



# **FishmiRNA: An evolutionarily supported microRNA annotation and expression database for ray-finned fishes**

Thomas Desvignes, Philippe Bardou, Jérôme Montfort, Jason Sydes, Cervin Guyomar, Simon George, John H Postlethwait, Julien Bobe

## **► To cite this version:**

Thomas Desvignes, Philippe Bardou, Jérôme Montfort, Jason Sydes, Cervin Guyomar, et al.. FishmiRNA: An evolutionarily supported microRNA annotation and expression database for ray-finned fishes. *Molecular Biology and Evolution*, 2022, 39 (2), pp.msac004. 10.1093/molbev/msac004 . hal-03528022

**HAL Id: hal-03528022**

**<https://hal.inrae.fr/hal-03528022>**

Submitted on 17 Jan 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

# FishmiRNA: An evolutionarily supported microRNA annotation and expression database for ray-finned fishes

Thomas Desvignes <sup>1#</sup>, Philippe Bardou <sup>2</sup>, Jérôme Montfort <sup>3</sup>, Jason Sydes <sup>1</sup>, Cervin Guyomar <sup>2</sup>, Simon George <sup>4</sup>, John H. Postlethwait <sup>1</sup>, Julien Bobe <sup>3#</sup>

1. Institute of Neuroscience, University of Oregon, Eugene OR 97403, USA
2. Sigene, GenPhySE, Université de Toulouse, INRAE, ENVT, F-31326, Castanet Tolosan, France
3. INRAE, LPGP, 35000, Rennes, France
4. MGX - Montpellier GenomiX, Univ. Montpellier, CNRS, INSERM, Montpellier France

<sup>#</sup> Correspondence to [tdesvign@uoregon.edu](mailto:tdesvign@uoregon.edu) and [julien.bobe@inrae.fr](mailto:julien.bobe@inrae.fr)

## Email addresses:

Thomas Desvignes: [tdesvign@uoregon.edu](mailto:tdesvign@uoregon.edu)

Philippe Bardou: [philippe.bardou@inrae.fr](mailto:philippe.bardou@inrae.fr)

Jérôme Montfort: [jerome.montfort@inrae.fr](mailto:jerome.montfort@inrae.fr)

Jason Sydes: [sydes@uoregon.edu](mailto:sydes@uoregon.edu)

Cervin Guyomar: [cervin.guyomar@inrae.fr](mailto:cervin.guyomar@inrae.fr)

Simon George: [simon.george@mgx.cnrs.fr](mailto:simon.george@mgx.cnrs.fr)

John H. Postlethwait: [jpostle@uoregon.edu](mailto:jpostle@uoregon.edu)

Julien Bobe: [julien.bobe@inrae.fr](mailto:julien.bobe@inrae.fr)

## ORCID:

Thomas Desvignes: 0000-0001-5126-8785

Philippe Bardou: 0000-0002-0004-0251

Jason Sydes: 0000-0003-4187-2828

Cervin Guyomar: 0000-0003-2707-2541

John H. Postlethwait: 0000-0002-5476-2137

Julien Bobe: 0000-0002-9355-8227

**Running title:** FishmiRNA, the fish microRNA annotation and expression database

**Competing interest statement:** The author(s) declare no competing interests.

**Data availability statement:** All data generated or analyzed during this study are included in the published article, in the FishmiRNA database, and/or deposited in NCBI.

## Abstract

microRNAs are important post-transcriptional regulators of gene expression involved in countless biological processes and are widely studied across metazoans. While miRNA research continues to grow, the large community of fish miRNA researchers lacks exhaustive resources consistent among species. To fill this gap, we developed FishmiRNA, an evolutionarily supported microRNA annotation and expression database for ray-finned fishes: [www.fishmirna.org](http://www.fishmirna.org). The self-explanatory database contains detailed, manually-curated miRNA annotations with orthology relationships rigorously established by sequence similarity and conserved syntenies, and expression data provided for each detected mature miRNA. In just few clicks, users can download the annotation and expression database in several convenient formats either in its entirety or a subset. Simple filters and BLAST search options also permit the simultaneous exploration and visual comparison of expression data for up to any ten mature miRNAs across species and organs. FishmiRNA was specifically designed for ease of use to reach a wide audience.

## Keywords

Teleost, Holostei, actinopterygian, non-coding RNA, Whole Genome Duplication, bowfin *Amia calva*

## Introduction

microRNAs (miRNAs) have emerged as key post-transcriptional regulators of gene expression that act by binding to the 3'-untranslated region of target messenger RNAs when incorporated into the RNA-induced silencing complex (Jonas and Izaurralde 2015; Bartel 2018). Regulatory functions of miRNAs have now been implicated in countless biological processes, including cell differentiation and proliferation, organ development and physiology, pathologies and diseases (Mendell and Olson 2012; Sun and Lai 2013), and genetic noise buffering, especially in stressful conditions (Schmiedel et al. 2015; Liufu et al. 2017). Furthermore, miRNAs evolve in lineage- and environment-specific manners and can modulate alternative developmental and physiological pathways that may influence adaptation, diversification, and speciation (Loh et al. 2011; Li and Zhang 2013; Quah et al. 2015).

