GBE

# The Chromosome-Scale Genome Assembly of the Redlip Blenny, *Ophioblennius macclurei* (Blenniidae)

Nikita Kulikov [1,2,†], Kim Joffroy [2,†], Anthony M. Bonacolta [3,4], Javier del Campo [3,4], Iker Irisarri [1,2,*]

[1]Department of Biodiversity and Evolutionary Biology, Museo Nacional de Ciencias Naturales (MNCN-CSIC), Madrid 28006, Spain
[2]Section Phylogenomics, Centre for Molecular Biodiversity Research, Leibniz Institute for the Analysis of Biodiversity Change (LIB), Zoological Museum Hamburg, Hamburg 20146, Germany
[3]Department of Marine Biology and Ecology, Rosenstiel School of Marine, Atmospheric and Earth Science, University of Miami, Miami, FL 33149, USA
[4]Biodiversity Program, Institut de Biologia Evolutiva (CSIC-Universitat Pompeu Fabra), Barcelona, Catalonia 08003, Spain

*Corresponding author: E-mail: iker.irisarri@mncn.csic.es.

[†]These authors contributed equally to this work.

**Accepted:** October 31, 2025

## Abstract

We present a high-quality, chromosome-scale genome assembly for the redlip blenny *Ophioblennius macclurei* (family Blenniidae). The assembly was generated using a combination of Oxford Nanopore long-read sequencing, Illumina short-read data, and Hi-C scaffolding technology. The assembled genome is 529.6 Mb in size, with a scaffold N50 of 23.7 Mb and a GC content of 43.49%. BUSCO analysis recovered 97.06% of expected genes in the scaffolded assembly and 89.2% from the annotated proteome. Automatic genome annotation identified a total of 18,927 protein-coding genes. This genome provides a valuable resource for understanding the diversity and evolutionary history of *Ophioblennius* and the fish family Blenniidae.

**Key words:** Blenniidae, combtooth blenny, Hi-C, HiFi, reference genome, platinum genome.

## Significance

Combtooth blennies are cryptobenthic fish inhabiting coastal and reef environments with high ecological importance in nutrient cycling. They are also evolutionarily intriguing due to their high diversity and fast diversification rates. Yet, the scarcity of genomic resources for this fish family is currently limiting our understanding of blennies' ecology and evolution. Specifically, the genome assembly of the redlip blenny could aid in a better understanding of the cryptic diversity and species boundaries in the genus *Ophioblennius*.

## Introduction

The redlip blenny *Ophioblennius macclurei* (Silvester, 1915) is a tropical marine fish of the family Blenniidae native to the western tropical Atlantic Ocean (Caribbean). It is found in shallow waters and coral reefs, where it is abundant, ecologically relevant, and space-controlling, and feeds mostly on algae and detritus (Nursall 1977). This species possesses two very long and sharp canine teeth, in addition to the

typical comb-like teeth of blennies. Originally, *O. macclurei* was considered a subspecies of *Ophioblennius atlanticus* (Valenciennes, 1836) with a distinct geographic distribution (Springer 1962) but was later recognized as a separate species (Rangel and Mendes 2009). To date, five species of *Ophioblennius* have been formally described, and genetic data show the presence of at least two additional lineages that might represent cryptic species (Lastrucci et al. 2018). The new genome assembly for *O. macclurei* could support genomic studies in the genus, aiding in the identification of cryptic species, clarifying species boundaries, and studying the patterns of isolation and gene flow across the Atlantic.

Combtooth blennies (Blenniidae) are a highly diverse fish family, with at least 413 valid species and 59 genera (Fricke et al. 2025). They are ecologically and evolutionarily highly relevant (Ord and Cooke 2016; Brandl et al. 2018), but their true potential remains underappreciated in part due to the scarcity of genomic resources and studies. So far, complete genomes for only seven combtooth blenny species have been sequenced, most by large genome sequencing consortia: *Blennius ocellaris*, *Coryphoblennius galerita*, *Lipophrys pholis*, *Parablennius gattorugine*, *Parablennius parvicornis*, *Salaria pavo*, and *Salarias fasciatus* (Cardoso et al. 2024; Adkins et al. 2025). Of these, annotated proteins are currently available only for *S. fasciatus* and *S. pavo*.

