Whole Genome Assembly and Annotation of Blackstripe Livebearer *Poeciliopsis prolifica*

Ying Zhang[1], Yuridia Reynoso[2], David Reznick[2], and Xu Wang [1,3,4,*]

[1]Department of Pathobiology, College of Veterinary Medicine, Auburn University, Alabama , USA
[2]Department of Evolution, Ecology and Organismal Biology, University of California, Riverside, California, USA
[3]Center for Advanced Science, Innovation and Commerce, Alabama Agricultural Experiment Station, Auburn, Alabama, USA
[4]HudsonAlpha Institute for Biotechnology, Huntsville, Alabama, USA

*Corresponding author: E-mail: xzw0070@auburn.edu.

Accepted: October 22, 2023

Abstract

The blackstripe livebearer *Poeciliopsis prolifica* is a live-bearing fish belonging to the family Poeciliidae with high level of post-fertilization maternal investment (matrotrophy). This viviparous matrotrophic species has evolved a structure similarly to the mammalian placenta. Placentas have independently evolved multiple times in Poeciliidae from nonplacental ancestors, which provide an opportunity to study the placental evolution. However, there is a lack of high-quality reference genomes for the placental species in Poeciliidae. In this study, we present a 674 Mb assembly of *P. prolifica* in 504 contigs with excellent continuity (contig N50 7.7 Mb) and completeness (97.2% Benchmarking Universal Single-Copy Orthologs [BUSCO] completeness score, including 92.6% single-copy and 4.6% duplicated BUSCO score). A total of 27,227 protein-coding genes were annotated from the merged datasets based on bioinformatic prediction, RNA sequencing and homology evidence. Phylogenomic analyses revealed that *P. prolifica* diverged from the guppy (*Poecilia reticulata*) ~19 Ma. Our research provides the necessary resources and the genomic toolkit for investigating the genetic underpinning of placentation.

Key words: long-read sequencing, genome assembly, genome annotation, Poeciliidae.

Significance

Placentas have independently evolved multiple times in the fish family Poeciliidae, which is ideal for investigating the evolution of placenta. However, there is no high-quality reference genome of placental species in Poeciliidae. Here, we generated a high-quality and well-annotated genome of *Poeciliopsis prolifica*, which is crucial for the evolutionary and comparative genomic studies in genus *Poeciliopsis*.

Introduction

Animals exhibit various reproductive strategies, from oviparity to viviparity, linked to the different levels of parental investment (Clutton-Brock 1991). Oviparity is known to be an ancestral reproductive mode in which females spawn eggs, and the nutrition for offspring is totally derived from the yolk provisioned before the egg is fertilized (lecithotrophy) (Wake 1992). In contrast, viviparity is a reproductive mode in which females give birth to live young. Viviparity may still be lecithotrophic, meaning that mothers may fully provision eggs before fertilization, but mothers may instead continue to provision the embryo after the egg is fertilized (matrotrophy) (Wourms et al. 1988; Pollux et al. 2009). Matrotrophic viviparity is not unique to therian

mammals. It is also found in some teleost fishes (Thibault and Schultz 1978; Wourms et al. 1988), reptiles (Guillette 1993; Braz et al. 2016), and amphibians (Wake 2015; Furness and Capellini 2019). All of these other taxa lack a uterus. Eggs are fertilized, and embryos develop in the ovarian cavity (Amoroso 1968; Kaye 1971; Wourms 1981; Rothchild 2003; Uribe et al. 2019). There are considerable variations among viviparous vertebrates in how much nutrition the pregnant female delivers to offspring during development. Some embryos of viviparous species receive nutrition entirely from the yolk (lecithotrophy), whereas others can acquire nutrition from pregnant females throughout development (matrotrophy). The mechanism of delivery can be through skin, gut, gill, oviduct, or placenta (Wake and Dickie 1998; Stewart and Thompson 2000; Exbrayat 2006; Uribe et al. 2021).

