FishmiRNA: An Evolutionarily Supported MicroRNA Annotation and Expression Database for Ray-Finned Fishes

Thomas Desvignes (**b**,^{*1} Philippe Bardou (**b**,² Jérôme Montfort,³ Jason Sydes (**b**,¹ Cervin Guyomar (**b**,² Simon George,⁴ John H. Postlethwait (**b**,¹ and Julien Bobe (**b**^{*3}

¹Institute of Neuroscience, University of Oregon, Eugene, OR, USA
²Sigenae, GenPhySE, Université de Toulouse, INRAE, ENVT, Castanet Tolosan, France
³INRAE, LPGP, Rennes, France
⁴MGX—Montpellier GenomiX, University of Montpellier, CNRS, INSERM, Montpellier, France
*Corresponding authors: E-mails: tdesvign@uoregon.edu; julien.bobe@inrae.fr.

Associate editor: Stuart Newfeld

Abstract

MicroRNAs (miRNAs) are important post-transcriptional regulators of gene expression involved in countless biological processes and are widely studied across metazoans. Although 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 miRNA annotation and expression database for ray-finned fishes: 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 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.

Key words: teleost, Holostei, actinopterygian, noncoding RNA, whole-genome duplication, bowfin, Amia calva.

Introduction

MicroRNAs (miRNAs) have emerged as key posttranscriptional 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. Although 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, although miRNA research continues to

© The Author(s) 2022. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (https://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com



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 miRNA 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 data sets 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 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 rayfinned 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 3,028 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 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).

	Table 1. Summary of	f Annotations and	l Sequencing	Data Included	in FishmiRNA.
--	---------------------	-------------------	--------------	---------------	---------------

Species	Genome Assembly	Annotation Citation	Sequencing Data NCBI Accession No.	Expression Data Citation
Spotted gar, Lepisosteus oculatus	LepOcu1 (GCA_000242695.1)	Braasch et al. (2016)	PRJNA296503	Braasch et al. (2016)
Bowfin, Amia calva	AmiCal1 (GCA_017591415.1)	Present study	PRJNA255850	Present study
Black bullhead, Ameiurus melas	AMELA_1.0 (GCA_012411365.1)	Present study	PRJNA730692	Present study
Striped catfish, Pangasianodon hypophthalmus	GENO_Phyp_1.0 (GCF_009078355.1)	Present study	PRJNA256963	Present study
Zebrafish, Danio rerio	GRCz11 (GCA_000002035.4)	Desvignes et al. (2014)	PRJNA240316	Desvignes et al. (2014)
Japanese medaka, Oryzias latipes	ASM223467v1 (GCA 002234675.1)	Desvignes et al. (2021)	PRJNA477647	Gay et al. (2018)
Molly, Poecilia mexicana	P_mexicana-1.0 (GCA_001443325.1)	Kelley et al. (2021)	PRJNA471100	Kelley et al. (2021)
European perch, Perca flavescens	GENO_Pfluv_1.0 (GCA_010015445.1)	Present study	PRJNA256973	Present study
Three-spined stickleback, Gasterosteus aculeatus	BROAD S1	Desvignes et al. (2019)	PRJNA486149	Desvignes et al. (2019)
Blasckfin icefish, Chaenocephalus aceratus	cace 20180227a pilon1	Kim et al. (2019)	PRJNA310135	Desvignes et al. (2016)



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.

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).