In this study, we used long (Oxford Nanopore) and short (Illumina) genomic reads together with chromatin conformation (Hi-C) and gene expression data (Illumina RNAseq) to generate and annotate the chromosome-scale genome assembly of *O. macclurei*. Genomic DNA reads were originally obtained for a study aiming to identify possible apicomplexan parasites in blood from four individual fish (Bonacolta et al. 2024). We here leveraged these data to assemble a new genome of combtooth blenny. To maximize sequencing depth of long reads, we merged the information from the four individuals, which derive from the same population in Curaçao. Hi-C and RNAseq data were derived from two further individuals of the same population.

## Results and Discussion

The nuclear genome assembly has a total length of 529.6 Mb, with contig N50 of 1.79 Mb and scaffold N50 of 23.7 Mb (Table 1, Fig. 1a). Hi-C data scaffolded contigs into 23 pseudochromosomes, which account for 99.1% of the assembly and agree with the expected karyotype of the closely related species *Ophioblennius trinitatis* (Galvão et al. 2011). Scaffold (pseudochromosome) sizes range from 14.28 to 37.24 Mb (Fig. 1a). The Hi-C contact map shows a near-complete assembly with all 23 pseudochromosomes being correctly oriented. The Merqury assembly spectrum plot shows that haplotypes were successfully collapsed after GenomeScope suggested a relatively high heterozygosity of 0.724% (Fig. 1b and c). Merqury indicated a QV (quality value) of 31.10, which corresponds to 99.9% per-basepair accuracy. The spectra copy number plot (Fig. 1b) displays a single prominent peak, indicating that the majority of k-mers from the sequencing reads were assembled once and not duplicated, suggesting a high-quality and nonredundant assembly. BUSCO results indicate high genome completeness (97.55%) for the "actinopterygii_odb10' dataset, consisting of 3,478 single-copy, 56 duplicated, and 17 fragmented BUSCOs (Fig. 1a). The presence of contaminants was assessed using BlobToolKit, which classifies contigs based on sequencing depth, GC %, and taxonomic assignment (Fig. 1d).

Assembly contiguity and BUSCO completeness of the *O. macclurei* genome are comparable to those of *S. fasciatus* (Table 1). The total genome size and chromosome numbers are comparable to those of other sequenced combtooth blenny genomes, including *S. fasciatus*. To further assess the Hi-C scaffolding step, we performed a synteny analysis with the *S. fasciatus* genome, which identified largely conserved synteny across of all 23 pseudochromosomes and the presence of several inversions and translocations, in addition to a putative fusion of two *S. fasciatus* pseudochromosomes into pseudochromosome 1 of *O. macclurei* (Fig. 1e). The presence of inversions and translocations might be expected given the relatively high phylogenetic distance among the two species (Hundt et al. 2014) and genome size differences (529.6 vs. 797.5 Mb), but detailed analyses with additional chromosome-scale assemblies would be required to fully understand synteny conservation across the Blenniidae.

RepeatMasker identified 14.98% of the genome as repetitive, consisting of SINE (2.85%), LTR (1.17%), and LINE (1.64%) retroelements as well as DNA transposons (3.0%) (Table 1). In agreement, GenomeScope identified 13.55% of the genome as repetitive. The proportion of repetitive elements is smaller compared to the 32.55% found in the 735.9 Mb genome of *S. pavo*, which is also 39% larger in size (Cardoso et al. 2024).

Structural annotation of the assembly was based on direct RNAseq data and existing proteins from actinopterygian fish (see Material and Methods). In total, 18.927 proteins were annotated (Table 1). Among them, we retrieved 3,278 (90.0%) complete BUSCOs ("actinopterygii_odb10"), of which 3,085 were single-copy, 164 duplicated, and 29 fragmented. The genome of *S. fasciatus* has twice as many annotated proteins and higher BUSCO completeness, despite retrieving fewer single-copy (2,467) and more duplicated (1,125) BUSCOs than in *O. macclurei*. This difference is probably explained by a higher redundancy in the genome sequence of *S. fasciatus* and the use of distinct annotation pipelines.