The placenta, an apposition of maternal and embryonic tissue, is a specialized form of matrotrophy (Mossman 1991), providing nutrient transport and gas exchange for embryo development and participating in physiological, endocrine, and immune interactions at the maternal–fetal interface (Faber et al. 1992; Moffett and Loke 2006; Dimasuay et al. 2016). In the live-bearing fish family Poeciliidae, the placenta has independently evolved multiple times from nonplacental ancestors. Placental species exhibit various degrees of postfertilization maternal provisioning (Reznick et al. 2002; Pollux et al. 2014). The genus *Poeciliopsis* alone includes more than 20 species that vary in maternal provisioning from lecithotrophy to extensive matrotrophy. There are three independent origins of placentation (Reznick et al. 2002), which makes the genus ideal for investigating placental evolution. However, there is no high-quality reference genome available for facilitating studies of the genomic basis of placentation in *Poeciliopsis*. Blackstripe Livebearer (*Poeciliopsis prolifica*) is a species with extensive matrotrophy (Wourms et al. 1988; Reznick et al. 2002). Our preliminary transcriptome sequencing in *P. prolifica* revealed that only 30–40% of the RNA-seq reads could be mapped to the guppy (*Poecilia reticulata*) genome, the closest species with a high-quality reference genome (Kunstner et al. 2016; Fraser et al. 2020) in Poeciliids.

van Kruistum et al. (2021), as part of a comparative study of live-bearing fish, assembled the *P. prolifica* genome using Illumina short-read sequencing, but the size and continuity were not ideal (593 Mb in 74,755 scaffolds with an N50 of 38.6 kb), with a Benchmarking Universal Single-Copy Orthologs (BUSCO) genome completeness score of 83.1%. Thus, a better reference genome of *P. prolifica* is needed for further transcriptome analyses. In this research, we reported a well-assembled and annotated genome of *P. prolifica* using the PacBio long-read sequencing and 10× Genomics linked-read sequencing. This high-quality genome provides the essential genetic toolkit for evolutionary and comparative genomic studies in *P. prolifica* and its closely related species and will facilitate the research of the placenta evolution.

Results and Discussion

Genome Assembly and Assessment

The total length of the estimated *P. prolifica* genome is 624,973,420 bp based on the K-mer distribution (K = 25; supplementary fig. S1, Supplementary Material online). The genome heterozygosity is 0.276%. A total of 99.1 Gb of PacBio Sequel CLS data and 39.4 Gb of Illumina linked-read data were generated for *P. prolifica* genome assembly (supplementary table S1, Supplementary Material online). Using PacBio long-reads alone, 721 contigs were assembled by CANU (see Materials and Methods), resulting in a 681,568,254 bp assembly with a contig N50 of 6,458,394 bp. BUSCO analysis revealed a genome completeness of 93.8%, with 88.9% single-copy and 4.9% duplicated BUSCOs. The 10× Genomics linked-reads scaffolds were generated and merged into the long-read assembly, resulting in 504 contigs in 415 scaffolds (Table 1). After merging, the final size of *P. prolifica* assembled genome is 674,152,735 bp, which a ~1% reduction in genome size. The contig N50 is 7.7 Mb, and the BUSCO completeness score is improved to 97.2%, with fewer duplicated BUSCOs (4.6%, see Table 1).

Compared with the reference guppy (GCA_904066995), the duplicated BUSCO is significantly higher in blackstripe livebearer (4.6% vs. 0.8%), which raises the concern whether the quickmerge procedure introduced duplicated regions. This is not the case because when BUSCO scores were checked for PacBio-only assembly, short-read assemblies from us and van Kruistum et al. (2021), we found that merging actually reduced the number of duplicated BUSCOs (from 4.9% to 4.6%), rather than introducing them (supplementary table S2, Supplementary Material online). In fact, well-assembled fish genomes may have variable duplicated BUSCOs, such as West African lungfish (*Protopterus annectens*; 94% completeness with 2.1% duplicated BUSCOs; GCF_019279795), and Japanese puffer (*Takifugu rubripes*; 98.3% completeness with 3.7% duplicated BUSCOs). Teleost experienced multiple rounds of whole-genome duplications, and taxa-specific gene loss and neofunctionalization may have occurred during evolutionary history. In this case, *P. prolifica* is from a distantly related genus compared with the reference guppy (*Poecilia*), so it is plausible for a higher duplicated BUSCO score (4.6% vs. 0.8%), given a higher total BUSCO score (97.2% vs. 93.9%; Table 1).