Although several databases and bioinformatic tools facilitate the study of miRNAs in cellular, developmental, physiological, and pathological contexts, they cover mostly humans and main laboratory model organisms, so researchers studying other species often lack appropriate resources to accurately facilitate miRNA studies in their systems. This situation is exacerbated in ray-finned fishes, actinopterygians. While several valuable miRNA annotation databases already exist, miRNA annotations in fish species are often unavailable, incomplete, or inconsistent across fish phylogeny, inhibiting the study of miRNAs in half of all vertebrate species. The legacy database miRBase was

created in 2006 but has not been updated since 2018 (Kozomara et al. 2019). miRBase contains partial miRNA annotations for several fish species, but none are exhaustive or curated and many new or improved annotations have not been incorporated. In 2015, MirGeneDB was created and covers a breadth of metazoan species, but this database provides miRNA annotation for only one fish species, zebrafish (Fromm et al. 2020). miRNAs, however, are of great interest in fish research with studies ranging from genomic evolution (Xiong et al. 2019; Desvignes et al. 2021), development (Giraldez et al. 2006; Kasper et al. 2017; Gay et al. 2018), medical models (Hsu et al. 2017), aquaculture (Herkenhoff et al. 2018; Blödorn et al. 2021; Cardona et al. 2021), pathology (Andreassen and Høyheim 2017; Wang et al. 2018), toxicology (Goodale et al. 2019; Ahkin Chin Tai and Freeman 2020), to adaptation and speciation (Franchini et al. 2019; Kelley et al. 2021). In most fish miRNA studies, the lack of consistent annotation resources leads to using, by default, combinations of annotations existing in other, often distantly-related species, or to *de novo* prediction of miRNAs using different algorithms in different research groups, leading to inconsistent results among species. Therefore, while miRNA research continues to grow rapidly, the large community of fish miRNA researchers is plagued by the lack of exhaustive and phylogenetically supported resources.

The complexity of studying miRNAs in fish emerges from two main sources. First, a whole-genome duplication, called the Teleost Genome Duplication or TGD, initially duplicated every miRNA gene (Amores et al. 1998; Taylor et al. 2003; Braasch et al. 2016). This genome duplication was followed by lineage-specific gene resolution events that confound orthology assignment among species (Postlethwait 2007). Second, because fish represent more than half of living vertebrate species and inhabit virtually all aquatic habitats on the planet, they harbor dramatic variations in gene repertoires, which makes miRNA gene and mature miRNA computational predictions and annotations unreliable without expert manual curation.

To fill this resource gap, we developed FishmiRNA: An evolutionarily supported microRNA annotation and expression database for ray-finned fishes.

## New approaches

FishmiRNA integrates several new approaches that provide accurate, consistent, and exhaustive annotations of evolutionarily-conserved miRNAs in ray-finned fishes, as well as the innovative ability to explore expression datasets across species and organs in a user-friendly interface.

Two novelties of FishmiRNA for achieving miRNA annotation consistency among species are the full integration of the teleost genome duplication (TGD) in gene orthology assignments and a broad phylogenetic context. For that purpose, we built on our recent establishment of genome-wide synteny-verified miRNA gene orthologies among several fish species, including the spotted gar

*Lepisosteus oculatus* whose lineage, the Holostei, diverged before the TGD. This work allowed us to infer the miRNA gene repertoire of the hypothetical Teleost-Holostei last common ancestor (TH-LCA) (Desvignes et al. 2021). In addition, FishmiRNA annotations rely on small RNA sequencing expression data to first, identify miRNA gene loci and second, to detect the most abundantly expressed mature miRNAs, leading to data-supported annotation of both the 5p and 3p strands for the majority of genes (76% across the database). The novel annotation of four teleost species and of the bowfin, *Amia calva*, a second holostean outgroup to the teleosts (Thompson et al. 2021), further increased confidence in the inferred TH-LCA miRNA gene repertoire and in intermediate teleost ancestors. This ancestral reconstruction approach, a cornerstone of the FishmiRNA database, allows the retracing of gene evolution across lineages. Each miRNA gene annotated in FishmiRNA is thus linked to its orthologs among other teleosts and with the TH-LCA. So far, FishmiRNA contains miRNA gene and mature miRNA annotations for 10 actinopterygian species, including eight teleost species (zebrafish, catfish, panga, medaka, molly, perch, stickleback, and icefish) and two Holostei species that diverged before the TGD (gar and bowfin) (Table 1). These species were selected based on their broad phylogenetic distribution within ray-finned fishes, on their importance in evolutionary, aquacultural, and biomedical research, and on the availability of high-quality genome sequences.

A unique approach of FishmiRNA unavailable in any other annotation database is to provide expression data for mature miRNAs. The consistent re-analysis of expression data for each species using the smallRNA-seq software *Prost!* (Desvignes et al. 2019) coupled with a graphical module enables users to compare the expression patterns of up to ten miRNAs from any species in the database, thus offering an innovative opportunity to incorporate in a study the evolutionary conservation of expression of any miRNA across species.

Finally, an important novel approach of FishmiRNA is the simplicity and efficient design of its user-friendly website. Contrary to other databases that incorporate hundreds of webpages, FishmiRNA database relies on a single webpage and two spreadsheets: an annotation table and an expression table, both of which can be interactively and quickly filtered, searched, and exported, in full or in part. This novel approach eases access to miRNA annotations and miRNA expression data and significantly increases the diversity of users reached by providing accessibility to miRNA data to anyone, with or without bioinformatic skills.

## Database features

### Graphic information

A summary of the annotation database is displayed at the top of the page in a phylogenetic context (Rabosky et al. 2018), including graphical representations of general statistics for each

species (Fig. 1). These immediate visual representations help users determine which species' annotation would be the most appropriate to apply to their experimental question or species of interest. For example, if FishmiRNA lacks a species, a user might select the phylogenetically most closely related species with the largest miRNA annotation. The phylogenetic coverage of FishmiRNA will grow as new species become annotated.