**Table 1** Summary table of raw sequencing data and nuclear genome assembly and annotation statistics for *O. macclurei* and *S. fasciatus*.

| | | *O. macclurei* | *S. fasciatus* |
|---|---|---|---|
| Genomic raw data | Total number of long reads | 1,057,526 (ONT) | 8,010,177 (PacBio)* |
| | Total bp (long reads) | 6,310,502,152 (ONT) | 65,323,603,844 (PacBio)* |
| | Total number of reads (Illumina) | 1,779,183,086 | 839,095,407* |
| | Total bp (Illumina) | 268,656,645,986 | 348,812,695,149* |
| Hi-C raw data | Total number of reads (Illumina) | 49,822,383 | NA* |
| | Total bp (Illumina) | 7,523,179,833 | NA* |
| RNAseq raw data | Total number of reads (Illumina) | 127,138,393 | NA |
| | Total bp (Illumina) | 19,197,897,343 | NA |
| Genome assembly statistics | Assembly size (bp) | 529,575,105 | 797,507,141 |
| | Number of scaffolds | 180 | 203 |
| | Scaffold N50 (bp) | 23,714,998 | 32,729,575 |
| | Longest scaffold (bp) | 37,241,069 | 41,705,463 |
| | Shortest scaffold (bp) | 521 | 3,075 |
| | Number of contigs | 689 | 805 |
| | Contig N50 (bp) | 1,791,215 | 2,597,836 |
| | Longest contig (bp) | 11,331,113 | 15,801,682 |
| | Shortest contig (bp) | 521 | 806 |
| | GC (%) | 43.49 | 44.34 |
| BUSCO completeness | Single-copy | 3,478 (95.5%) | 3,367 (92.5%) |
| (actinopterygii_odb10, 3,640 genes) | Duplicated | 58 (1.6%) | 233 (6.4%) |
| | Fragmented | 16 (0.4%) | 10 (0.3%) |
| | Missing | 88 (2.5%) | 10 (0.8%) |
| Genome annotation | Protein-coding genes | 18,927 | 39,295 |
| | BUSCO single-copy | 3,085 (84.8%) | 2,467 (67.8%) |
| | BUSCO duplicated | 164 (4.5%) | 1,125 (30.9%) |
| | BUSCO fragmented | 29 (0.8%) | 7 (0.2%) |
| | BUSCO missing | 362 (9.9%) | 41 (1.1%) |
| | Repetitive DNA | 14.98% | NA |
| | SINE | 2.85% | NA |
| | LTR | 1.17% | NA |
| | LINE | 1.64% | NA |
| | DNA transposons | 3.00% | NA |

Assemblies are based, respectively, on a combination of Oxford Nanopore Technology (ONT) or SEQUEL Pacific Biosciences (PacBio) long reads and Illumina short reads. In the absence of a publication for the *S. fasciatus* genome, we report raw sequencing data of SRA accessions associated with NCBI BioSample SAMEA4966329, marked by asterisks

The mitochondrial genome was assembled using GetOrganelle (Jin et al. 2020) into a circularized molecule of 16,491 bp and annotated with MITOS2 (Bernt et al. 2013). It showed the gene content and order typical of most other vertebrates, including *S. fasciatus*. The cytochrome b gene sequence was used to confirm the species identification as *O. macclurei* (Fig. S1a), which further agrees with its inferred geographic range in the Caribbean (Lastrucci et al. 2018).

## Conclusion

In this study, we show how genomic data generated for other purposes can be leveraged to generate a nuclear chromosome-scale and mitochondrial genomes. This new genome adds to a still scarce set of genome assemblies for combtooth blennies, which represent key resources for understanding the ecology, behavior, and evolution of this group of cryptobenthic fish.

## Materials and Methods

### DNA and RNA Extraction and Sequencing

Blood samples were obtained from four *O. macclurei* individuals (12E, 3E, 9-3E, 1S) obtained from the pet trade, all originally from the same population in Curaçao. DNA was extracted using the QIAGEN DNeasy Blood and Tissue Kit (CAT# 69504). All four samples were sequenced with Oxford Nanopore Technology using the Genomic DNA by Ligation protocol (Version, GDE_9063_v109_revAB_14Aug2019; Kit, SQK-LSK109). Sequencing was conducted using a MINion FLO-MIN106 Flowcell. All the samples were basecalled using guppy and the fastq_hac model within it. In addition, we generated $2 \times 150$ bp Illumina reads (100 M) for one individual (E9).