We aligned 16 RNA-seq data of *P. prolifica* samples from two independently and artificially inbred lines (Line A and Line B), and the average unique alignment rate was more

Table 1
Summary Statistics of *Poeciliopsis prolifica* and the Reference Guppy *Poecilia reticulata* Genome Assemblies

Genome assembly	*P. prolifica* (this assembly)	*Poecilia reticulata* (female) (GCA_000633615)	*Poecilia reticulata* (male) (GCA_904066995)
Sequencing data and coverage			
PacBio sequencing data	99.1 Gb PacBio Sequel CLS reads	N/A	50.3 Gb PacBio RS II CLS reads
Illumina sequencing data	39.4 Gb NovaSeq	89.9 Gb GAII	40.7 Gb HiSeq reads
Genome coverage	PacBio: 147×; Illumina: 58×	Illumina: 110×	PacBio: 72×; Illumina: 58×
Assembly statistics			
Total contig length	674,107,802 bp	664,638,573 bp	696,229,266 bp
Number of contigs	504	40,144	171
% of gap	0.007%	9.16%	0.212%
Contig N50	7.7 Mb	0.042 Mb	8.2 Mb
Contig L50	25	4,424	27
Maximum contig length	26.3 Mb	0.374 Mb	25.6 Mb
Completeness (actinopterygii_odb10)			
BUSCO completeness	97.2%	96.1%	93.9%
Single-copy BUSCO	92.6%	95.5%	93.1%
Duplicated BUSCO	4.6%	0.6%	0.8%
Fragmented BUSCO	0.5%	1.5%	1.4%
Missing BUSCO	2.3%	2.4%	4.7%
Annotation statistics			
Number of protein-coding genes	27,227	22,842	

N/A, not available.

than 90% (supplementary table S3, Supplementary Material online). The assembly and mapping statistics indicate both excellent contiguity and completeness.

Genome Annotation

A total of 176 Mb repetitive regions, accounting for 26.2% of *P. prolifica* genome, were annotated (supplementary table S4, Supplementary Material online, supplementary data S1 and S2, Supplementary Material online), which is similar to *Poecilia reticulata* (27.9%). DNA transposons subfamily Tc1-IS630-Pogo (Gao et al. 2020) is the most abundant class of transposable elements in both genomes. *Poecilia reticulata* has more than 2-fold greater abundance of RTE/Bov-B retrotransposons, whereas L1/CIN4 elements are more abundant in *P. prolifica* (supplementary table S4, Supplementary Material online). Noticeably, rolling circles are absent in the *Poecilia reticulata* genome, but 118 kb is found in the *P. prolifica* genome (supplementary table S4, Supplementary Material online). Based on the repeat-masked genome, 28,799 protein-coding genes were annotated by the MAKER pipeline. In addition, 24,714 protein-coding gene models were identified based on the *Poecilia reticulata* genome (GCA_000633615) using GeMoMa (Keilwagen et al. 2019). After merging the two gene sets, a total of 28,632 gene models with the best support were retained as the final gene models (supplementary data S3, Supplementary Material online). Among these, 27,227 (95.1%) genes are complete with a start codon and a stop codon, of which 23,515 genes (86.4%) were assigned to orthogroups by OrthoFinder, and 21,526 orthologs (79.1%) were shared with *Poecilia reticulata* (94.2%). In addition, a total of 5,549 noncoding RNAs (ncRNAs) were annotated in *P. prolifica* genome (supplementary table S5, Supplementary Material online). The *P. prolifica* genome is also annotated using the NCBI Eukaryotic Genome Annotation Pipeline (version 10.1). A total of 23,101 protein-coding genes and 3,820 noncoding genes were identified and characterized in RefSeq (GCF_027474105.1-RS_2023_04), and 22,700 (98.3%) protein-coding genes among these were assigned to orthogroups.

Synteny Analysis between *P. prolifica* and *Poecilia reticulata*

Of the 415 *P. prolifica* scaffolds, the top 55 largest scaffolds (520 Mb in 674 Mb) were mapped to the 23 chromosomes in *Poecilia reticulata* (Fraser et al. 2020), accounting for 77.2% of the total genome length in *P. prolifica*. The results showed a high level of synteny between the genome of the two species, and each of the *Poecilia reticulata* chromosomes corresponds to *P. prolifica* scaffolds (fig. 1). A small number of translocation and inversion events can be detected and visualized in the Circos plot (fig. 1), suggesting a moderate level of genome rearrangement between the two species. The *Poecilia reticulata* sex chromosome (Wright et al. 2017) corresponds to SCAFFOLD15 and SCAFFOLD35 in *P. prolifica* (chromosome 12 in the guppy genome).

Phylogenomic Analysis with Other Teleost Species

To understand the phylogenetic relationship among *P. prolifica* and other 14 Actinopterygii species, 4,301 single-copy orthologs shared by 15 species were exploited to construct the phylogenetic tree. The results showed that the species in the fish family Poeciliidae were clustered in the same clade, and *P. prolifica* has a closer evolutionary relationship with the southern platyfish *Xiphophorus maculatus* rather than the guppy *Poecilia reticulata* (supplementary fig. S2, Supplementary Material online), which is consistent with the results of previous multigenome alignment studies (Van Kruistum et al. 2021).