### The FishmiRNA annotation database

The FishmiRNA annotation database consists of a single table containing 38 columns and as many lines as there are miRNA genes in the database, currently 3028 genes. Four columns present the taxonomy of the species and its reference genome assembly; six columns describe the miRNA gene with a name, potential previous names, a unique gene identifier, links to miRBase, Ensembl, and to other databases when available; six columns describe the miRNA hairpin with a name, its sequence, position, and strand in the genome assembly; two columns report gene orthology among teleosts and with the Teleost-Holostei last common ancestor (TH-LCA); three columns inform on gene clustering; one column summarizes 5p and 3p strand annotations, each detailed in eight columns providing the mature strand name, a unique mature identifier, a reference sequence, its position in the genome, and whether this mature miRNA can be produced by more than one miRNA gene. All of this information can be simultaneously visualized for each gene in an “individual gene ID card” by clicking on the magnifier button associated with the miRNA gene name in the annotation table (Fig. 2).

For rendering purposes, only nine columns are displayed by default, but users can select their own column preferences or display the entire table (Fig. 2). Users can also filter the database by species, by FishmiRNA Gene ID, hairpin and mature miRNA names, and by orthology relationships among teleosts and with the TH-LCA. In addition, users can search the database by genomic location or by a BLAST search of sequences on hairpins or mature miRNA sequences (Fig. 2). The entirety or filtered parts of the annotation database can be immediately exported as a spreadsheet or a GFF file. Hairpin and mature miRNA sequences can also be exported in a click in FASTA format for ready use with smallRNA-seq analysis software or other purposes (Fig. 2).

The filtering of miRNA genes reduces the portion of the annotation table displayed to only genes of interest, which can then be individually selected using tick boxes to conveniently filter their respective mature miRNAs and explore their expression patterns in a single click (Fig. 2).

### The FishmiRNA expression database

The FishmiRNA expression database consists of a single table containing 23 columns and as many lines as mature miRNAs detected in the smallRNA-seq libraries analyzed. Three columns provide information on the mature miRNA name, its unique mature identifier, and the sequence of

the most-highly expressed isomiR; and 10 columns that contain the raw counts and 10 columns that contain normalized read counts (reads per million, RPM) for each mature miRNA in a selection of major organs (brain, gills, heart ventricle, skeletal muscle, intestine, liver, ovary, testis, hematopoietic kidney, and spleen). Not all species have expression data for all of the selected organs; the entry “No\_Data” signals these cases. For rendering purposes, only the mature miRNA name, its sequence, and normalized read counts are displayed by default. Similar to the annotation table, users can configure columns shown in the expression table, filter by species, mature miRNA name, unique identifier, or explore expression data by BLAST. The entire expression database or filtered parts of it can also be immediately exported (Fig. 3).

One of the most useful features of the FishmiRNA expression database is its graphical module. The expression profile of each mature miRNA can be visualized by a click on the blue histogram button associated with its name (Fig. 3). This link opens a modal box displaying expression data in a histogram with the organ of highest expression in red and the mean expression across organs marked by a horizontal red line. Furthermore, by selecting mature miRNAs using tick boxes next to the mature miRNA name, users can instantly plot on the same graph the expression profiles of up to any 10 mature miRNAs in the database. All graphs generated can also be exported in various image formats using the dropdown menu located at the top-right corner of the graph (Fig. 3).

### Quick download links and origin of analyzed data

To facilitate the dissemination, re-use, and transparency of data provided in FishmiRNA, the download section of the FishmiRNA database provides quick links for downloading, per species or for all species, annotation files in FASTA format, and raw output files from the smallRNA-seq software *Prost!* used for the annotation and expression analyses (Desvignes et al. 2019) (Fig. 4). Furthermore, links to the publication of each original annotation are provided along with links to the NCBI BioProject of the expression data analyzed and the article that published these original data (Fig. 4).

## Materials and Methods

Small RNA sequencing data presented in FishmiRNA are all publicly available in NCBI (Table 1). Organ sampling and library preparation protocols as well as sequencing platforms may differ between species and therefore expression patterns may not be fully comparable. All species were, however, re-analyzed the same way and for each species, Illumina sequencing libraries were simultaneously analyzed using *Prost!* (Desvignes et al. 2019), selecting for read length from 17 to 25 nucleotides with a minimum of five identical reads. Publicly available genome assemblies were used for each species (Table 1). Gene and mature miRNA annotations were performed as previously described (Desvignes et al. 2019), using orthology and ohnology relationships established across



species (Desvignes et al. 2021), and following nomenclature rules established for zebrafish (Desvignes et al. 2015; Ruzicka et al. 2019; Desvignes et al. 2020). FishmiRNA was developed based on the RumimiR web interface (Bourdon et al. 2019).

## Authors' contributions

Study concept and design: TD, JHP, and JB

Acquisition of data: TD, SG, and JB

Analysis and interpretation of data: TD, JM, JS, CG, JHP, and JB

Web interface development: PB

Wrote the manuscript: TD and JB

Critical revision of the manuscript: TD, PB, CG, JHP, and JB

Obtained funding: JHP, TD, PB, and JB

Study supervision: TD, JHP, and JB

## Acknowledgments

Authors would like to thank all our past and present fish miRNA collaborators for interactions that revealed the need for FishmiRNA. This work was supported by the National Institutes of Health (grant numbers NIH R24 OD011199, NIH 5R01 OD011116, and NIH R01 GM085318 to JHP), the National Science Foundation Office of Polar Program (NSF OPP-1543383 to JHP and TD), and the Agence Nationale de la Recherche (ANR-18-CE20-0004 to JB). This work benefited from access to the University of Oregon high performance computer, Talapas. MGX acknowledges financial support from France Génomique National infrastructure, funded as part of "Investissement d'Avenir" program managed by Agence National pour la Recherche (ANR-10-INBS-09).