For Hi-C, we obtained a fifth individual, and a muscle tissue sample was flash frozen using liquid nitrogen and then stored at −80°C. Hi-C data were generated at Phase Genomics following their Proximo Hi-C workflow.

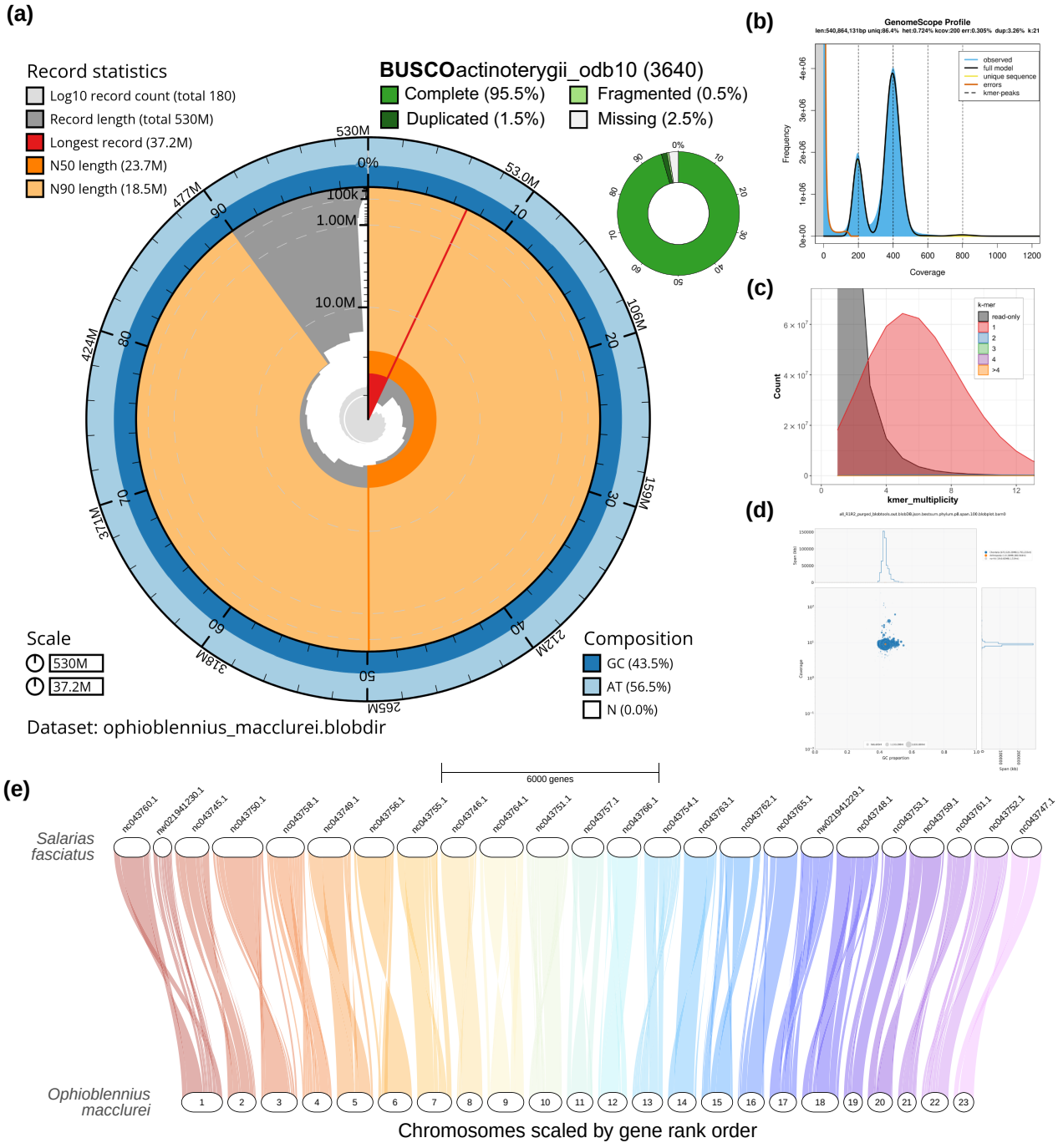


**Fig. 1.** a) The genome assembly of *Ophioblennius macclurei* has high contiguity and BUSCO completeness as summarized in the snail plot generated with BlobToolKit. b) The GenomeScope k-mer profile plot based on the 9-3E Illumina reads displays the fit of the model (black) to the observed k-mer frequencies (blue), showing two peaks typical of diploid organisms, a k-mer-based estimated genome size of 540 Mb and a heterozygosity rate of 0.724%. c) The Merqury copy number spectrum plot shows low redundancy of the assembly, indicating that the heterozygous regions were properly collapsed. d) BlobPlot indicating that the final genome assembly is free of contaminants. e) Macro-synteny between *O. macclurei* and *S. fasciatus* genomes.

For RNAseq, we obtained seven tissues (brain, eyes, fins, gills, heart, liver, skin) from another individual, RNA extracted following a TriZoL protocol. Individual libraries were generated using the KAPA RNA HyperPrep Kit with RiboErase (HMR) (KR1351; Roche) and sequenced in the Illumina NovaSeq X Plus platform 2 × 150 bp reads (20 M reads each).

## Nuclear Genome Assembly

Nanostats v.1.6.0 (De Coster et al. 2018) and FastQC v.0.11.9 (Andrews 2011) were used for assessing the quality of raw sequencing reads. For long reads, adapter sequences were removed using Porechop v.0.2.4 (Wick et al. 2017). Due to the low sequencing depth of individual samples, the Oxford Nanopore reads from all four individuals were used for de novo genome assembly using Flye v.2.9.2 (Kolmogorov et al. 2019) with default parameters. To improve the per-base accuracy of the long-read–based assembly, we performed two rounds of polishing using Racon v.1.5.0 (Vaser et al. 