

## Proposal Status

Status As of Today Dated: **08/13/09**

This proposal has been declined by NSF.

### **Comments from the cognizant Program Officer:**

Proposal summary: Funding was requested to examine pollution tolerance in the killifish *Fundulus heteroclitus*; populations differ by orders of magnitude in their tolerance. The majority of the proposal addresses sequencing and annotating the killifish genome, which is necessary in order to identify genes associated with pollution tolerance. With the sequence available, a variety of methods will be applied to identify genes associated with tolerance, evaluate whether the same genes are used for tolerance in different populations, and to evaluate gene expression.

The Evolutionary Genetics Panel considered this proposal in terms of the intellectual merit and the broader impacts of the proposed research.

Results of prior support:

Intellectual merit. Both PIs have been very productive with previous support.

Broader impacts. As a result of work conducted by PI Shaw and others, *Daphnia* is being accepted as a new model species. No contributions to human resources or outreach were described.

Intellectual merit: The trait of pollution tolerance is clearly adaptive, and thus this system should be a good one for investigating the genetic basis of adaptive traits. The questions raised by the PIs are important and entirely appropriate. The weakness of this proposal is that very little of it is devoted to addressing the questions that were provided as the framework. Instead, the proposal provides substantial detail about sequencing and annotating the genome; this part of the proposal is very convincing. The big problem is that it deals with sequencing, not hypothesis testing.

Broader impacts: The broader impacts of this project are primarily the production of an annotated genome, which will serve a substantial community. A workshop during the final stage of genome annotation will provide opportunities for students and postdocs to learn how to do annotation. No other broader impacts were stated, and the budgets do not provide support for postdocs, graduate students, or undergraduates. While the scientific benefit of the work is likely to be meaningful, the PIs need to consider their research in the broader contexts of training and outreach.

Panel recommendation:

After study of all reviews and discussion during the panel meeting, the panel rated this proposal as "Superior"

Program recommendation: However, despite this high rating by the panel, funding was not possible because the proposed work is largely related to sequencing a genome rather than hypothesis testing.

**Proposal Number:** 0919183

**Panel Summary:**

Panel Summary

Population and Evolutionary Processes Cluster  
Spring 2009 Advisory Panel

Results of Prior NSF Support (If Applicable):

Results of prior support Criterion I (Intellectual Merit):

Whitehead holds a recently funded NSF grant, and lists four manuscripts published during the first year. Shaw was co-PI on a grant that ended 9/08; fifteen manuscripts resulted from this collaborative effort.

Results of prior support Criterion II (Broader Impacts):

Shaw's work with *Daphnia* contributed to the acceptance of this animal as a new model species. A special issue of *Science*, that includes a manuscript describing metal tolerance in *Daphnia*, will be launched in 2009. Shaw helped establish the *Daphnia* Genomics Consortium and associated web resources.

If this a resubmission, how have previous criticisms been addressed?

N/A

CRITERION I: [Intellectual Merit]:

Intellectual Strengths:

The PIs propose to study adaptive phenotypic change in the marine teleost *Fundulus heteroclitus*. This area of research has implications for understanding fundamental mechanisms of molecular evolutionary change (e.g., parallel evolution at the molecular level). The attempt to address the question of whether selection predominately acts on standing genetic variation or de novo mutants in natural populations was considered a particularly novel aspect of the research.

Intellectual Weaknesses:

Several reviewers were disappointed with the lack of detail for individual experiments. Much time was devoted to the acquisition of the annotated genome, but comparatively little was devoted to hypothesis testing.

## CRITERION II: [Broader Impacts]:

### Broader Impact Strengths:

The acquisition of a robust annotated genome benefits a large community of researchers across a diverse array of scientific disciplines. The PIs will host a genome annotation workshop and a collaboration wiki. These vehicles would be important resources for the genomics community and serve as a training device for students and postdocs.

### Broader Impact Weaknesses:

The panel felt that the sequenced genome would aid all members of the Fundulus Genomics Consortium. However, how this resource would aid student training was not well developed.

## SYNTHESIS AND RECOMMENDATION

The PIs propose to assemble a robust annotated genome for *Fundulus heteroclitus*, a teleost fish that serves as a model for addressing questions of adaptive phenotypic change. Acquisition of this genomic resource would aid a large community of scientists comprising the Fundulus Genome Consortium. However, the panel felt the 'Testing Hypotheses' sections of the proposal lacked sufficient detail.

The panel recommendation is: Superior

This summary was read by the assigned panelists and they concurred that the summary accurately reflects the panel discussion.

**Panel Recommendation:** Superior

## Review #1

---

**Proposal Number:** 0919183  
**Performing Organization:** La St U  
**NSF Program:** Evolutionary Processes Cluster  
**Principal Investigator:** Whitehead, Andrew  
**Proposal Title:** Collaborative Research: Genomic basis of dramatic, rapid, convergent evolution in the killifish *Fundulus heteroclitus*  
**Rating:** Excellent

---

## REVIEW:

What is the intellectual merit of the proposed activity?

