

# De novo assemblies, annotations, and gene expression profiling of gill epithelium from 16 species of *Fundulus* killifish in response to salinity change.



Lisa K. Johnson<sup>1,2,3\*</sup> and Andrew Whitehead<sup>3</sup>  
<sup>1</sup>School of Veterinary Medicine, <sup>2</sup>Molecular, Cellular, Integrative Physiology Graduate Group, <sup>3</sup>Department of Environmental Toxicology, University of California Davis  
 \* Email: ljcohen@ucdavis.edu @monsterbashseq



## ABSTRACT:

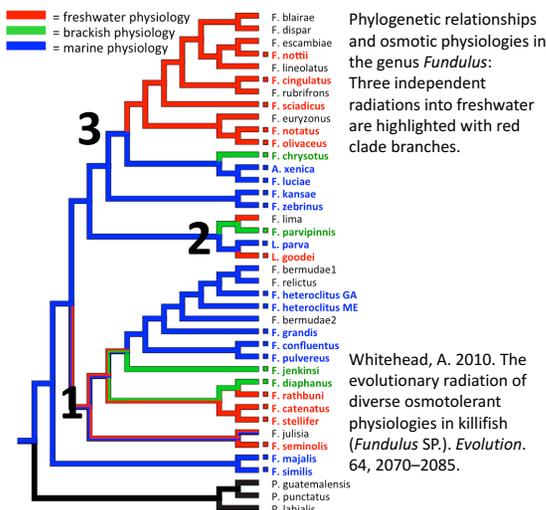
Estuaries are characterized by periodic fluxes in salinity. Many species of North American killifish (*Fundulus*) are estuarine specialists and harbor euryhaline phenotypes. Three clades within *Fundulus* independently radiated into freshwater environments and have lost their abilities to tolerate high salinity. We use *Fundulus* as a comparative model system for studying the physiological and genetic mechanisms that diverge between euryhaline and freshwater species. We examined 16 estuarine and freshwater species with representation from each of three clades. Fish from all species were acclimated to either brackish or fresh water then exposed to an acute brackish water challenge. Gill transcriptome data were collected. To enable multi-species comparisons, reference transcriptome assemblies were generated *de novo* for each species then used to analyze transcriptional responses to salinity change by clade and physiology. We find differences in the gene expression between euryhaline and freshwater species, some of which are shared across clades, implicating molecular mechanisms that contribute to divergent osmoregulatory physiologies.

## OBJECTIVES:

To address comparative physiology questions across species using transcriptomics, specifically:

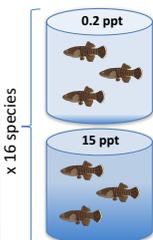
*Is there evidence of parallel or divergent osmoregulatory evolution in this system?*

1. Build reference transcriptomes and reproducible infrastructure for analyzing RNAseq data across multiple species
2. What are the “apples-to-apples” orthologous comparisons?
3. Test hypotheses for gene expression patterns across clades



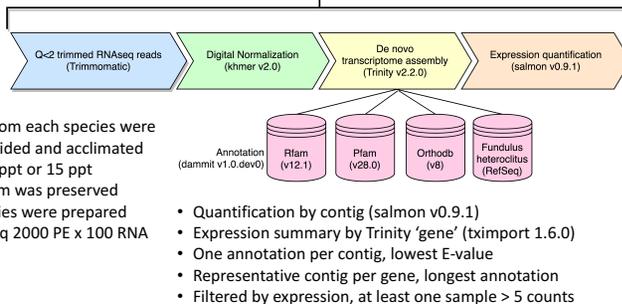
## METHODS:

### Osmotic Challenge Experiment



- Individuals from each species were randomly divided and acclimated to either 0.2 ppt or 15 ppt
- Gill epithelium was preserved
- polyA+ libraries were prepared
- Illumina HiSeq 2000 PE x 100 RNA sequencing