2017) and the short Illumina reads of individual 9-3E. Illumina reads were also used to assess genome size and heterozygosity rate using GenomeScope (Ranallo-Benavidez et al. 2020) with k = 21 and a maximum coverage threshold (max_kmercov = 10,000,000). BlobToolKit (Challis et al. 2020) was used to check for possible contaminants, assigning taxonomy to contigs by BLASTN searchers against the NCBI nt database.

We performed a specific decontamination step to identify and remove any possible apicomplexan parasite sequences, which were known to be present in the source samples. To do this, we performed a sequence similarity search with MMseqs2 v.11-e1a1c (Steinegger and Söding 2017) to screen each contig against a custom decontamination database that included genomes from seven apicomplexan species (*Eimeria tenella*, *Eimeria brunetti*, *Eimeria necatrix*, *Eimeria maxima*, *Eimeria. mitis*, *Eimeria acervulina*, and *Babesia bovis*) plus the sequences of the *S. fasciatus* genome, serving as positive reference set. Thirty four contigs showed low (<70%) sequence similarity to *S. fasciatus* sequences and these were further scrutinized by manual BLASTN searchers against the NCBI nt database. Contig 417 was identified as *E. coli* and removed manually. All other 33 contigs represented nonconserved genome fragments from fish, with no obvious similarity to apicomplexans. Thus, we considered the genome to be free of apicomplexan contaminants, in agreement with the final BlobPlot analysis (Fig. 1d). Mitochondrial contigs were identified by BLASTN against the *S. fasciatus* mitogenome and removed manually from the assembly.

Haplotigs and overlaps were removed from the decontaminated genome using purge_dups v.1.2.6 (Guan et al. 2020). A final assembly quality check was performed using Meryl v.1.4.1 (Rhie et al. 2020) with k = 21. Mercury v.1.3 (Rhie et al. 2020) was used to perform a final evaluation of the assembly using k = 21.