In this proposal, the authors propose to sequence the genome of the killifish (*Fundulus heteroclitus*) and use the resulting genomic resources to help understand the dramatic evolution of resistance to toxic environments, which has occurred independently in multiple populations. I thought that this was an exciting and ambitious proposal, using new genomic-scale data to address fundamental questions in evolution. The authors make a compelling case that they will be able to successfully carry out the research. I would have appreciated additional details about the testing of some of the evolutionary hypotheses, but I nevertheless found the proposal to be generally convincing.

What are the broader impacts of the proposed activity?

The authors make a compelling case that the genomics tools generated here will be useful to dozens of other researchers who use this species as a model system. On the negative side, they say very little about other Broader Impacts, such as educational components.

## Review #2

---

**Proposal Number:** 0919183  
**Performing Organization:** La St U  
**NSF Program:** Evolutionary Processes Cluster  
**Principal Investigator:** Whitehead, Andrew  
**Proposal Title:** Collaborative Research: Genomic basis of dramatic, rapid, convergent evolution in the killifish *Fundulus heteroclitus*  
**Rating:** Excellent

---

### REVIEW:

What is the intellectual merit of the proposed activity?

Very high

What are the broader impacts of the proposed activity?

Very important, will affect many fields (evolutionary ecology, applied pollution biology)

Summary Statement

The proposed research would use the latest pyrosequencing technology to produce a full genome sequence for *Fundulus heteroclitus*, and then annotate this genome.

I found this to be a very well written and well defended proposal. One of the positive aspects of this proposal is that it clearly shows that it is now possible to produce a fully annotated, complete sequence of a vertebrate within a very reasonable budget of about 500K. This seems well worth it given the benefits outlined in this proposal.

And these benefits are laid out well. *Fundulus heteroclitus* has been at the forefront of several types of research. This species was one of the first for which it was shown that differences between populations in expression differences could control adaptive differences. Also, as explained in this proposal, this fish is an important model for examining adaptation and tolerance to extreme pollutants, such as dioxin and PCBs. *Fundulus* researchers have begun addressing important and classic questions in evolutionary biology, e.g., is parallel adaptation due to the same or different genes, how often is regulator genes versus coding genes involved in adaptation, and the role of standing variation versus new mutations in allowing adaption. All of these projects will be helped by a full sequence, with high-level annotation.

The authors of the proposal also did a good job of outlining the resources that are already available to support and make the full genome more valuable, including EST libraries, cDNA and oligonucleotide microarrays, BAC libraries, mapping markers, and knockout technologies.

The team that has been assembled for this project has an amazing breadth and depth of knowledge in genomics, from the actual sequencing to the annotation. The method of combining several of the latest types of sequencing seems guaranteed to produce a good final sequence. The proposal outlines a community approach to annotation. Key to this is the 'genome annotation steering committee,' which will adopt annotation guidelines and ensure common reporting standards. I think it would have been helpful to discuss in greater detail how this committee is formed, how it will be ensured that the *Fundulus* community will 'sign on' to this committee's authority, and how any possible conflicts within the committee will be resolved.

This is a very important model organism for tolerance research, related to pollution and other applied questions, and for ecological and evolutionary understanding of adaptation. The utility of this organism for such research will be greatly aided by a full genome sequence, with good annotation. I hope NSF will rank this type of whole genome sequencing as a high priority for funding.

### Review #3

---

**Proposal Number:** 0919183  
**Performing Organization:** La St U  
**NSF Program:** Evolutionary Processes Cluster  
**Principal Investigator:** Whitehead, Andrew  
**Proposal Title:** Collaborative Research: Genomic basis of dramatic, rapid, convergent evolution in the killifish *Fundulus heteroclitus*  
**Rating:** Excellent

---

### REVIEW:

What is the intellectual merit of the proposed activity?

Sequencing the *Fundulus heteroclitus* genome would bring significant advances to a broad community of scientists with research programs in ecological, evolutionary, physiological, and environmental science. Whitehead makes an excellent case for the value of a *Fundulus* genome sequence. The bottom line is that a genome sequence is imperative to nearly all biologists working on an organism.

This collaborative proposal has an excellent team. Wes Warren and the WUGSC have extensive experience in genome sequencing. Joseph Shaw has extensive experience in genome annotation. Andrew Whitehead is a young, enthusiastic PI fully capable of organizing this effort.

The plan for genome sequencing is 10x of long, single 454 reads, 1x of shorter, paired 454 reads, and 6x of very short paired Illumina reads. The strategy makes sense and was developed by experts in genome sequencing. The de novo assembly of a chicken genome suggests that the strategy should also work for *Fundulus*. However, it was not discussed why single long reads of ~400bp would be better than shorter paired reads of ~200bp each. The long reads are proposed to attain sufficient coverage yet two short paired reads should provide the same level of coverage. Paired reads are critical of assembly contiguity. Given that the two types of reads have the same costs, the case for single versus paired 454 reads was not well supported.