	Show/Hi	de columns		Search			/	Expo	ort
J	configure th selecting col	e table bein a by umn name(s)	ス	Hearby searce and search by BLAST aga	or by position inst hairpin and/or r	mature sequences		For the data, s	e filtered row(s) export sequences or annotations
	_	_							
Mur	9 items selec	ted •	Bys	source By posit	ion By BLAST			Copy Exc	cel CSV
-	inder of visible c	olomins, av ale.	Sp	Nothing sele	cted •	Species Nothing	g selected *	The Gata.	
~ 	All 3 Reset	set to default	Sele	ect species common r	iame(s).	Select species scien	tific name(s).	Hairpin	of the bairpin sequence(s)
-	columns or Re	set to delabit.	Fis	shmiRNA					of the numper sequences.
			Filte	er by FishmiRNA Gene	e ID.			All 5p 3	3p
8	Remove all fil	ters	Ha	airpin mir-144		Mature			
ler	move all filters' i	in the search box.	Filte	r by haripin name.	1	Filter by mature mil	RNA name.	GFF A GFF file of t	the miRNA(s).
			Те	leost		TH-LCA*			
					a tolooctr.	Filter by orthology y	with the THEFA"		
	10 v entri Species common	es Species scientific	Filte	Disp	olay th	e gen	e ID Card	Search:	3p miRNA
	10 v entri Species common name 11	Species scientific name	Filte Hairpin name	Orthology among teleosts	orthology with Teleost-Holostei LCA	e gen ^{Sp} miRNA name 11	e ID Card	Search: 3p miRNA name †1	3p miRNA sequence
	10 v entri Species common name †1 Black bullhead	species scientific name 11 Amelurus melas	Filte	Orthology mir144	Orthology with Teleost-Holostei LCA TH-LCA-mir144	e gen sp miRNA name 11 arme- miR-144-5p	e ID Card sp miRNA sequence	Search: 3p miRNA name miR-144-3p	3p miRNA sequence CTACAGTATAGATGATGTA
	10 ventri Species common name 11 Black bullhead Blackfin icefish	es Species scientific name Amelurus melas Chaenocephalus aceratus	Hairpin name ame-mir-144 (2) cac-mir-144 (2)	Orthology amon Orthology among teleosts mir144 mir144	Orthology with Teleost-Holostei LCA TH-LCA-mir144 TH-LCA-mir144	e gen sp miRNA name miR-114-5p cac- miR-144-5p	e ID Card	Search: 3p miRNA name miR-144-3p Cac- miR-144-3p	3p miRNA sequence CTACAGTATAGATGATGTA CTACAGTATAGATGATGTA
	10 v entri species common name 11 Black bullhead Blackfin licefish Bowfin	Species scientific name 1 Ameiurus melas Chaencephalus aceratus Amia calva	Hairpin name ame-mir-144 (2) cac-mir-144 (2) aca-mir-144 (2)	Orthology amon Orthology among teleosts m mir144 mir144 mir144	Orthology with Teleast-Holostei LCA TH-LCA-mir144 TH-LCA-mir144 TH-LCA-mir144	e gen sp miRNA name miR-144-5p aca- miR-144-5p aca- miR-144-5p	e ID Card Sp miRNA sequence GGATATCATCATATACTGTAAGT AGGATATCATCATATACTGTAAGT	Search: 3p miRNA name miR-144-3p cac- miR-144-3p aca- miR-144-3p	3p miRNA sequence CTACAGTATAGATGATGTA CTACAGTATAGATGATGTA TACAGTATAGATGATGATGTAC