Materials and Methods

Sample Collection

Poeciliopsis prolifica (NCBI Taxonomy ID 188132) Line A and Line B used in this study were collected from the Rio El Palillo (Sinaloa State, Mexico) by Dr David Reznick and have been developed in laboratory by mating siblings for up to 12 generations (Line A) or 9 generations (Line B) of sib-sib matings. All procedures were approved by the University of California, Riverside Institutional Animal Care and Use Committee with protocol number 20170006. The blood samples and whole body of juvenile fish were collected after euthanasia using 100 mg/L buffered MS-222 (tricaine methanesulfonate, Syndel Inc., Ferndale, WA, USA).

Genomic DNA Extraction and Genome Sequencing

High molecular weight DNA samples were extracted using the Qiagen Genomic-Tip 20/G kit (Qiagen, Redwood City, CA, USA) from blood samples of a single male fish (*P. prolifica* line A). The quality, size distribution, and DNA integrity were determined by TapeStation 4200 using the Genomics ScreenTape (Agilent Technologies, Santa Clara, CA, USA). PacBio single-molecule real-time (SMRT) library was prepared on 20 μg male *P. prolifica* HMW genomic DNA using SMRTbell prep kit (Pacific Biosciences, Menlo Park, CA, USA) and sequenced in a total of eight SMRT cells (supplementary table S1, Supplementary Material online) on PacBio Sequel at the HudsonAlpha Genome Sequencing Center (Huntsville, AL, USA). A 10× Genomics linked-read library was performed on 1.1 ng of genomic DNA from the same male sample, using the Chromium Genome Reagent Kit v2 and Chromium Genome Library & Gel Bead Kit version 2 on a 10× Genomics Chromium Controller (10× Genomics, Inc., San Francisco, CA, USA). The final library concentration was measured on a Qubit 3.0 Fluorometer (Thermo Fisher Scientific, Waltham, MA, USA). The library size distribution was determined by LabChip GX Touch HT Nucleic Acid Analyzer (PerkinElmer, Hopkinton, MA, USA).

Total RNA Extraction and RNA-Seq

Total RNA samples were extracted from the whole body of juvenile *P. prolifica* ($n = 3$ for Line A and $n = 3$ for Line B), embryos and placenta ($n = 1$ for Line A and $n = 1$ for Line B), using the RNeasy Mini Plus kit (Qiagen, Redwood City, CA, USA). Standard double-strand RNA-seq libraries were constructed using the Illumina TruSeq RNA Library Prep Kit (Illumina, San Diego, CA, USA) and the NEXTFLEX Rapid RNA-Seq Kit for Illumina Sequencing (PerkinElmer, Hopkinton, MA, USA). To obtain information on the direction of transcription, single-strand RNA-seq libraries were prepared using the NEXTFLEX Rapid Directional RNA-Seq Kit (PerkinElmer, Hopkinton, MA, USA) and TruSeq Stranded mRNA Library Prep kit (Illumina, San Diego, CA, USA). The libraries were sequenced in a 2×100 bp paired-end (PE) setting on an Illumina HiSeq2000 machine (supplementary table S3, Supplementary Material online).

Genome Assembly

The reads sequenced by Illumina were trimmed by Trimmomatic (version 0.36; Bolger et al. 2014) and used for genome size, heterozygosity, and repeat length estimated using GenomeScope (Vurture et al. 2017) based on K-mer (Marcais and Kingsford 2012) frequency distributions with default parameters. A de novo assembly of 10× Genomics linked-reads was generated using Supernova version 2.1.1 with default parameters (Weisenfeld et al. 2017). The trimmed PacBio reads were assembled using CANU version 2.0 (Koren et al. 2017). The PacBio long-read and 10× Genomics linked-read libraries were combined using quickmerge version 0.3.0 with default parameters (Chakraborty et al. 2016) to obtain the scaffold assembly, and a final high-quality assembly was generated by Pilon (version 1.24); Walker et al. 2014) using 58× of Illumina 150-bp PE reads. The contigs were aligned and inspected for overlap using the Geneious software version 11.1.15 using the most stringent de novo assembly option (Kearse et al. 2012), and assembly summary statistics were computed by the stats.sh script in BBMap (Bushnell; Table 1). The completeness of the *P. prolifica* assembly was determined with BUSCO version 5.3.2 (Seppey et al. 2019) against the actinopterygii_odb10 database. The BUSCO scores were also evaluated on the reference guppy (*Poecilia reticulata*) genome assemblies GCA_000633 615.2 (Kunstner et al. 2016) and GCA_904066995.1 (Fraser et al. 2020).