We used chromatin conformation information (Hi-C) to scaffold the assembly contigs, following the VGP assembly pipeline in the Galaxy environment (Rhie et al. 2021; Larivière et al. 2024). Briefly, the Hi-C reads were mapped against the assembled genome with BWA-MEM2 v.2.2.1 (Vasimuddin et al. 2019) and a first Hi-C contact map was generated using PretextMap v.0.1.9 (https://github.com/sanger-tol/PretextMap) and Pretext Snapshot v.0.0.5 (https://github.com/sanger-tol/PretextSnapshot). The genome was scaffolded with YaHS v.1.2 (Zhou et al. 2023).

The final genome assembly completeness was assessed with BUSCO v.4.0.6 (Manni et al. 2021) with default parameters and the "actinopterygii_odb10' database. Genome assembly metrics were visualized in a snail plot generated with BlobToolKit v.4.4.5 (Challis et al. 2020).

## Nuclear Genome Annotation

Structural annotation followed a Snakemake pipeline originally used for fish genome annotation (Oriowo et al. 2025; https://github.com/YanisChrys/phoxinus_annotation). Briefly, repetitive elements were identified de novo with RepeatModeler v.2.0.1. Repetitive DNA and masking was performed with RepeatMasker version v.4.1.0 (Smit et al. 2013) using the repeat library previously identified via RepeatModeler and using rmblastn v.2.10.0 (Flynn et al. 2020) as a search engine. The masked genome assembly was used for structural annotation of proteins using direct RNAseq evidence and homology to known fish proteins using the BRAKER v.3.0.7 (Gabriel et al. 2024). RNAseq alignment maps were generated by mapping the newly generated RNAseq data onto the masked genome using STAR v.2.7.10 (Dobin et al. 2013). The protein database consisted of annotated proteins from 11 fish species: *Acanthochromis polyacanthus* (GCF_021347895.1), *Cyprinodon variegatus* (GCF_000732505.1), *Danio rerio* (GCF_000002035.6), *Gouania willdenowi* (GCF_900634775.1), *Lepisosteus oculatus* (GCF_000242695.1), *Mugil cephalus* (GCF_022458985.1), *Oreochromis niloticus* (GCF_001858045.2), *Oryzias latipes* (GCF_000313675.1), *S. fasciatus* (GCF_902148845.1), *Stegastes partitus* (GCF_000690725.1), and *Xiphophorus maculatus* (GCF_002775205.1), all downloaded from NCBI.

## Chromosome-Scale Collinearity

To further assess the quality of the genome scaffolding, we performed a synteny analysis of the new *O. macclurei* genome with the existing chromosome-level *S. fasciatus* genome assembly (GCF_902148845.1), using GENESPACE (Lovell et al. 2022).

## Mitochondrial Genome

The mitochondrial genome was assembled using GetOrganelle (Jin et al. 2020) from the raw Illumina reads (9-3E) and then annotated with MITOS v.2.1.9 (Bernt et al. 2013). The species identification was confirmed by a maximum likelihood analysis of cytochrome b sequences obtained from Lastrucci et al. (2018). A multiple sequence alignment was inferred with MAFFT v7.427 (Katoh and

Standley 2013), which was later analyzed with IQ-TREE3 v.3.0.1 (Wong et al. 2025) under the BIC-selected best-fit evolutionary model and 1000 bootstrap pseudoreplicates. Tree rooting follows (Lastrucci et al. 2018).

## Supplementary Material

Supplementary material is available at *Genome Biology and Evolution* online.

## Acknowledgments

We thank Paul Sikkel for sampling and export permits. We thank Ioannis Chrysostomakis for advice in genome annotation. Computations were performed at the HPC of the Leibniz Institute for the analysis of Biodiversity Change (LIB), for which we acknowledge support by Alexander Donath and Sebastian Martin.

## Funding

This work was supported by the Leibniz Institute for the Analysis of Biodiversity Change (LIB), the Spanish Agencia Estatal de Investigación MICIU/AEI/10.13039/501100011033, European Social Fund Plus (ESF+), and European Regional Development Fund, European Union (ERDF/EU) (Grants RyC2022-038245-I and PID2023-152168NB-I00 to I.I.) and the Spanish Consejo Superior de Investigaciones Científicas (PIE Grant 20243AT022 to I.I.). AMB and JdC were supported by startup funds from the University of Miami. N.K. was supported by a scholarship by the Studienstiftung des Deutsches Volkes.