	10 Y entri Species common name 1: Black bullhead Blackfin icefish Bowfin European perch	es Species scientific name Ameiurus melas Ameiurus melas Amai calva Perca fluviatilis	Filte Hairpin name ame-mir-144 Q cac-mir-144 Q pfl-mir-144 Q	Orthology amon Orthology among teleosts 11 mir144 mir144 mir144	Orthology with Teleost-Holostel LCA TH+LCA-mir144 TH+LCA-mir144 TH+LCA-mir144 TH+LCA-mir144	E GEN Sp miRNA name miR.144-5p cac- miR.144-5p aca- miR.144-5p pfl- miR.144-5p	e ID Card Sp miRNA sequence	Search: 3p miRNA name 1 ame. miR-144-3p aca- miR-144-3p miR-144-3p	3p miRNA sequence CTACAGTATAGATGATGTA CTACAGTATAGATGATGTA TACAGTATAGATGATGTAC CTACAGTATAGATGATGATGTA
	10 v entri Species common name 1 Black bullhead Blackin Ledish Blowlin European perch Bapanese medaka	Species scientific name Ameiurus melas Chaencephalus aceratus Amia calva Perco fluviatilis Oryzias latipes	File Hairpin name ame-mir-144 (2) ccac-mir-144 (2) pfl-mir-144 (2) ola-mir-144 (2)	Orthology amon Orthology among teleosts 11 mir144 mir144 mir144 mir144 mir144	Orthology with Teleost-Holostei LCA TH-LCA-mir144 TH-LCA-mir144 TH-LCA-mir144 TH-LCA-mir144 TH-LCA-mir144	E Gen miRNA name miR-144-5p aca- miR-144-5p pfl- miR-144-5p ola- miR-144-5p	e ID Card Sp miRNA sequence GGATATCATCATATACTGTAAGT AGGATATCATCATATACTGTAAGT GGATATCATCTTATACTGTAAGT GGATATCATCTTATACTGTAAGT	Search: Sp miRNA name miR-144-3p aca- miR-144-3p pfl. miR-144-3p ola- miR-144-3p	3p miRNA sequence CTACAGTATAGATGATGAT CTACAGTATAGATGATGATGAT TACAGTATAGATGATGATGATGAT CTACAGTATAGATGATGATGATGATGATGATGATGATGATGATG
	10 v entri Species common Black bullhead Blacklin Blacklin bullhead Blacklin Blackli	es Species scientific name Ameiurus melas Chaenocepholus acceratus Amia calva Perca fluviatilis Oryzias latipes Poecilia mexicona	Filtz	Orthology amon Orthology among teleosts 12 mir144 mir144 mir144 mir144 mir144 mir144	Orthology with Teleost-Holostel LCA TH-LCA-mir144 TH-LCA-mir144 TH-LCA-mir144 TH-LCA-mir144 TH-LCA-mir144 TH-LCA-mir144	E Gen miRNA name arme- miR:144-5p cac- miR:144-5p pfl- miR:144-5p pfl- miR:144-5p pdi- miR:144-5p	e ID Card Sp miRNA sequence 11 GGATATCATCATATACTGTAAGT AGGATATCATCATATACTGTAAGT GGATATCATCATATACTGTAAGT GGATATCATCATATACTGTAAGT GGATATCATCTTATACTGTAAGT GGATATCATCTTATACTGTAAGT	Search: 3p miRNA name 1 ame- miR-144-3p (ac- miR-144-3p (br- miR-144-3p (br- miR-144-3p (br- miR-144-3p (br- miR-144-3p (br-))))))))))))))))))))))))))))))))))))	3p miRNA sequence CTACAGTATAGATGATGTAT CTACAGTATAGATGATGTAT TACAGTATAGATGATGTAC CTACAGTATAGATGATGTATGTA CTACAGTATAGATGATGTATGTA