Gene Prediction and Annotation

Repetitive elements annotation in our *P. prolifica* assembly was performed using RepeatModeler version 2.0.1 (Flynn et al. 2020), and the interspersed repeats sequences and low-complexity DNA sequences were masked with

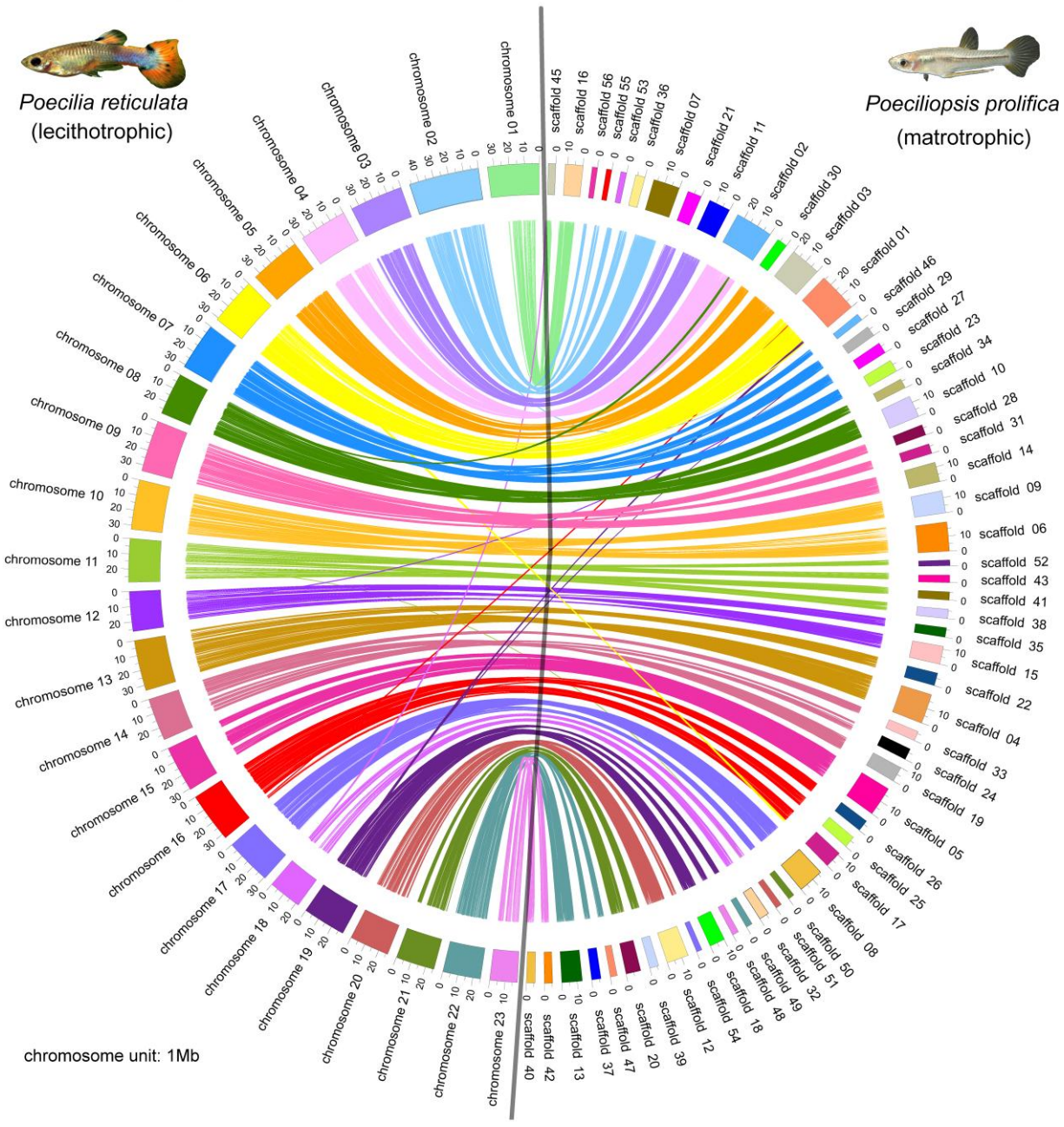

Fig. 1.—Genome comparisons between *Poeciliopsis prolifica* and the reference guppy *Poecilia reticulata* based on orthologous genes. A total of 55 largest scaffolds in the *P. prolifica* genome assembly showed a one-to-one relationship with 23 chromosomes in the *Poecilia reticulata* genome. The chromosomes on the left of the circle represent *Poecilia reticulata* chromosomes, and the scaffolds on the right of the circle represent *P. prolifica* scaffolds.