## References cited

- Ahkin Chin Tai JK, Freeman JL. 2020. Zebrafish as an integrative vertebrate model to identify miRNA mechanisms regulating toxicity. *Toxicology Reports* 7:559–570.
- Amores A, Force A, Yan Y-L, Joly L, Amemiya C, Fritz A, Ho RK, Langeland J, Prince V, Wang Y-L, et al. 1998. Zebrafish hox Clusters and Vertebrate Genome Evolution. *Science* 282:1711–1714.
- Andreassen R, Høyheim B. 2017. miRNAs associated with immune response in teleost fish. *Developmental & Comparative Immunology* 75:77–85.
- Bartel DP. 2018. Metazoan MicroRNAs. *Cell* 173:20–51.



- Blödorn EB, Domingues WB, Nunes LS, Komninou ER, Pinhal D, Campos VF. 2021. MicroRNA roles and their potential use as selection tool to cold tolerance of domesticated teleostean species: A systematic review. *Aquaculture* 540:736747.
- Bourdon C, Bardou P, Aujean E, Le Guillou S, Tosser-Klopp G, Le Provost F. 2019. RumimiR: a detailed microRNA database focused on ruminant species. *Database* [Internet] 2019. Available from: <https://doi.org/10.1093/database/baz099>
- Braasch I, Gehrke AR, Smith JJ, Kawasaki K, Manousaki T, Pasquier J, Amores A, Desvignes T, Batzel P, Catchen J, et al. 2016. The spotted gar genome illuminates vertebrate evolution and facilitates human-teleost comparisons. *Nat Genet* 48:427–437.
- Cardona E, Guyomar C, Desvignes T, Montfort J, Guendouz S, Postlethwait JH, Skiba-Cassy S, Bobe J. 2021. Circulating miRNA repertoire as a biomarker of metabolic and reproductive states in rainbow trout. *BMC Biology* 19:235.
- Desvignes T, Batzel P, Berezikov E, Eilbeck K, Eppig JT, McAndrews MS, Singer A, Postlethwait JH. 2015. miRNA Nomenclature: A View Incorporating Genetic Origins, Biosynthetic Pathways, and Sequence Variants. *Trends in Genetics* 31:613–626.
- Desvignes T, Batzel P, Sydes J, Eames BF, Postlethwait JH. 2019. miRNA analysis with Prost! reveals evolutionary conservation of organ-enriched expression and post-transcriptional modifications in three-spined stickleback and zebrafish. *Scientific Reports* 9:3913.
- Desvignes T, Beam MJ, Batzel P, Sydes J, Postlethwait JH. 2014. Expanding the annotation of zebrafish microRNAs based on small RNA sequencing. *Gene* 546:386–389.
- Desvignes T, Detrich III HW, Postlethwait JH. 2016. Genomic conservation of erythropoietic microRNAs (erythromiRs) in white-blooded Antarctic icefish. *Marine Genomics* 30:27–34.
- Desvignes T, Loher P, Eilbeck K, Ma J, Urgese G, Fromm B, Sydes J, Aparicio-Puerta E, Barrera V, Espín R, et al. 2020. Unification of miRNA and isomiR research: the mirGFF3 format and the mirtop API. *Bioinformatics* 36:698–703.
- Desvignes T, Sydes J, Montfort J, Bobe J, Postlethwait JH. 2021. Evolution after Whole-Genome Duplication: Teleost MicroRNAs. *Molecular Biology and Evolution* 38:3308–3331.
- Franchini P, Xiong P, Fruciano C, Schneider RF, Woltering JM, Hulsey CD, Meyer A. 2019. MicroRNA Gene Regulation in Extremely Young and Parallel Adaptive Radiations of Crater Lake Cichlid Fish. *Mol Biol Evol* 36:2498–2511.
- Fromm B, Domanska D, Høye E, Ovchinnikov V, Kang W, Aparicio-Puerta E, Johansen M, Flatmark K, Mathelier A, Hovig E, et al. 2020. MirGeneDB 2.0: the metazoan microRNA complement. *Nucleic Acids Res* 48:D132–D141.

