

NATIONAL SCIENCE FOUNDATION
Panel Summary Review

Proposal:1120512

PI Name:Whitehead, Andrew

INSTITUTION: Louisiana State University & Agricultural and Mechanical College
NSF PROGRAM: EVOLUTIONARY GENETICS
PROPOSAL TITLE: Collaborative Research: The genomic basis of dramatic, rapid, convergent evolution in the killifish *Fundulus heteroclitus*

PANEL SUMMARY:

Panel Summary

Evolutionary Genetics Spring 2011 Panel

Results Of Prior NSF Support (If Applicable):

Prior awards have been made separately to Whitehead, Shaw, Crawford, and a joint award made to Crawford, Whitehead and Oleksiak.

Results of prior support Criterion I (Intellectual Merit):

These grants have resulted in multiple manuscripts in quality peer-reviewed journals.

Results of prior support Criterion II (Broader Impacts):

All have produced invaluable genomic resources for the communities. The funded projects have also trained a number of postdocs and students, including members of minority groups.

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

Prior reviews were enthusiastic, but concerns were raised about the ambitiousness of the entire proposal, including the timeline for sequencing, assembly, and annotation. This iteration of the proposal brought on additional PIs to strengthen annotation, development of bioinformatics tools, and databases.

CRITERION I: [Intellectual Merit]:

Intellectual Merit Strengths:

A major strength of the proposal is the community-based, collaborative approach amongst multiple labs studying the genetics of pollution tolerance of this species. Moreover, the investigators integrate results from QTL scans, population genomics, and gene expression studies to ask the question about convergent evolution in multiple population exposed to different chemical environments. The PIs have a proven track record of productivity in *Fundulus* genomics, which is an important model for both basic and applied research in ecological and toxicological genomics. The proposed research on the evolution and genetic architecture of pollution tolerance is interesting and timely.

Intellectual Merit Weaknesses:

The panel still felt that the proposed research was extremely ambitious. There was some concern that the genome sequencing would fail to produce a quality assembly, but the panel was split on whether or not this was a real issue in providing the necessary tools for Objective 2. There was also some concern over whether the community-based annotation of the genome would produce reliable annotations, and justification for the quality of work produced from this approach would be useful. Perhaps the biggest concern raised, particularly with the ambitiousness of the proposal, is why there are no funds budgeted

NATIONAL SCIENCE FOUNDATION
Panel Summary Review

Proposal:1120512

PI Name:Whitehead, Andrew

for graduate students or post-docs. Only three undergraduates are funded under one investigator (Whitehead).

CRITERION II: [Broader Impacts]:

This proposal creates an invaluable resource for the community and comparative genomics with the genome sequence of *Fundulus*. The community annotation will provide great training for the community, postdocs, and students.

Broader Impact Strengths:

The strength lies in the resources provided and the teaching opportunities through community annotation.

Broader Impact Weaknesses:

The major weakness identified is the limited number of funds provided in the budget to fund students for this ambitious project. It is assumed that the PIs will obtain students and postdocs by other means?

POST-DOC MENTORING PLAN:

N/A

SYNTHESIS AND RECOMMENDATION

The PIs propose to evaluate the genetic basis of pollution tolerance in the killifish, in a large, multi-investigator collaboration. The integrated approach is sure to identify genome regions and possibly genes associated with adaptation in multiple populations to test hypotheses about convergent evolution. The genome sequence will be a valuable resource for the community and beyond. There are concerns with the ambitiousness of the proposal, and the lack of funding of students and postdocs on the project, but overall, the research question is extremely timely and important.

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

PANEL RECOMMENDATION KEY:

N:Not Competitive, M:Meritorious, S:Superior, O:Outstanding

NATIONAL SCIENCE FOUNDATION
All Reviews (Jacket and PI copies)

Proposal:1120512

PI Name:Whitehead, Andrew

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

Institution:Louisiana State University & Agricultural and Mechanical College

NSF Program:Evolutionary Genetics

Principal Investigator:Whitehead, Andrew

Rating:Excellent, Very Good

Review:

What is the intellectual merit of the proposed activity?

The PIs propose an interesting project seeking to understand the evolution of pollution tolerance in *Fundulus heteroclitus*, a trait that has evolved independently at least four times. The ultimate goal is to identify the "genomic architecture and candidate genes underpinning adaptive convergent phenotypes."

The first aim is to generate a whole genome reference sequence for *F. heteroclitus*. While not the primary goal of the research, it is clear that there is a large community of *Fundulus* researchers and a genome sequence would be of broad value. The community seems mobilized to reach this goal - the plan to achieve it seems sound, but this is a pretty large genome and there is certainly a risk that the approach (including 40kb mate pairs which are not, from what I hear, a slam dunk) will come up a bit short of complete assembly and this will impact the rest of the proposed work.

The second aim is to map loci based on multiple existing lines of experimental evidence (QTL mapping to transcriptome scans of tolerant and sensitive populations) to the genome to identify candidate genes involved in adaptation. The third aim will further validate candidate gene identification by screening whole-genome patterns of polymorphism in population samples (new data).

The proposed work will test the hypothesis that adaptive convergence resulted from selective sweeps of a small set of ancestral polymorphisms, and that the same genes account for the adaptive phenotype among independently derived populations. It should be noted, however, that so far, existing QTL data does not strongly support this view; only few of the markers associated with tolerance have been the same across all the tolerant populations.

