

NATIONAL SCIENCE FOUNDATION
Panel Summary Review

Proposal:1025432

PI Name:Whitehead, Andrew

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

PANEL SUMMARY:

Panel Summary

Evolutionary Genetics Spring 2010 Panel

Results Of Prior NSF Support (If Applicable):

This is a collaborative project with several investigators who have had previous grants.

Results of prior support Criterion I (Intellectual Merit):

It is hard to tell from the limited space but grants to Whitehead, Shaw and Crawford have led to a large number of papers.

Results of prior support Criterion II (Broader Impacts):

Broader impact's are listed only for a grant to Shaw, in which case the impact of a *Daphnia* grant seems to have been substantial.

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

A previous submission received high marks for the general goal of developing the *Fundulus* model for studying adaptation to pollution and to develop the *Fundulus* genome. There was concern that the experimental design for mapping and identifying genes was not well explained. These concerns have still not been adequately addressed.

The previous submission was criticized on broader impacts for a lack of emphasis on student training. This new version now places a much stronger emphasis on education and training.

CRITERION I: [Intellectual Merit]:

Intellectual Merit Strengths:

The main strengths of the proposal are the collaboration by several strong investigators, and the plan to sequence the *Fundulus* genome. Reviewers were in agreement on the fundamental value of *Fundulus* as a model system for studying adaptation to polluted environments, and that having a genome sequence would really make it possible to get down to the individual genes.

Intellectual Merit Weaknesses:

Reviewers were also in agreement that the plans for mapping genes, and for selected resequencing of adapted lines are vague and confusingly heterogeneous. A variety of genome-related projects are

NATIONAL SCIENCE FOUNDATION
Panel Summary Review

Proposal:1025432

PI Name:Whitehead, Andrew

described as in progress, but it is not really clear how ongoing AFLP -based and QTL work will be joined up with the genome sequence to identify genes. It seems that different investigators included components of their current work at the expense of actually developing a clear and joint plan for how the genome will be used.

CRITERION II: [Broader Impacts]:

Broader Impact Strengths:

There is an extensive plan to use web-based tools, workshops and conferences to ensure that the *Fundulus* genome is annotated and becomes fully useful to the broader community. The plan for an annotation workshop is interesting. Despite some questions on its feasibility, it may be a way to efficiently make a lot of progress in a typically slow and expensive task. The genome-related tools and workshops will have a big impact on the *Fundulus* community. The workshops will also include a large number of undergraduates.

Broader Impact Weaknesses:

No weaknesses were mentioned.

POST-DOC MENTORING PLAN: N/A

SYNTHESIS AND RECOMMENDATION

The grant brings together a strong group of investigators who work on an important model system for the study of the genetic basis of tolerance to pollution. The goals and plan to sequence the *Fundulus* genome makes a lot of sense.

This grant was criticized previously for a lack of clear convincing plan to use the genome that would be obtained. In this latest round of review these general concerns persist and reviewers were far from convinced that the mapping studies would actually identify genes and SNPs responsible for the adaptations.

The panel recommendation is Meritorious.

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

PANEL RECOMMENDATION: Meritorious

PANEL RECOMMENDATION KEY:

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

**NATIONAL SCIENCE FOUNDATION
All Reviews (Jacket and PI copies)**

Proposal:1025432

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:Poor

Review:

What is the intellectual merit of the proposed activity?

The proposed work will examine adaptation to environmental stress in *Fundulus heteroclitus*. Through sequencing (Aim 1) and analysis the investigators will generate candidate genes (Aim 2) test whether there is adaptive convergence from standing variation or new mutations (Aim 3). Although *Fundulus* is a great system in many ways it wasn't clear how this system would provide additional knowledge beyond other systems, e.g. beach mice, sticklebacks.

The team of investigators are well suited to conduct the proposed research.

The proposed research is not really creative or original but is potential transformative to the *Fundulus* community since so much could be gained with a genome sequence.