One area that wasn't addressed is how data will be made accessible to the community concurrent with its production. Even raw traces can be a valuable resource before an assembly is generated. Once an assembly is generated it should be made public and

distributed to the community before annotation since annotation can take considerable time. A conference is proposed at the final stages of the annotation. However, the data, in as many forms as possible, should immediately be made public so that it is not accessible to some individuals but not others in the community.

What are the broader impacts of the proposed activity?

A workshop will be conducted to train the community to use bioinformatic tools for genome annotation. Travel awards will be made for both students and postdocs. No plans are given for how the workshop should broaden the participation of underrepresented groups. However, the workshop and the overall project should significantly enhance the infrastructure for research and education. The broader impact of the project is likely to be highly significant.

#### Summary Statement

A *Fundulus* genome sequence would be of enormous value to the scientific community and will without doubt have a broad impact. The research plan is well developed and likely to succeed. However, there were two concerns. No plan was given for the dissemination of the data concurrent to its production. Making the data public only after publication would greatly limit its near-term value to others in the community. Another concern is that the authors should consider substituting paired end reads rather than single 454 reads. However, addressing these concerns in another proposal would not be useful as it would delay the availability of a *Fundulus* genome sequence to a large community of scientists. The proposal should be immediately funded and these concerns should be taken into consideration. The bottom line is that there is no reason why a model system as well developed as *Fundulus* should not have a genome sequence.



<b>Proposal Number:</b>	0919183
<b>Performing Organization:</b>	La St U
<b>NSF Program:</b>	Evolutionary Processes Cluster
<b>Principal Investigator:</b>	Whitehead, Andrew
<b>Proposal Title:</b>	Collaborative Research: Genomic basis of dramatic, rapid, convergent evolution in the killifish <i>Fundulus heteroclitus</i>
<b>Rating:</b>	Good

---

## REVIEW:

What is the intellectual merit of the proposed activity?

I think the PIs are (mostly) asking for funds to generate a genome sequence in *Fundulus heteroclitus*, although this is never stated explicitly. I inferred it from the first few pages of text. However, that is not a Specific Aim of the proposal nor their overarching goal, which I find a bit bizarre. I am pretty sure the PIs want to sequence the genome, then map QTLs (although they request no funds for this), and finally look at gene expression.

It is not clear to me why they will sequence a single inbred individual instead of an outbred individual. The PIs cite a half-dozen papers where genome sequences have been derived from heterozygous organisms, and the advantages to doing so (e.g., preliminary SNP characterization) seem substantial in the face of relatively little "cost". I don't see the justification for using an inbred fish.

What are the broader impacts of the proposed activity?

Relatively weak; most listed are narrow impacts on the scientific community that is already engaged in this type of research.

Summary Statement

I think this proposal would be improved if the PIs made their goals, and their respective hierarchy, more clear.

## Review #5

---

**Proposal Number:** 0919183  
**Performing Organization:** La St U  
**NSF Program:** Evolutionary Processes Cluster  
**Principal Investigator:** Whitehead, Andrew  
**Proposal Title:** Collaborative Research: Genomic basis of dramatic, rapid, convergent evolution in the killifish *Fundulus heteroclitus*  
**Rating:** Very Good

---

## REVIEW:

What is the intellectual merit of the proposed activity?

What are the broader impacts of the proposed activity?

### Summary Statement

Whitehead and Shaw

Collaborative Research: Genomic basis of dramatic, rapid, convergent evolution in the killifish *Fundulus heteroclitus*

Whitehead and Shaw propose to study pollution tolerance in the coastal killifish, *Fundulus heteroclitus*. Killifish populations are reasonably distinct genetically throughout their range, and disjunct populations display quite remarkable differences in tolerance to environmental stress. The PIs propose to extend previous research by elucidating the genomic basis of adaptive phenotypes by anchoring QTLs, genomic scans and transcriptome profiling to a complete genome sequence (the focus of the present proposal). A complete genome sequence would lead to:

1. Identification of the number and physical location of loci associated with the tolerant phenotype.
2. A test of the hypothesis that populations have evolved tolerance convergently.
3. A test of the association between mapped loci and diverged patterns of gene expression.

The PIs present a compelling argument for sequencing the entire *Fundulus* genome, a non-model yet ecologically and evolutionarily important marine animal. The idea that

selection might act on pre-existing versus do novo diversity in disjunct populations is a fascinating twist to the proposed research. Similarly, the cadre of scientists using *Fundulus* as a model is impressive and further supports the contention that a whole genome sequence would be a valuable advance.

However, there are some issues that need to be addressed. First, testing hypotheses of convergent evolution across disparate populations within a species is secondary relative to the time spent discussing the specifics of genome sequencing — hypothesis testing should be the primary focus of this proposal. Second, the PIs should address if the sequenced teleost genomes available (i.e., puffer and medaka in particular) are useful for the proposed research or if these genomes are from animals too distantly related to killifish to be useful in any regard. Last, beyond the utility of a sequenced genome to the *Fundulus* community, broader impacts are poorly described.

Overall: Very Good