## Data Availability

The assembled genome is available at ENA under Study ID PRJEB95139. Raw sequencing data are available under NCBI BioProject PRJNA1041302 and ENA under Study ID PRJEB95139. The annotated nuclear proteins, multiple sequence alignment, and tree are available in Zenodo under https://doi.org/10.5281/zenodo.17371133.

## Literature Cited

Adkins P, et al. The genome sequence of the butterfly blenny, *Blennius ocellaris* Linnaeus, 1758. Wellcome Open Res. 2025:10:87. https://doi.org/10.12688/wellcomeopenres.23748.1.

Andrews S. FastQC. High Throughput Sequence QC Report. 2011. www.bioinformatics.babraham.ac.uk/projects/.

Bernt M, et al. MITOS: improved de novo metazoan mitochondrial genome annotation. Mol Phylogenet Evol. 2013:69:313–319. doi: http://dx.doi.org/10.1016/j.ympev.2012.08.023.

Bonacolta AM, Krause-Massaguer J, Smit NJ, Sikkel PC, del Campo J. A new and widespread group of fish apicomplexan parasites. Curr Biol. 2024:34:2748–2755.e3. https://doi.org/10.1016/j.cub.2024.04.084.

Brandl SJ, Goatley CHR, Bellwood DR, Tornabene L. The hidden half: ecology and evolution of cryptobenthic fishes on coral reefs. Biol Rev Camb Philos Soc. 2018:93:1846–1873. https://doi.org/10.1111/brv.12423.

Cardoso SD, Sun L, Gonçalves D. Chromosome-level genome assembly of the highly-polymorphic peacock blenny (*Salaria pavo*). Sci Data. 2024:11:1424. https://doi.org/10.1038/s41597-024-04242-8.

Challis R, Richards E, Rajan J, Cochrane G, Blaxter M. BlobToolKit—interactive quality assessment of genome assemblies. G3 (Bethesda). 2020:10:1361–1374. https://doi.org/10.1534/g3.119.400908.

De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. NanoPack: visualizing and processing long-read sequencing data. Bioinformatics. 2018:34:2666–2669. https://doi.org/10.1093/bioinformatics/bty149.

Dobin A, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013:29:15–21. https://doi.org/10.1093/bioinformatics/bts635.

Flynn JM, et al. RepeatModeler2 for automated genomic discovery of transposable element families. Proc Natl Acad Sci U S A. 2020:117:9451–9457. https://doi.org/10.1073/pnas.1921046117.

Fricke R, Eschmeyer WN, van der Laan R. Eschmeyer's catalog of fishes: Genera, species, references. 2025. http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp. Electronic version accessed dd 16/10/2025.

Gabriel L et al. BRAKER3: Fully automated genome annotation using RNA-seq and protein evidence with GeneMark-ETP, AUGUSTUS and TSEBRA. bioRxiv 2023.06.10.544449. 2024. https://doi.org/10.1101/2023.06.10.544449

Galvão T, Bertollo L, Molina W. Chromosomal complements of some Atlantic Blennioidei and Gobioidei species (Perciformes). CCG. 2011:5:259–275. https://doi.org/10.3897/compcytogen.v5i4.1834.

Guan D, et al. Identifying and removing haplotypic duplication in primary genome assemblies. Bioinformatics. 2020:36:2896–2898. https://doi.org/10.1093/bioinformatics/btaa025.

Hundt PJ, Iglésias SP, Hoey AS, Simons AM. A multilocus molecular phylogeny of combtooth blennies (Percomorpha: Blennioidei: Blenniidae): multiple invasions of intertidal habitats. Mol Phylogenet Evol. 2014:70:47–56. https://doi.org/10.1016/j.ympev.2013.09.001.

Jin J-J, et al. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. Genome Biol. 2020:21:241. https://doi.org/10.1186/s13059-020-02154-5.

Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013:30:772–780. https://doi.org/10.1093/molbev/mst010.

Kolmogorov M, Yuan J, Lin Y, Pevzner PA. Assembly of long, error-prone reads using repeat graphs. Nat Biotechnol. 2019:37:540–546. https://doi.org/10.1038/s41587-019-0072-8.