Aim 1. One genome will be sequenced to 60x and assembled and another 8 genomes will be sequenced to 25x. In general the sequencing strategy is good. However, I could not figure out whether the eight strains would be done at 25x each or together. From what I could tell there will be 2 full flow cells plus 2 lanes to generate 60x coverage budgeted at 57k and 4 full flow cells budgeted at 125k. The cost makes sense but the coverage doesn't. If 2 cells is 60x then four cells should be about 120x which is 15x across 8 genomes. A minor note. HiSeq can generate 200 Gb in a single run, enough for the entire experiment. Also, there was no indication that WU-genome center offers competitive pricing with other Next-Gen service providers. The whole advantage of Next-Gen is that you don't need a genome center to sequence a genome.

Aim 2. The goal of this aim is to generate candidate genes by integration with ongoing QTL mapping, genome sweeps scans and expression analysis. Although a compelling case can be made for the general value of the *Fundulus* genome sequence it was difficult to evaluate these other projects since they were briefly described, e.g. what is the sample size of the F2 population, what is cM/Mb for *Fundulus*, etc. For the population mapping data it wasn't entirely clear whether a Scan will be conducted with the data generated by the proposed work (in which case there was little information to evaluate the method and value of the research) or whether a scan is being done and would be mapped onto the genome. Of greater concern is that most of Aim 2 seems to be generating gene lists. This is not really compelling science as it won't answer a specific question and its not clear that the candidates can be experimentally tested by transgenics or RNAi. Replication of the scan in multiple population may not get down to the gene especially if selection is strong.

NATIONAL SCIENCE FOUNDATION
All Reviews (Jacket and PI copies)

Proposal:1025432

PI Name:Whitehead, Andrew

Aim 3. The central question of whether adaptation occurs from standing variation or de novo mutations seems mute in the *Fundulus* system. The resistant populations are thought to evolve over dozens of generations so I fail to see how de novo mutations are contributing. Even under the strongest selection it would take 100s of generations for a de novo mutation to reach appreciable frequencies. In general I imagine that most adaptations that occur due to an environmental shift are from standing variation until all those are fixed and then mutation becomes the limiting factor.

I was also concerned about the use of population genetic tests as a means of validation since these tests have low power true positives may not replicate. The sequencing should provide some interesting comparative data but I wasn't sure they would really answer the questions proposed, typically a QTG and usually a QTN is needed, e.g. beach mice and sticklebacks.

What are the broader impacts of the proposed activity?

The plan is to include three minority undergrads at the workshop and to pay for them to work 20 hr/week on the project during the year. This seems like a lot of effort for just three undergrads. The impact would be much larger if more students were incorporated for a shorter period of time. Also, at least at my own institution where we have minority programs it would be difficult to find three minority students interested in 20 hr/week, even if paid.

Another part of the plan is to start a *Fundulus* consortium and have a workshop for annotation and training. This part of the plan is fine.

Summary Statement

The proposal attempts to address previous concerns by addressing specific hypotheses about adaptation using *Fundulus* genome sequencing. While the motivation for the *Fundulus* genome sequence is still solid I found the second and third aims problematic. Aim 2 is weak since it will generate candidate genes and the strategy for validation may not work. Also, the main question of whether adaptation occurs because of de novo or standing variation seems mute in a system where adaptation has only occurred for a couple of dozen generations. A de novo mutation could not contribute much to resistance even if it occurred within the first generation.

NATIONAL SCIENCE FOUNDATION
All Reviews (Jacket and PI copies)

Proposal:1025432

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 to sequence, annotate and map the genome of *Fundulus heteroclitus*, one of the primary model systems for adaptation to pollution and physiological stress response. The three aims of the study focus on three steps in the methodology, genome sequencing, mapping and validation of candidate genes, but after revision from the last version, also test specific hypotheses.