RepeatMasker version 4.0.6 (Tarailo-Graovac and Chen 2009). To assemble a set of prior transcript models from RNA-seq data, RNA-seq reads from Line A and Line B (supplementary table S3, Supplementary Material online) were trimmed by Trimmomatic (version 0.36) (Bolger et al. 2014) and then mapped to repeat-masked assembly with Tophat version 2.1.1 (Kim et al. 2013). Transcript isoform information was extracted using cufflinks version 2.2.1 (Trapnell et al. 2012) and inferred from de novo transcript assembly by Trinity version 2.4.0 (Haas et al. 2013). De novo gene prediction was performed by MAKER annotation pipeline version 2.31.9 (Cantarel et al. 2008) based on the transcriptome evidence from our RNA-seq assembly and homology evidence of annotated teleost protein sequences in the OrthoDB database v10 (Zdobnov et al. 2017). For homology-based gene prediction, Gene Model

Mapper (GeMoMa) annotation pipeline version 1.8 (Keilwagen et al. 2019) was applied to identified *P. prolifica* protein-coding genes based on the guppy genome (GCA_000633615.2). Finally, MAKER and GeMoMa gene sets were merged, and the genes with a start codon and a stop codon were selected to generate the final gene model (supplementary data S3, Supplementary Material online). For noncoding gene annotation, rRNAs were predicted using Barrnap version 0.9 with default parameters (Seemann 2013) and INFERNAL version 1.1.2, tRNAs were predicted by using tRNAscan-SE version 2.0.9 (Chan and Lowe 2019), and miRNAs, snoRNAs, and snRNAs were identified by Rfam/INFERNAL version 1.1.2 based on Rfam database version 14.7 (Kalvari et al. 2021; supplementary data S4, Supplementary Material online).

Orthologs Identification and Comparative Genomic Analysis

OrthoFinder (version 2.5.4; Emms and Kelly 2019) was employed to identify the orthologs shared by *P. prolifica* and the guppy genome (Kunstner et al. 2016). To analyze chromosome structural changes, MCScanX was used to identify the homologous regions of *P. prolifica* genome and *Poecilia reticulata* genome with default parameters. The putative gene pairs and linked relationships were visualized by Circos with required configuration files (Krzywinski et al. 2009). Phylogenetic relationships were analyzed for *P. prolifica* and five species in the family Poeciliidae, as well as nine additional species of teleost fish. The protein sequences of 4,301 single-copy orthologs shared in all 15 genomes were aligned by MAFFT (Katoh and Standley 2013), and concatenated sequences were used to construct a phylogenetic tree using the maximum-likelihood method in FastTree (Price et al. 2010) with 1,000 bootstraps. The phylogenetic tree was rerooted and visualized in FigTree version 1.4.4 (Rambaut).

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online (http://www.gbe.oxfordjournals.org/).

Acknowledgments

X.W. is supported by the National Science Foundation EPSCoR RII Track-4 award (1928770), the USDA National Institute of Food and Agriculture Hatch project USDA-NIFA-ALA05-2-18041, an Alabama Agriculture Experiment Station (AAES) Agriculture Research Enhancement, Exploration, and Development (AgR-SEED) award, and a laboratory start-up fund from Auburn University College of Veterinary Medicine. Y.Z. is partially supported by the China Scholarship Council. D.R. was supported by the National Science Foundation (DEB-0416085, DEB-1754669). We thank the HudsonAlpha Genome Sequencing Center for assistance with the PacBio sequencing and Gloria (Xiaozhu) Wang for the genomic DNA extraction. We acknowledge the Auburn University Easley Cluster for support of this work. We thank Françoise Thibaud-Nissen at NCBI for assistance with the RefSeq annotation.

Author Contributions

X.W. and D.R. contributed to the conception and design of the study. Y.R. and D.R. performed fish breeding and sample dissections. Y.Z. performed the DNA and RNA sequencing experiments. Y.Z. and X.W. performed the genomic data analysis. X.W. and D.R. provided samples, resources, and analysis tools. Y.Z. and X.W. wrote the first draft of the manuscript. D.R. wrote sections of the manuscript. All authors contributed to the manuscript revision, read, and approved the submitted version.