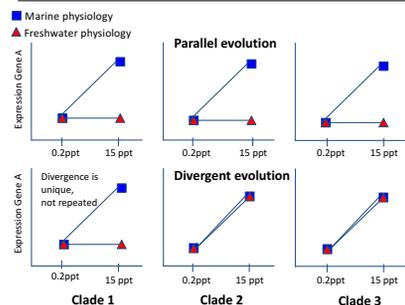
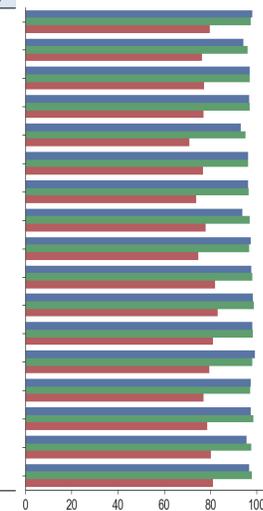
### Data Analysis x 16 species



## TRANSCRIPTOME ASSEMBLIES:

Species	Native Physiology BW = Brackish Water FW = Freshwater M = Marine	Clade	N	Quality Trimmed (Q>2) Reads	% Kept After Diginorm	Trinity contigs	Annotated contigs	Unique gene Names	<i>Fundulus heteroclitus</i> (RefSeq) gene names
<i>Adenia xenica</i>	M	3	9	250,627,759	24.8	362,783	148,980	50,123	21,806
<i>Fundulus catanatus</i>	FW	1	7	328,807,408	21.6	405,866	164,596	47,298	22,435
<i>Fundulus chrysotus</i>	BW	3	8	258,850,289	25.6	396,400	155,206	46,327	22,455
<i>Fundulus diaphanus</i>	BW	1	7	137,246,213	34	384,218	128,323	42,877	21,624
<i>Fundulus grandis</i>	M	1	9	467,432,867	23.2	809,060	234,074	63,741	24,275
<i>Fundulus heteroclitus 1</i>	M	1	9	319,925,008	24.8	592,419	188,268	54,653	22,849
<i>Fundulus heteroclitus 2</i>	M	1	9	275,951,932	28.2	668,487	186,798	54,253	22,994
<i>Fundulus notatus</i>	FW	3	9	349,630,701	21.6	416,299	167,061	46,884	22,530
<i>Fundulus nottii</i>	FW	3	2	46,463,472	47	159,771	69,247	30,427	18,280
<i>Fundulus olivaceus</i>	FW	3	8	202,133,952	27.8	350,265	134,207	42,194	21,600
<i>Fundulus parvapis</i>	BW	2	8	184,254,591	27.5	352,346	126,200	46,368	20,647
<i>Fundulus rathbuni</i>	FW	1	9	348,759,075	22.7	501,222	176,367	48,801	22,718
<i>Fundulus sciadicus</i>	FW	3	5	101,937,160	37.2	241,279	98,332	36,027	20,382
<i>Fundulus similis</i>	M	1	9	207,444,577	30.5	520,319	154,675	48,509	22,338
<i>Fundulus zebrinus</i>	M	3	4	98,327,251	36.6	266,978	102,046	37,405	20,439
<i>Lucania goodei</i>	FW	2	9	219,175,639	28	385,476	138,485	46,599	22,631
<i>Lucania parva</i>	M	2	9	255,219,214	25	409,543	145,529	46,914	22,601

BUSCO Lineage Database v3.0 used to evaluate the completeness of transcriptomes, indicated similar results across assemblies.



## SUMMARY:

1. Reference *de novo* transcriptome assemblies and annotations for gill from 16 species of *Fundulus* killifish were generated and are available.
2. 53K genes x 16 species expression quantification table will be used for analysis of patterns across clades for the osmotic challenge experiment.
3. Reproducible, automated scripts are available to generate *de novo* transcriptome assemblies, annotations, and merge gene expression tables across species.

## DATA:

- *Fundulus* Multispecies Osmotic Transcriptome Sequencing Project (FMOTSP): <http://doi.org/10.17605/OSF.IO/M4XEG>
- Analysis scripts: [https://github.com/ljcohen/RNAseq\\_15killifish](https://github.com/ljcohen/RNAseq_15killifish)
- Raw data: <https://www.ncbi.nlm.nih.gov/bioproject/473009>

## ACKNOWLEDGEMENTS:

Reid Brennan performed the osmotic challenge experiment. Jen Roach prepared the RNAseq libraries. Thank you to Noah Reid and Titus Brown for guidance and assistance with data analysis. Thanks to the DIB lab and Whitehead lab members at UC Davis for helpful discussions.