Strengths

The main strength of this proposal is the collaborative approach of PIs from five institutions. Five of the PIs have extensive experience with the genetic basis of pollution tolerance in *Fundulus*, with more specific expertise on genome sequencing (Warren) brought in as well. The consortium represents an extremely powerful alliance to undertake this big task in an efficiently and timely fashion. In fact, the idea of getting an entire genome and identified candidate genes for little over 800k would have been ludicrous a few years ago. However, the PIs went to great length to plan the study and make sure to consider possible difficulties that may arise along the way. There appears little doubt that the consortium will be able to sequence and annotate the genome and to identify genes involved in pollution tolerance.

Another strength of the proposal is the data that have already been accumulated in *Fundulus*. Detailed studies on pollution response and tolerance are available, possible candidate genes have been identified and hypotheses of mechanisms of adaptation exist. The proposed genome sequence will therefore be placed against a background of extensive physiological, ecological, genetic and phylogeographic data. It will therefore be more feasible to evaluate the biological significance of genetic patterns that are identified. Furthermore, as specific crosses for QTL mapping already exist, phenotypes and genotypes can be effectively integrated.

One of the most generally applicable implications of this study were not discussed, and that is the importance of conserving genetic diversity in wild population, even if it is not adaptive in extant environments. If pollution tolerance indeed arises from existing polymorphisms in ancestral sensitive populations, possibly even conveying a fitness cost in those populations, it would affect conservation practices worldwide and give credence to the idea that such diversity must be preserved.

Weaknesses

Perhaps necessarily because of the size of the task at hand, the proposal is still based on methodological approaches rather than testable hypotheses. The overarching hypothesis (rapid convergent evolution is caused by selection from polymorphisms in 'natural' sensitive populations) is exciting and has considerable implication for the long-standing idea of conserving genetic variability in wild populations. Specific hypotheses could be more clearly stated, though the project is by its nature

NATIONAL SCIENCE FOUNDATION
All Reviews (Jacket and PI copies)

Proposal:1025432

PI Name:Whitehead, Andrew

exploratory, and some of the specific issues in *Fundulus* are clearly described.

What are the broader impacts of the proposed activity?

Broader impacts are primarily centered on undergraduate involvement in the research project and a workshop with travel support for students and postdocs. The idea of getting students actively involved and making each responsible for a gene family in the annotation is an interesting one and may have great benefits.

In addition, the availability of a genome sequence would be of great significance in further *Fundulus* research. What was perhaps somewhat missing is outreach activity to local communities to explain the effects of pollution on local species and the interaction between environmental change, local adaptation and genetic diversity. Data from this and other *Fundulus* projects seem to present an excellent opportunity for such activities.

Summary Statement

This project proposes to produce and apply a genome sequence for *Fundulus heteroclitus*, a primary model system for evolutionary toxicology. While much of the approach appears methodological, some interesting hypotheses are proposed and the resulting genome will be of great value for ecotoxicology and evolutionary research. The broader impacts are interesting, but not outstanding, although the project may have great potential for outreach.

NATIONAL SCIENCE FOUNDATION
All Reviews (Jacket and PI copies)

Proposal:1025432

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:Very Good

Review:

What is the intellectual merit of the proposed activity?

This is a large collaborative proposal to sequence the *Fundulus heteroclitus* genome and to try to identify the genes responsible for adaptation to pollution. *Fundulus* has long been a model system for the study of adaptation, particularly in the context of pollution because of their ability to adapt and survive in highly polluted environments. The collaboration includes a genome center (Warren at Wash U), and the application includes a number of letters of support.

Fundulus does show extraordinary adaptability and there seems little doubt that a genome sequence would greatly facilitate the discovery of the genes involved in adaptation to pollution. With the exception of a large set of EST sequences there are not yet many genetic resources for *Fundulus* (e.g. no genetic or physical map). This grant application describes a number of large projects in progress including SNP discovery and mapping projects. The project would be easier to evaluate if some of these were in print.