	10 v entri Species common 1 mame 1 Black bullhead Blackan bullhead Blackin Blackin Blackin bullhead Blackin Black bullhead Shortin Spotted Gar	Species scientific name a Ameiurus metas Chaencephalus aceratus Amia calva Perca fluviatilis Oryzias latipes Poecila mexicana Lepisosteus	Filte Hairpin name ame-mir-144 (3) cac-mir-144 (3) pfl-mir-144 (3) pla-mir-144 (3) pme-mir-144 (3) loc-mir-144 (3)	Orthology amon Orthology amon teleosts 11 mir144 mir144 mir144 mir144 mir144 mir144 mir144	Orthology with LCA TH-LCA-mir144	E Generation of the second sec	E ID Card Spmirna sequence GGATATCATCATATACTGTAAGT AGGATATCATCATATACTGTAAGT GGATATCATCTTATACTGTAAGT GGATATCATCTTATACTGTAAGT GGATATCATCTTATACTGTAAGT GGATATCATCTTATACTGTAAG	Search: 3p mIRNA name mIR-144-3p cac- mIR-144-3p pfl- mIR-144-3p pme- mIR-144-3p loc- mIR-144-3p	3p miRNA sequence CTACAGTATAGATGATGAT CTACAGTATAGATGATGATGAT TACAGTATAGATGATGATGATGAT CTACAGTATAGATGATGATGAT CTACAGTATAGATGATGATGTA CTACAGTATAGATGATGATGATGA
	10 v entri Species common 1 Black bullhead Blackinu Blackinu bullhead Blackinu Blach	es Species scientific name Ameiurus melas Chaenocephalus acceratus Amia calva Perca fluviatilis Oryzias latipes Poecilia mexicana Lepisosteus culatus Pangasiandon hypophthalmus	Filtz	Orthology amon Orthology among teleosts 11 mir144 mir144 mir144 mir144 mir144 mir144 mir144 mir144	Orthology with Teleost-Holostel LCA 7H-LCA-mir144 7H-LCA-mir144 7H-LCA-mir144 7H-LCA-mir144 7H-LCA-mir144 7H-LCA-mir144 7H-LCA-mir144 7H-LCA-mir144	E Geometry i space of the second sec	E ID Card SpmiRNA sequence	Search:	3p miRNA sequence CTACAGTATAGATGATGTAU CTACAGTATAGATGATGTAU TACAGTATAGATGATGTAU CTACAGTATAGATGATGTAU CTACAGTATAGATGATGTAU CTACAGTATAGATGATGTAU TACAGTATAGATGATGTAU
	10 v entri Species common 1 11 Black bullhead Blackin 2 Black bullhead Blackin 2 Black bullhead Black Black bullhead Black	Species Scientific name Ameiurus melas Chaencephalus aceratus Amia calva Perca fluvtatils Oryzias latipes Poglia Pagsianadan Pypophthalmus Gasterostus Gasterostus	Hairpin name ame-mir-144 (a) cac-mir-144 (a) pfl-mir-144 (a) pla-mir-144 (a) pme-mir-144 (a) phy-mir-144 (a) gac-mir-144 (a)	Orthology amon Orthology amon teleosis 11 mir144 mir144 mir144 mir144 mir144 mir144 mir144 mir144 mir144 mir144	Orthology with releast. Holostei LCA TH-LCA-mir144 TH-LCA-mir144	E Genner spinner Spinner miRNA name miR144-5p cac miR-144-5p cac miR-144-5p pfl- miR-144-5p	e ID Card SpmiRNA sequence GGATATCATCATATACTGTAAGT AGGATATCATCATATACTGTAAGT GGATATCATCTTATACTGTAAGT GGATATCATCTTATACTGTAAGT GGATATCATCTTATACTGTAAGT GGATATCATCTTATACTGTAAGT GGATATCATCTTATACTGTAAGT GGATATCATCTTATACTGTAAGT	Search: 3p mRNA name mR-144.3p cac- mR-144.3p pfl- mR-144.3p ola- mR-144.3p pme- mR-144.3p pome- mR-144.3p phy- mR-144.3p phy- mR-144.3p mR-144.3p	3p miRNA sequence CTACAGTATAGATGATGTAG CTACAGTATAGATGATGTAG TACAGTATAGATGATGTAG CTACAGTATAGATGATGTAG CTACAGTATAGATGATGTAG CTACAGTATAGATGATGTAG TACAGTATAGATGATGTAG TACAGTATAGATGATGTAG