Larivière D, et al. Scalable, accessible and reproducible reference genome assembly and evaluation in Galaxy. Nat Biotechnol. 2024:42:367–370. https://doi.org/10.1038/s41587-023-02100-3.

Lastrucci NS, Nunes LT, Lindner A, Floeter SR. An updated phylogeny of the redlip blenny genus *Ophioblennius*. J Fish Biol. 2018:93:411–414. https://doi.org/10.1111/jfb.13732.

Lovell JT, et al. GENESPACE tracks regions of interest and gene copy number variation across multiple genomes. eLife. 2022:11:e78526. https://doi.org/10.7554/eLife.78526.

Manni M, Berkeley MR, Seppey M, Simão FA, Zdobnov EM. BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. Mol Biol Evol. 2021:38:4647–4654. https://doi.org/10.1093/molbev/msab199.

Nursall JR. Territoriality in redlip blennies (*Ophioblennius atlanticus*-Pisces: Blenniidae). J Zool. 1977:182:205–223. https://doi.org/10.1111/j.1469-7998.1977.tb04156.x.

Ord TJ, Cooke GM. Repeated evolution of amphibious behavior in fish and its implications for the colonization of novel environments. Evolution. 2016:70:1747–1759. https://doi.org/10.1111/evo.12971.

Oriowo TO, et al. A chromosome-level, haplotype-resolved genome assembly and annotation for the Eurasian minnow (Leuciscidae: *Phoxinus phoxinus*) provide evidence of haplotype diversity. GigaScience. 2025:14:giae116. https://doi.org/10.1093/gigascience/giae116.

Ranallo-Benavidez TR, Jaron KS, Schatz MC. GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. Nat Commun. 2020:11:1432. https://doi.org/10.1038/s41467-020-14998-3.

Rangel CA, Mendes LF. Review of blenniid fishes from Fernando de Noronha Archipelago, Brazil, with description of a new species of *Scartella* (Teleostei: Blenniidae). Zootaxa. 2009:2006:51–61. https://doi.org/10.11646/zootaxa.2006.1.4.

Rhie A, et al. Towards complete and error-free genome assemblies of all vertebrate species. Nature. 2021:592:737–746. https://doi.org/10.1038/s41586-021-03451-0.

Rhie A, Walenz BP, Koren S, Phillippy AM. Merqury: reference-free quality, completeness, and phasing assessment for genome assemblies. Genome Biol. 2020:21:245. https://doi.org/10.1186/s13059-020-02134-9.

Smit AFA, Hubley R, Green P. RepeatMasker Open-4.0. 2013. Available from: http://www.repeatmasker.org

Springer VG. A review of the blenniid fishes of the genus *Ophioblennius* Gill. Copeia. 1962:1962:426–433. https://doi.org/10.2307/1440912.

Steinegger M, Söding J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. Nat Biotechnol. 2017:35:1026–1028. https://doi.org/10.1038/nbt.3988.

Vaser R, Sovic I, Nagarajan N, Sikic M. Fast and accurate de novo genome assembly from long uncorrected reads. Genome Res. 2017:27:737–746. https://doi.org/10.1101/gr.214270.116.

Vasimuddin M, Misra S, Li H, Aluru S. Efficient architecture-aware acceleration of BWA-MEM for multicore systems. 2019 IEEE international parallel and distributed processing symposium (IPDPS). 2019, pp. 314–324. Rio de Janeiro, Brazil. https://doi.org/10.1109/IPDPS.2019.00041.

Wick RR, Judd LM, Gorrie CL, Holt KE. Completing bacterial genome assemblies with multiplex MinION sequencing. Microb Genom. 2017:3:e000132. https://doi.org/10.1099/mgen.0.000132.

Wong TKF et al. IQ-TREE 3: phylogenomic inference software using complex evolutionary models. 2025. https://ecoevorxiv.org/repository/view/8916/. Accessed August 15, 2025.

Zhou C, McCarthy SA, Durbin R. YaHS: yet another Hi-C scaffolding tool. Bioinformatics. 2023:39:btac808. https://doi.org/10.1093/bioinformatics/btac808.

**Associate editor**: Diego Cortez