In the second major data gathering activity (aside from assembling the genome), pooled DNA from 30 individuals from each of four tolerant and sensitive populations will be sequenced. These sequences will provide allele frequency estimates to identify regions of genomes under recent selection that are shared among converged populations - i.e., genome regions with reduced levels of within-population polymorphism, and elevated levels of population subdivision (F_{ST}).

Because the polluted sites where tolerant populations reside have been contaminated with tolerance-associated chemicals only in the last fifty years, the PIs predict that pollution tolerance in *F.*

NATIONAL SCIENCE FOUNDATION
All Reviews (Jacket and PI copies)

Proposal:1120512

PI Name:Whitehead, Andrew

heteroclitus has evolved by fixation of standing variation; if this turns out to be the case, I'm sure this will become a textbook example of how rapid evolution to environmental stress occurs.

What are the broader impacts of the proposed activity?

Strong both in terms of developing a community resource (the Fundulus genome) and in terms of the broad training opportunities and inclusion of underrepresented groups.

Summary Statement

Important problem, strong group, strong research plan and strong broader impacts

NATIONAL SCIENCE FOUNDATION
All Reviews (Jacket and PI copies)

Proposal:1120512

PI Name:Whitehead, Andrew

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

Institution:Louisiana State University & Agricultural and Mechanical College

NSF Program:Evolutionary Genetics

Principal Investigator:Whitehead, Andrew

Rating:Excellent

Review:

What is the intellectual merit of the proposed activity?

This is a large, multi-investigator proposal that examines the genomic basis (using genome scans, QTL analysis, and transcriptome profiling) of tolerance to pollution in the killifish, *Fundulus heteroclitus*. A great strength of this proposal is in the unification of many individual research labs and projects. The genetic basis of adaptation to pollution is an important one, as many populations and species are exposed to anthropogenic inputs of contaminants. The PIs are all well qualified to conduct the work, and preliminary work in this system strengthens the readiness of the work proposed herein.

What are the broader impacts of the proposed activity?

The proposal will train students and provide genome sequence, but perhaps the biggest impact is during the community-based manual curation of the genome sequence. The curation exercises will provide invaluable integration across the community and training for scientists at all stages of their education and/or careers.

Summary Statement

This is a well-written, and well-integrated proposal that aims to evaluate the genomic architecture of pollution tolerance in killifish. The research is broadly applicable as a model for adaptation to polluted environments. The broader impacts integrate the research with training of students and post-docs, and is particularly strong in doing so through the community-based curation objective.

NATIONAL SCIENCE FOUNDATION
All Reviews (Jacket and PI copies)

Proposal:1120512

PI Name:Whitehead, Andrew

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

Institution:Louisiana State University & Agricultural and Mechanical College

NSF Program:Evolutionary Genetics

Principal Investigator:Whitehead, Andrew

Rating:Good

Review:

What is the intellectual merit of the proposed activity?

Clearly stated hypothesis that adaptive convergence emerged from selective sweeps of a small set of ancestral polymorphisms and that the same genes account for the adaptive phenotype of toxin tolerance among independently derived population of killifish that encounter diverse but mechanistically-related pollutants.

The rapid, recent and repeated evolution of toxin tolerance in *Fundulus* allows a powerful test of evolutionary hypotheses in an ecologically relevant model with implications for ecotoxicology and possibly conservation.

This is a 3 year multi-PI proposal including Whitehead: Louisiana , Oleksiak and Crawford: Florida, Warren: WUSTL, Hahn: WHOI , Shaw and Colbourne: Indiana. This research group has been at the forefront of ecological genomics, consistently developing the *Fundulus* model system for several years. The more recent pairing with WUSTL and IU increases access to technology and bioinformatic tools necessary for this proposal.

The first aim describes genome sequencing and assembly using on male from an inbred stock. The work will be done using Illumina paired end reads primarily for short insert (200-300bp) libraries. Research groups at WUGC and CGB have experience with assembly for this type of sequence data for de Novo (?) assembly of a vertebrate genome. The annotation efforts are proposed to be a combination of 'jamboree' and 'cottage industry' models in which the 'community' is first trained and inspired at a 5 day workshop. The PIs note the potential duplication of effort and variation in quality. This may raise more of a stumbling block than they anticipate.

With the genome in hand the group intends to map pre-existing datasets to the genome in order to address the question of structural vs. regulatory changes for adaptive phenotypes. They currently have QTL data, gene expression data, and genome scan data from multiple tolerant and sensitive populations. By focusing on genomic regions implicated by multiple datasets and shared among tolerant populations they will likely bias results in favor of the hypothesized 'standing variation' explanation for repeated evolution.

In a less biased approach the third aim will use population level resequencing for 30 individuals from each of 8 populations in order to quantify polymorphism and associate regions of the genome that share similar patterns of diversity between tolerant and sensitive populations.

NATIONAL SCIENCE FOUNDATION
All Reviews (Jacket and PI copies)

Proposal:1120512

PI Name:Whitehead, Andrew

.....

As the success and value of the project rest so greatly on members of the fundulus research community and their participation in annotation, additional letters of support demonstrating the strength and involvement of this community would seem important. Who will be the 40 workshop participants?

What are the broader impacts of the proposed activity?

Three undergraduate students will be employed for one year in the lab of AW to conduct bioinformatic research (it is unclear exactly what portion of the genome annotation will be done by these students or what projects they will do.)

.....

It is curious that no graduate student or postdoc training will be funded through this research proposal beyond the training provided at the annotation workshop.

Summary Statement