Select miRNA genes and explore the expression of their mature products

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.

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 ten columns that contain the raw counts and ten 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



Fig. 3. Exploring the FishmiRNA Expression database. Screenshot of the Expression section of the database. Like the Annotation section, users can choose to add or remove columns in the display and can filter and search the database by species, mature miRNA name, and unique identifier. The entire expression database, or subsets of it, can also be exported in convenient table formats. The histogram icon displays the corresponding mature miRNA expression pattern in a variety of organs. Users can also select up to any 10 mature miRNAs using tick boxes on the left to display their expression patterns on the same graph.

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 ten



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.

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, reanalyzed 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 described previously (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, 2020; Ruzicka et al. 2019). FishmiRNA was developed based on the RumimiR web interface (Bourdon et al. 2019).

Acknowledgments

The authors 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 Nos. NIH R24 OD011199, NIH 5R01 OD011116, and NIH R01 GM085318 to J.H.P.), the National Science Foundation Office of Polar Program (NSF OPP-1543383 to J.H.P. and T.D.), and the Agence Nationale de la Recherche (ANR-18-CE20-0004 to J.B.). 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).

Author Contributions

Study concept and design: T.D., J.H.P., and J.B. Acquisition of data: T.D., S.G., and J.B. Analysis and interpretation of data: T.D., J.M., J.S., C.G., J.H.P., and J.B. Web interface development: P.B. Wrote the manuscript: T.D. and J.B. Critical revision of the manuscript: T.D., P.B., C.G., J.H.P., and J.B. Obtained funding: J.H.P., T.D., P.B., and J.B. Study supervision: T.D., J.H.P., and J.B.

Data Availability

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

References

- Ahkin Chin Tai JK, Freeman JL. 2020. Zebrafish as an integrative vertebrate model to identify miRNA mechanisms regulating toxicity. *Toxicol Rep.* 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(5394):1711–1714.
- Andreassen R, Høyheim B. 2017. miRNAs associated with immune response in teleost fish. *Dev Comp Immunol*. 75:77–85.

Bartel DP. 2018. Metazoan microRNAs. Cell 173(1):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* 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 humanteleost comparisons. *Nat Genet.* 48(4):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 Biol.* 19(1):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 Genet.* 31(11):613–626.
- Desvignes T, Batzel P, Sydes J, Eames BF, Postlethwait JH. 2019. miRNA analysis with Prost! reveals evolutionary conservation of organenriched expression and post-transcriptional modifications in three-spined stickleback and zebrafish. *Sci Rep.* 9(1):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(2):386–389.
- Desvignes T, Detrich HW, III, Postlethwait JH. 2016. Genomic conservation of erythropoietic microRNAs (erythromiRs) in white-blooded Antarctic icefish. *Mar 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(3):698–703.
- Desvignes T, Sydes J, Montfort J, Bobe J, Postlethwait JH. 2021. Evolution after whole-genome duplication: teleost microRNAs. *Mol Biol Evol.* 38(8):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(11):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(D1):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(9):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(5770):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. *Aquat Toxicol.* 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. 9:88.
- 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. *Dis Model Mech.* 10(11):1323–1332.
- Jonas S, Izaurralde E. 2015. Towards a molecular understanding of microRNA-mediated gene silencing. Nat Rev Genet. 16(7):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. *Dev Cell*. 40(6):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. J Evol Biol. 34(6):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. *Nat Ecol Evol*. 3(3):469–478.
- Kozomara A, Birgaoanu M, Griffiths-Jones S. 2019. miRBase: from microRNA sequences to function. Nucleic Acids Res. 47(D1):D155–D162.
- Li J, Zhang Z. 2013. MiRNA regulatory variation in human evolution. Trends Genet. 29(2):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(10):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(6):1172–1187.
- Postlethwait JH. 2007. The zebrafish genome in context: ohnologs gone missing. J Exp Zool B Mol Dev Evol. 308(5):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(5):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(7714):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(D1):D867–D873.
- Schmiedel JM, Klemm SL, Zheng Y, Sahay A, Blüthgen N, Marks DS, Oudenaarden A. V. 2015. MicroRNA control of protein expression noise. *Science* 348(6230):128–132.
- Sun K, Lai EC. 2013. Adult-specific functions of animal microRNAs. *Nat Rev Genet.* 14(8):535–548.
- Taylor JS, Braasch I, Frickey T, Meyer A, Peer Y. D. 2003. Genome duplication, a trait shared by 22,000 species of ray-finned fish. *Genome Res.* 13(3):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.* 53:1373–1384.
- 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.* 9:2801.
- Xiong P, Schneider RF, Hulsey CD, Meyer A, Franchini P. 2019. Conservation and novelty in the microRNA genomic landscape of hyperdiverse cichlid fishes. *Sci Rep.* 9(1):13848.