- Gay S, Bugeon J, Bouchareb A, Henry L, Delahaye C, Legeai F, Montfort J, Le Cam A, Siegel A, Bobe J, et al. 2018. MiR-202 controls female fecundity by regulating medaka oogenesis. *PLoS Genet.* 14:e1007593.
- Giraldez AJ, Mishima Y, Rihel J, Grocock RJ, Dongen SV, Inoue K, Enright AJ, Schier AF. 2006. Zebrafish MiR-430 Promotes Deadenylation and Clearance of Maternal mRNAs. *Science* 312:75–79.
- Goodale BC, Hampton TH, Ford EN, Jackson CE, Shaw JR, Stanton BA, King BL. 2019. Profiling microRNA expression in Atlantic killifish (*Fundulus heteroclitus*) gill and responses to arsenic and hyperosmotic stress. *Aquatic Toxicology* 206:142–153.
- Herkenhoff ME, Oliveira AC, Nachtigall PG, Costa JM, Campos VF, Hilsdorf AWS, Pinhal D. 2018. Fishing Into the MicroRNA Transcriptome. *Front. Genet.* [Internet] 0. Available from: <https://www.frontiersin.org/articles/10.3389/fgene.2018.00088/full>
- Hsu AY, Wang D, Gurol T, Zhou W, Zhu X, Lu H-Y, Deng Q. 2017. Overexpression of microRNA-722 fine-tunes neutrophilic inflammation by inhibiting Rac2 in zebrafish. *Disease Models & Mechanisms* 10:1323–1332.
- Jonas S, Izaurralde E. 2015. Towards a molecular understanding of microRNA-mediated gene silencing. *Nature Reviews Genetics* 16:421–433.
- Kasper DM, Moro A, Ristori E, Narayanan A, Hill-Teran G, Fleming E, Moreno-Mateos M, Vejnar CE, Zhang J, Lee D, et al. 2017. MicroRNAs Establish Uniform Traits during the Architecture of Vertebrate Embryos. *Developmental Cell* 40:552–565.e5.
- Kelley JL, Desvignes T, McGowan KL, Perez M, Rodriguez LA, Brown AP, Culumber Z, Tobler M. 2021. microRNA expression variation as a potential molecular mechanism contributing to adaptation to hydrogen sulphide. *Journal of Evolutionary Biology* 34:977–988.
- Kim B-M, Amores A, Kang S, Ahn D-H, Kim J-H, Kim I-C, Lee JH, Lee SG, Lee H, Lee J, et al. 2019. Antarctic blackfin icefish genome reveals adaptations to extreme environments. *Nature Ecology & Evolution* 3:469.
- Kozomara A, Birgaoanu M, Griffiths-Jones S. 2019. miRBase: from microRNA sequences to function. *Nucleic Acids Res* 47:D155–D162.
- Li J, Zhang Z. 2013. miRNA regulatory variation in human evolution. *Trends in Genetics* 29:116–124.
- Liufu Z, Zhao Y, Guo L, Miao G, Xiao J, Lyu Y, Chen Y, Shi S, Tang T, Wu C-I. 2017. Redundant and incoherent regulations of multiple phenotypes suggest microRNAs' role in stability control. *Genome Res.* 27:1665–1673.
- Loh Y-HE, Yi SV, Streelman JT. 2011. Evolution of microRNAs and the diversification of species. *Genome Biol Evol* 3:55–65.

- Mendell JT, Olson EN. 2012. MicroRNAs in Stress Signaling and Human Disease. *Cell* 148:1172–1187.
- Postlethwait JH. 2007. The zebrafish genome in context: ohnologs gone missing. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution* 308B:563–577.
- Quah S, Hui JHL, Holland PWH. 2015. A Burst of miRNA Innovation in the Early Evolution of Butterflies and Moths. *Mol Biol Evol* 32:1161–1174.
- Rabosky DL, Chang J, Title PO, Cowman PF, Sallan L, Friedman M, Kaschner K, Garilao C, Near TJ, Coll M, et al. 2018. An inverse latitudinal gradient in speciation rate for marine fishes. *Nature* 559:392–395.
- Ruzicka L, Howe DG, Ramachandran S, Toro S, Van Slyke CE, Bradford YM, Eagle A, Fashena D, Frazer K, Kalita P, et al. 2019. The Zebrafish Information Network: new support for non-coding genes, richer Gene Ontology annotations and the Alliance of Genome Resources. *Nucleic Acids Res* 47:D867–D873.
- Schmiedel JM, Klemm SL, Zheng Y, Sahay A, Blüthgen N, Marks DS, Oudenaarden A van. 2015. MicroRNA control of protein expression noise. *Science* 348:128–132.
- Sun K, Lai EC. 2013. Adult-specific functions of animal microRNAs. *Nat Rev Genet* 14:535–548.
- Taylor JS, Braasch I, Frickey T, Meyer A, Peer YV de. 2003. Genome Duplication, a Trait Shared by 22,000 Species of Ray-Finned Fish. *Genome Res.* 13:382–390.
- Thompson AW, Hawkins MB, Parey E, Wcisel DJ, Ota T, Kawasaki K, Funk E, Losilla M, Fitch OE, Pan Q, et al. 2021. The bowfin genome illuminates the developmental evolution of ray-finned fishes. *Nat Genet*:1–12.
- Wang M, Jiang S, Wu W, Yu F, Chang W, Li P, Wang K. 2018. Non-coding RNAs Function as Immune Regulators in Teleost Fish. *Front. Immunol.* [Internet] 0. Available from: <https://www.frontiersin.org/articles/10.3389/fimmu.2018.02801/full>
- Xiong P, Schneider RF, Hulsey CD, Meyer A, Franchini P. 2019. Conservation and novelty in the microRNA genomic landscape of hyperdiverse cichlid fishes. *Scientific Reports* [Internet] 9. Available from: <http://www.nature.com/articles/s41598-019-50124-0>

## Tables

Table 1: Summary of annotations and sequencing data included in FishmiRNA

| Species  | Genome Assembly                     | Annotation Citation     | Sequencing Data NCBI Acc # | Expression Data Citation |
|--|-------------------------------------|-------------------------|----------------------------|--------------------------|
| <b>Spotted gar, <i>Lepisosteus oculatus</i></b>                | LepOcu1<br>(GCA_000242695.1)        | (Braasch et al. 2016)   | PRJNA296503                | (Braasch et al. 2016)    |
| <b>Bowfin, <i>Amia calva</i></b>                               | AmiCal1<br>(GCA_017591415.1)        | Present Study           | PRJNA255850                | Present Study            |
| <b>Black bullhead, <i>Ameiurus melas</i></b>                   | AMELA_1.0<br>(GCA_012411365.1)      | Present Study           | PRJNA730692                | Present Study            |
| <b>Striped catfish, <i>Pangasianodon hypophthalmus</i></b>     | GENO_Phyp_1.0<br>(GCF_009078355.1)  | Present Study           | PRJNA256963                | Present Study            |
| <b>Zebrafish, <i>Danio rerio</i></b>                           | GRCz11<br>(GCA_000002035.4)         | (Desvignes et al. 2014) | PRJNA240316                | (Desvignes et al. 2014)  |
| <b>Japanese medaka, <i>Oryzias latipes</i></b>                 | ASM223467v1<br>(GCA_002234675.1)    | (Desvignes et al. 2021) | PRJNA477647                | (Gay et al. 2018)        |
| <b>Molly, <i>Poecilia mexicana</i></b>                         | P_mexicana-1.0<br>(GCA_001443325.1) | (Kelley et al. 2021)    | PRJNA471100                | (Kelley et al. 2021)     |
| <b>European perch, <i>Perca flavescens</i></b>                 | GENO_Pfluv_1.0<br>(GCA_010015445.1) | Present Study           | PRJNA256973                | Present Study            |
| <b>Three-spined stickleback, <i>Gasterosteus aculeatus</i></b> | BROAD S1                            | (Desvignes et al. 2019) | PRJNA486149                | (Desvignes et al. 2019)  |
| <b>Blasckfin icefish, <i>Chaenocephalus aceratus</i></b>       | cace 20180227a<br>pilon1            | (Kim et al. 2019)       | PRJNA310135                | (Desvignes et al. 2016)  |