Data Availability

The draft genome assembly of *Poeciliopsis prolifica* has been deposited at NCBI under Assembly number JAPDDK000000000. The eight SMAT cells of long-read sequencing data generated by the PacBio Sequel instrument and the linked-read Illumina sequencing data were deposited at NCBI Sequence Read Archive (SRA) under accession number PRJNA893701. The Illumina RNA-seq data are available at NCBI GEO (Gene Expression Omnibus) databases under the accession number GSE221844. The NCBI RefSeq annotation can be accessed at annotation release GCF_027474105.1-RS_2023_04 with assembly accession number GCF_027474105.1. Supplemental Data files, including gene/repeat/ncRNA annotations, can be accessed at https://doi.org/10.35099/aurora-577.

Literature Cited

Amoroso EC. 1968. The evolution of viviparity. Proc R Soc Med. 61: 1188–1200.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120.

Braz HB, Scartozzoni RR, Almeida-Santos SM. 2016. Reproductive modes of the south American water snakes: a study system for the evolution of viviparity in squamate reptiles. Zool Anz J Compar Zool. 263:33–44.

Bushnell B. BBMap: a fast, accurate, splice-aware aligner. 2014. Available from: https://www.osti.gov/biblio/1241166. Accessed on November 4, 2021.

Cantarel BL, et al. 2008. MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. Genome Res. 18:188–196.

Chakraborty M, Baldwin-Brown JG, Long AD, Emerson JJ. 2016. Contiguous and accurate de novo assembly of metazoan genomes with modest long read coverage. Nucleic Acids Res. 44:e147.

Chan PP, Lowe TM. 2019. tRNAscan-SE: searching for tRNA genes in genomic sequences. Methods Mol Biol. 1962:1–14.

