed by [Tereza Manousaki](#), 05 June 2024

The paper by Schoenle et al. present the first genome assembly and annotation of the Caucasian dwarf goby *Knipowitschia cf. caucasica*, an invasive species of the river Rhine. This resource is critical for studying the species biology and the invasion processes.

The genome reference is of high quality and at chromosome level. Thus, I would encourage the authors to change the title to something similar to "Chromosome level genome reference of the Caucasian dwarf goby *Knipowitschia cf. caucasica*, a new alien Gobiidae invading the River Rhine" to make sure that the title represents better the quality of the resource produced.

From all sections, M&Ms, results and discussion everything is mostly clear. My only concern is the way BUSCO score is presented at the genome level which confuses me as I would expect to see the total percentage of single and duplicated genes instead of the percentage of the single and duplicated from the identified BUSCOs.

Finally, the gene prediction seems to have returned moderate results compared to the genome. For a direct comparison it would be easier to have both genome and annotation BUSCOed with the same database (e.g. Actinopterygii). For improving the gene prediction, maybe the authors could consider to add protein hints from closely related species.

With best regards,

Tereza Manousaki

Reviewed by **Ruiqi Li** , 12 May 2024

The genome note is well-written and provides a very clear description of rationale for sequencing its genome, the methods used, etc. The genome itself is a valuable resource for further studies. I have only a few minor comments to add.

-----  
Specific Comments:  
-----

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

**Lines 28-32: A transition/connection is needed to clarify why a reference genome is important for subsequent studies, and subsequent analyses on local adaptations and population genetics.**

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? [X] Yes, [ ] No (please explain), [ ] I don't know

Results

In the case of negative results, is there a statistical power analysis (or an adequate Bayesian analysis or equivalence testing)? [X] Yes, [ ] No (please explain), [ ] I don't know

Are the results described and interpreted correctly? [X] Yes, [ ] No (please explain), [ ] I don't know

**1. Line 104: It would be great if the authors add a table comparing the three assemblies to detail their differences and the rationale for selecting the best assembly.**

**2. Line 106-107, 115-116: I think you should use the percentage of total BUSCO for single and duplicated BUSCOs in parentheses, or at least ensure consistency in how these figures are presented.**

Discussion

Have the authors appropriately emphasized the strengths and limitations of their study/theory/methods/argument? [X] Yes, [ ] No (please explain), [ ] I don't know

Are the conclusions adequately supported by the results (without overstating the implications of the findings)? [X] Yes, [ ] No (please explain), [ ] I don't know

Reviewed by anonymous reviewer 1, 24 June 2024

-Title and abstract

.\_Does the title clearly reflect the content of the article? Yes,

.\_Does the abstract present the main findings of the study? Yes

-Introduction

.\_Are the research questions/hypotheses/predictions clearly presented? Yes

.\_Does the introduction build on relevant research in the field? Yes

-Materials and methods

.\_Are the methods and analyses sufficiently detailed to allow replication by other researchers? Yes

.\_Are the methods and statistical analyses appropriate and well described? Yes

-Results

.\_In the case of negative results, is there a statistical power analysis (or an adequate Bayesian analysis or equivalence testing)? Not applicable

.\_Are the results described and interpreted correctly? Yes

-Discussion

.\_Have the authors appropriately emphasized the strengths and limitations of their study/theory/methods/argument? Yes

.\_Are the conclusions adequately supported by the results (without overstating the implications of the findings)?  
Yes

This study reports a new chromosome level genome assembly and transcriptome analysis of the dwarf goby *Knipowitschia cf. caucasica*. This species is a recent invader in the River Rhine coming from the Caucasian region, predicted to have the potential to seriously impact the local species community due to its ecological habits. The generation of a high-quality genome will be of great benefit for thorough analyses of adaptation, phylogeographic reconstructions, and coevolutionary processes with local prey or competitor species.

Overall, I find the study well done and well presented. I only miss a comparative genome analysis with an existing genome of a closely related species. In the discussion the authors compare some parameters with the round goby *Neogobius melanostomus*.