## Figure legends

Fig. 1. General statistics of the FishmiRNA database.

Screen shot of the home page. On the left, the species present in FishmiRNA database are displayed in their phylogenetic context (Rabosky et al. 2018). On the right, the miRNA gene (in blue) and mature miRNA (in red) annotation statistics are given for each species.

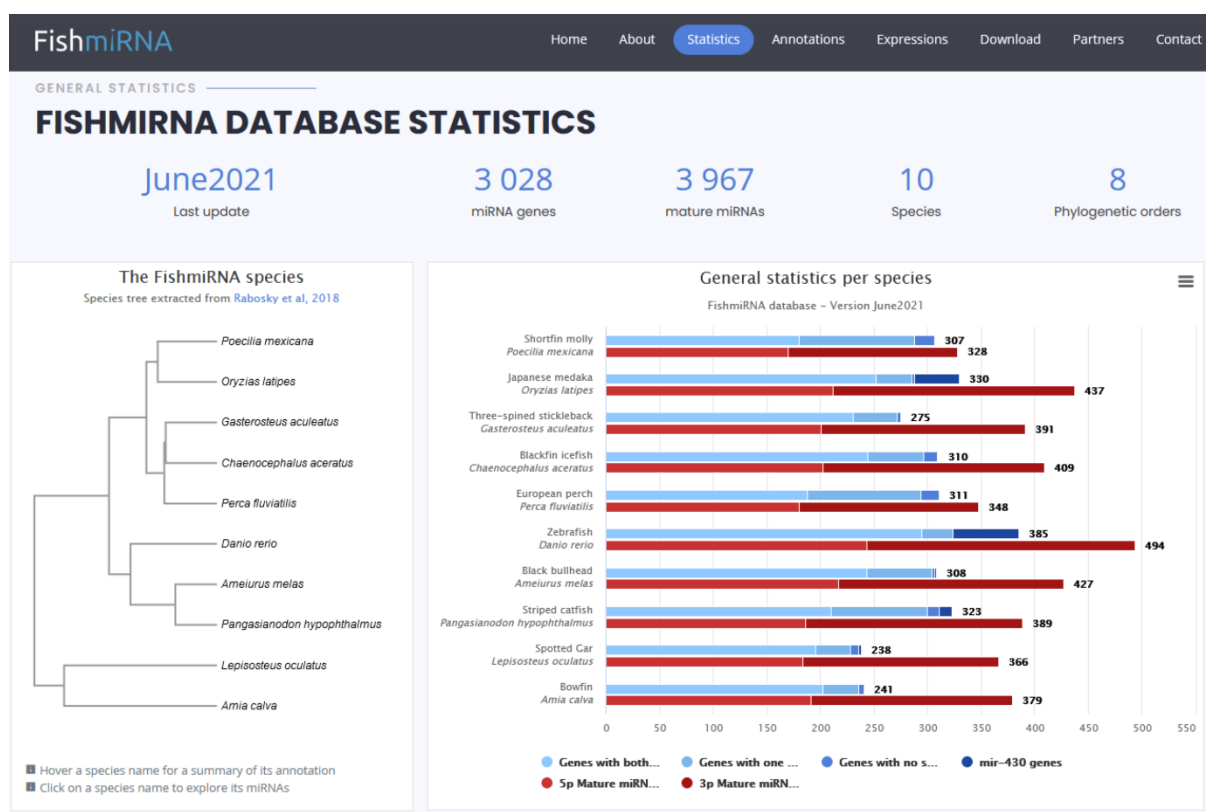


Fig. 2. Exploring the FishmiRNA Annotation database.

Screen shot of the Annotation section of the database. Users can choose to add or remove columns in the display and can filter and search the database by species, miRNA gene and mature miRNA names, and orthology among species. The entire annotation database, or subsets of it, can be exported in various convenient formats. The magnifier icon opens the corresponding gene identification card that contains all the FishmiRNA information related to this gene. Selecting genes using tick boxes on the left allows the filtering of the FishmiRNA expression database for mature products of selected genes.

**EXPLORE FISHMIRNA DATABASE**

**Show/Hide columns**  
Configure the table below by selecting column name(s)  
9 items selected  
Number of visible columns: 9/38.  
All Reset  
'All' columns or 'Reset to default'.  
Remove all filters  
'Remove all filters' in the search box.

**Search**  
Filter by source and/or by position  
Search by BLAST against hairpin and/or mature sequences  
By source By position By BLAST  
Species Nothing selected Species Nothing selected  
Select species common name(s). Select species scientific name(s).  
FishmiRNA  
Filter by FishmiRNA Gene ID.  
Hairpin mir-144 Mature  
Filter by hairpin name. Filter by mature miRNA name.  
Teleost TH-LCA\*  
Filter by orthology among teleosts. Filter by orthology with the TH LCA\*.