Clutton-Brock T. 1991. The evolution of parental care. Princeton (NJ): Princeton University Press.
Dimasuay KG, Boeuf P, Powell TL, Jansson T. 2016. Placental responses to changes in the maternal environment determine fetal growth. Front Physiol. 7:12.
Emms DM, Kelly S. 2019. Orthofinder: phylogenetic orthology inference for comparative genomics. Genome Biol. 20:238.
Exbrayat J-M. 2006. Reproductive biology and phylogeny of Gymnophiona: Caecilians. Enfield (NH): Science Publishers Enfield.
Faber JJ, Thornburg KL, Binder ND. 1992. Physiology of placental transfer in mammals. Am Zool. 32:343–354.
Flynn JM, et al. 2020. Repeatmodeler2 for automated genomic discovery of transposable element families. Proc Natl Acad Sci U S A. 117: 9451–9457.
Fraser BA, et al. 2020. Improved reference genome uncovers novel sex-linked regions in the guppy (*Poecilia reticulata*). Genome Biol Evol. 12:1789–1805.
Furness AI, Capellini I. 2019. The evolution of parental care diversity in amphibians. Nat Commun. 10:4709.
Gao B, et al. 2020. Evolution of pogo, a separate superfamily of IS630-tc1-mariner transposons, revealing recurrent domestication events in vertebrates. Mob DNA. 11:25.
Guillette LJ. 1993. The evolution of viviparity in lizards. BioScience 43: 742–751.
Haas BJ, et al. 2013. De novo transcript sequence reconstruction from RNA-Seq using the Trinity platform for reference generation and analysis. Nat Protoc. 8:1494–1512.
Kalvari I, et al. 2021. Rfam 14: expanded coverage of metagenomic, viral and microRNA families. Nucleic Acids Res. 49:D192–D200.
Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 30:772–780.
Kaye M. 1971. The evolution of placentation. Aust N Z J Obstet Gynaecol. 11:197–207.
Kearse M, et al. 2012. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647–1649.
Keilwagen J, Hartung F, Grau J. 2019. Gemoma: homology-based gene prediction utilizing intron position conservation and RNA-Seq data. Methods Mol Biol. 1962:161–177.
Kim D, et al. 2013. Tophat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 14:R36.
Koren S, et al. 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Res. 27:722–736.
Krzywinski M, et al. 2009. Circos: an information aesthetic for comparative genomics. Genome Res. 19:1639–1645.
Kunstner A, et al. 2016. The genome of the Trinidadian guppy, *Poecilia reticulata*, and variation in the Guanapo population. PLoS One. 11: e0169087.
Marcais G, Kingsford C. 2012. Jellyfish: a fast k-mer counter. Tutorialis e Manuais. 1:1–8.
Moffett A, Loke C. 2006. Immunology of placentation in eutherian mammals. Nat Rev Immunol. 6:584–594.
Mossman HW. 1991. Classics revisited: comparative morphogenesis of the fetal membranes and accessory uterine structures. Placenta 12: 1–5.
Pollux BJ, Meredith RW, Springer MS, Garland T, Reznick DN. 2014. The evolution of the placenta drives a shift in sexual selection in livebearing fish. Nature 513:233–236.
Pollux B, Pires M, Banet A, Reznick D. 2009. Evolution of placentas in the fish family Poeciliidae: an empirical study of macroevolution. Ann Rev Ecol Evol Syst. 40:271–289.
Price MN, Dehal PS, Arkin AP. 2010. Fasttree 2–approximately maximum-likelihood trees for large alignments. PLoS One. 5:e9490.
Rambaut A. FigTree v1.4.4. 2014. Available from: http://tree.bio.ed.ac.uk/software/figtree/. Accessed on February 28, 2022.
Reznick DN, Mateos M, Springer MS. 2002. Independent origins and rapid evolution of the placenta in the fish genus Poeciliopsis. Science 298:1018–1020.
Rothchild I. 2003. The yolkless egg and the evolution of eutherian viviparity. Biol Reprod. 68:337–357.
Seemann T. barrnap 0.9: rapid ribosomal RNA prediction. 2013. Available from: https://github.com/tseemann/barrnap. Accessed on March 30, 2022.
Seppey M, Manni M, Zdobnov EM. 2019. BUSCO: assessing genome assembly and annotation completeness. Methods Mol Biol. 1962:227–245.
Stewart JR, Thompson MB. 2000. Evolution of placentation among squamate reptiles: recent research and future directions. Comp Biochem Physiol A Mol Integr Physiol. 127:411–431.
Tarailo-Graovac M, Chen N. 2009. Using RepeatMasker to identify repetitive elements in genomic sequences. Curr Protoc Bioinformatics. 25: 4.10.1–4.10.14. doi:10.1002/0471250953.bi0410s25
Thibault RE, Schultz RJ. 1978. Reproductive adaptations among viviparous fishes (Cyprinodontiformes: Poeciliidae). Evolution 32:320–333.
Trapnell C, et al. 2012. Differential gene and transcript expression analysis of RNA-Seq experiments with TopHat and Cufflinks. Nat Protoc. 7:562–578.
Uribe MC, Cerda-Jardon PI, Blackburn DG. 2021. Morphological basis for maternal nutrient provision to embryos in the viviparous fish *Ataeniobius toweri* (Teleostei: Goodeidae). J Morphol. 282: 1575–1586.
Uribe MC, De la Rosa Cruz G, García Alarcón A, Campuzano Caballero JC, Guzmán Bárcenas MG. 2019. Structures associated with oogenesis and embryonic development during intraovarian gestation in viviparous teleosts (Poeciliidae). Fishes 4:35.
van Kruistum H, et al. 2021. Parallel genomic changes drive repeated evolution of placentas in live-bearing fish. Mol Biol Evol. 38: 2627–2638.
Vurture GW, et al. 2017. Genomescope: fast reference-free genome profiling from short reads. Bioinformatics 33:2202–2204.
Wake MH. 1992. Evolutionary scenarios, homology and convergence of structural specializations for vertebrate viviparity. Am Zool. 32: 256–263.
Wake MH. 2015. Fetal adaptations for viviparity in amphibians. J Morphol. 276:941–960.
Wake MH, Dickie R. 1998. Oviduct structure and function and reproductive modes in amphibians. J Exp Zool. 282:477–506.
Walker BJ, et al. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One. 9:e112963.
Weisenfeld NI, Kumar V, Shah P, Church DM, Jaffe DB. 2017. Direct determination of diploid genome sequences. Genome Res. 27:757–767.
Wourms JP. 1981. Viviparity: the maternal-fetal relationship in fishes. Am Zool. 21:473–515.
Wourms JP, Grove BD, Lombardi J. 1988. Fish Physiology. Vol. 11B. San Diego (CA): Academic Press. p. 1–134.
Wright AE, et al. 2017. Convergent recombination suppression suggests role of sexual selection in guppy sex chromosome formation. Nat Commun. 8:14251.
Zdobnov EM, et al. 2017. OrthoDB v9.1: cataloging evolutionary and functional annotations for animal, fungal, plant, archaeal, bacterial and viral orthologs. Nucleic Acids Res. 45:D744–D749.

Associate editor: Christopher Wheat