**Export**  
For the filtered row(s) export data, sequences or annotations  
Copy Excel CSV  
The data.  
Hairpin  
A FASTA file of the hairpin sequence(s).  
All 5p 3p  
A FASTA file of the mature sequence(s).  
GFF  
A GFF file of the miRNA(s).

Show 10 entries

|                                     | Species common name      | Species scientific name            | Hairpin name | Orthology among teleosts | Orthology with Teleost-Holostei LCA | 5p miRNA name  | 5p miRNA sequence       | 3p miRNA name  | 3p miRNA sequence   |
|-------------------------------------|--------------------------|------------------------------------|--------------|--------------------------|-------------------------------------|----------------|-------------------------|----------------|---------------------|
| <input type="checkbox"/>            | Black bullhead           | <i>Ameiurus melas</i>              | ame-mir-144  | mir144                   | TH-LCA-mir144                       | ame-miR-144-5p | GGATATCATCATATACTGTAAGT | ame-miR-144-3p | CTACAGTATAGATGATGTA |
| <input checked="" type="checkbox"/> | Blackfin icefish         | <i>Chionocephalus aceratus</i>     | cac-mir-144  | mir144                   | TH-LCA-mir144                       | cac-miR-144-5p | AGGATATCATCTTACTGTA     | cac-miR-144-3p | CTACAGTATAGATGATGTA |
| <input type="checkbox"/>            | Bowfin                   | <i>Amia calva</i>                  | aca-mir-144  | mir144                   | TH-LCA-mir144                       | aca-miR-144-5p | GGATATCATCATATACTGTAAGT | aca-miR-144-3p | TACAGTATAGATGATGTA  |
| <input type="checkbox"/>            | European perch           | <i>Perca fluviatilis</i>           | pfl-mir-144  | mir144                   | TH-LCA-mir144                       | pfl-miR-144-5p | GGATATCATCTTACTGTAAGT   | pfl-miR-144-3p | CTACAGTATAGATGATGTA |
| <input checked="" type="checkbox"/> | Japanese medaka          | <i>Oryzias latipes</i>             | ola-mir-144  | mir144                   | TH-LCA-mir144                       | ola-miR-144-5p | GGATATCATCTTACTGTAAGT   | ola-miR-144-3p | CTACAGTATAGATGATGTA |
| <input type="checkbox"/>            | Shortfin Molly           | <i>Poecilia mexicana</i>           | pme-mir-144  | mir144                   | TH-LCA-mir144                       | pme-miR-144-5p | GGATATCATCTTACTGTAAGT   | pme-miR-144-3p | CTACAGTATAGATGATGTA |
| <input checked="" type="checkbox"/> | Spotted Gar              | <i>Lepisosteus oculatus</i>        | loc-mir-144  | mir144                   | TH-LCA-mir144                       | loc-miR-144-5p | GGATATCATCATATACTGTAAGT | loc-miR-144-3p | CTACAGTATAGATGATGTA |
| <input type="checkbox"/>            | Striped catfish          | <i>Pangasianodon hypophthalmus</i> | phy-mir-144  | mir144                   | TH-LCA-mir144                       | phy-miR-144-5p | GGATATCATCATATACTGTAAGT | phy-miR-144-3p | TACAGTATAGATGATGTA  |
| <input type="checkbox"/>            | Three-spined stickleback | <i>Gasterosteus aculeatus</i>      | gac-mir-144  | mir144                   | TH-LCA-mir144                       | gac-miR-144-5p | GGATATCATCTTACTGTAAGT   | gac-miR-144-3p | CTACAGTATAGATGATGTA |
| <input checked="" type="checkbox"/> | Zebrafish                | <i>Danio rerio</i>                 | dre-mir-144  | mir144                   | TH-LCA-mir144                       | dre-miR-144-5p | GGATATCATCGTACTGTAAGT   | dre-miR-144-3p | CTACAGTATAGATGATGTA |

Showing 10 of 10 entries (filtered from 3,028 total entries) 4 rows selected

Select all Deselect all For the selected row(s): Explore expression

Select miRNA genes and explore the expression of their mature products





**FishmiRNA** Home About Statistics Annotations Expressions Download Partners Contact

## EXPLORE EXPRESSION DATA

**Show/Hide columns**  
Configure the table below by selecting column name(s)

12 items selected  
Number of visible columns: 12/23.  
[All] [Reset]  
'All' columns or 'Reset to default'.  
[Remove all filters]  
'Remove all filters' in the search box.

**Search**  
Filter by source  
Search by BLAST against mature sequences

By source: **By BLAST**

Species: Nothing selected  
Select one or more species scientific name.

Mature: ame-miR-202-5p,aca-miR-202-5p,pfl-miR-202-5p,ola-miR-202-5p,  
Filter by miRNA name.

FishmiRNA  
Filter by FishmiRNA Mature ID.

**Export**  
For the filtered row(s) export data, sequences or annotations

[Copy] [Excel] [CSV]  
The data.

Show 10 entries **Display individual miRNA expression** Search:

| miRNA name   | miRNA sequence         | Brain RPM | Gills RPM | Heart RPM | Muscle RPM | Intestine RPM | Liver RPM | Ovary RPM | Testis RPM | Head-kidney RPM | Spleen RPM |
|--|------------------------|-----------|-----------|-----------|------------|---------------|-----------|-----------|------------|-----------------|------------|
| <input checked="" type="checkbox"/> aca-miR-202-5p | TTCTATGCATATACCTCTTT   | 190.0     | 149.5     | 321.3     | 119.7      | 276.2         | 369.7     | 92808.3   | 192186.1   | 310.1           | No_Data    |
| <input type="checkbox"/> ame-miR-202-5p            | TTCTATGCATATACCTCTTT   | 186.2     | 70.9      | 60.7      | 9.5        | 29.4          | 5.6       | 297898.2  | 84492.1    | No_Data         | No_Data    |
| <input checked="" type="checkbox"/> dre-miR-202-5p | TTCTATGCATATACCTCTTTGA | 8.6       | No_Data   | 29.9      | No_Data    | No_Data       | No_Data   | 39266.5   | 26612.8    | No_Data         | No_Data    |
| <input type="checkbox"/> gac-miR-202-5p            | TTCTATGCATATACCTTTTC   | 54.0      | No_Data   | 40.6      | No_Data    | No_Data       | No_Data   | 54368.3   | 33247.3    | No_Data         | No_Data    |
| <input checked="" type="checkbox"/> loc-miR-202-5p | TTCTATGCATATACCTCTTT   | 28.5      | No_Data   | 94.1      | No_Data    | No_Data       | No_Data   | 12020.4   | 64826.5    | No_Data         | No_Data    |
| <input checked="" type="checkbox"/> ola-miR-202-5p | TTCTATGCATATACCTCTTT   | 62.5      | 294.8     | 133.2     | 29.7       | 31.0          | 24.6      | 207245.4  | 215827.7   | 68.0            | No_Data    |
| <input type="checkbox"/> pfl-miR-202-5p            | TTCTATGCATATACCTTTTC   | 20.9      | 8.0       | 6.9       | 1.9        | 13.1          | 2.8       | 33102.9   | 69471.5    | No_Data         | No_Data    |
| <input checked="" type="checkbox"/> phy-miR-202-5p | TTCTATGCATATACCTCTTT   | 23.6      | 10.8      | 5.3       | 8.2        | 13.0          | 41.0      | 119844.6  | 186988.8   | No_Data         | No_Data    |

Showing 1 to 8 of 8 entries (filtered from 3,855 total entries) 5 rows selected

☒ Select all ☐ Deselect all For the selected row(s):  Explore annotation ☒ **Compare expression profile**

Select up to 10 miRNAs to compare their expressions

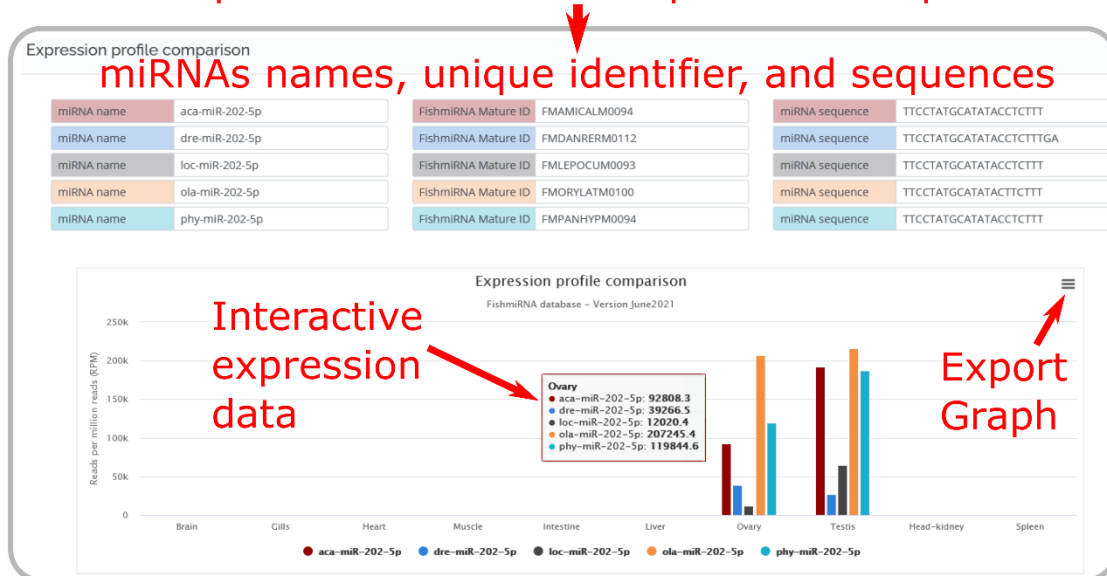


Fig. 4. Downloads and links.

Screenshot of the Download section of the database. Quick links for downloading, per species or for all species, annotations files in FASTA format and raw output files from the smallRNA-seq software *Prost!* used for the annotation and expression analyses. Links to the publication of each original annotation are also provided along with links to the NCBI Project of the expression data analyzed and the article that published these original data.

**FishmiRNA**

Home About Statistics Annotations Expressions **Download** Partners Contact

DOWNLOAD

## DOWNLOAD DATA

If you make use of the data presented in FishmiRNA, please cite the FishmiRNA paper but also the original articles for the specific annotations and/or sequencing data used in your study.

Data and citations are available for FishmiRNA as a whole or by species:

HAIRPIN FASTA: [Download](#)

MATURE FASTA: [Download](#)

PROST! FILES: [Download](#)

ANNOTATION CITATION: [MOLECULAR BIOLOGY AND EVOLUTION](#)

SEQUENCING DATA SRA: [PRJNA477647](#)

EXPRESSION DATA CITATION: [PLOS GENETICS](#)

Expression data in NCBI

Japanese medaka — *Oryzias latipes*

Species-specific citations

miRNA sequences  
Expression data  
analyzed with *Prost!*