The main idea is to generate a high quality genome sequence using an inbred line, as well as eight lower coverage genomes from lines selected from polluted and non-polluted environments. Then with the genome in hand investigators will be able to bring a variety of mapping approaches to discover the genes associated with adaptation to pollution.

The plan for genome sequencing and annotation seems well motivated, clear and feasible. *Fundulus* is an important model and a genome would greatly facilitate the discovery of genes associated with adaptation.

The grant is more difficult to follow on just how the genome will be used to find genes. For example, it is hard to picture just how the genome will be used to identify QTLs (section G.2.1). Are these already known but unpublished QTLs? If so the work would presumably require returning to the DNAs of already phenotyped individuals. Similarly the use of transcriptome and population genomic data are pretty sketchy. A lot of genome-level mapping work is described as in progress, but the only genome-related *Fundulus* papers in print are about gene expression or AFLPs. Also there is some text that reads like it was written by an ad agency - e.g. "Preliminary experiments...have found that up to 17% of metabolic genes have evolved adaptive changes in gene expression".

NATIONAL SCIENCE FOUNDATION
All Reviews (Jacket and PI copies)

Proposal:1025432

PI Name:Whitehead, Andrew

What are the broader impacts of the proposed activity?

Apparently the previous grant was criticized for not saying enough about training under broader impacts. This latest grant now says more about this. It is hard to see how a grant that convincingly proposes to sequence a genome for a model organism could come up short on broader impact û at this time every new teleost genome will have huge positive effect on many fields.

Summary Statement

This is a very strong collaborative project that would greatly advance the Fundulus model system by providing a genome sequence. The plan to identify genes associated with adaptation to pollution is hard to follow and there really is not much preliminary genetic work done that is relevant and published.

NATIONAL SCIENCE FOUNDATION
All Reviews (Jacket and PI copies)

Proposal:1025432

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:Very Good

Review:

What is the intellectual merit of the proposed activity?

The lead-PI has assembled a strong team of first-rate collaborators. Together they show that they have command of the various steps involved in sequencing, assembling and annotating a novel eukaryotic genome and to make this resource available to the larger community.

My enthusiasm is somewhat dampened by the hurried and occasionally sloppy writing, particularly since this is a revision of an earlier submission. Clearly, several individuals have contributed sections to this proposal, yet little effort seems to have gone into making it readable. Specific comments:

- The PIs propose to sequence one individual from an inbred strain, but do not say whether it will be male or female. Which is the heterogametic sex in this species? How inbred is this strain really?
- Do any BAC libraries exist for this species? Are any BAC end sequences available for this project? If not, would this not be a useful resource for genome assembly?
- Much of the experimental work appears to depend on research to be conducted by a collaborator at the US EPA. Given the crucial role of several of these experiments within the scope of this proposal, it would have been useful to learn more about how these experiments will be carried out, what the expected results are, etc.
- It is not clear why an annotation "jamboree" would still be necessary at this point. Also, what is the evidence that the distributed "cottage industry" model of annotation will in fact work well?
- Why do the PIs propose to generate a separate genome browser for the fundulus genome instead of taking advantage of existing and widely used genome browser environments (e.g., UCSC genome browser)? Maybe I misunderstand the rationale here?
- The PIs have responded to the criticism that their original proposal lacked specific hypothesis-driven aims, by adding two aims, genome-wide mapping/transcriptomics as well as candidate gene validation. This entire section comes across as hastily put together, even though it has clear merit once the reader manages to cut through the jumbled writing.
- For the transcriptomics experiments, we are told about a completed experiment, though no details are provided on experimental design, whether juveniles or adults were used, males or females. And do the results provide any (preliminary) evidence as to "whether tolerant populations have converged on similar compensatory responses to PCBs"?
- What would it tell us if TATAA sites, GC boxes, etc would differ across treatments? I can guess, but not everyone may be familiar with these concepts.
- In section G.2.4. data release policies are mentioned, which is laudable. However, they appear to be

NATIONAL SCIENCE FOUNDATION
All Reviews (Jacket and PI copies)

Proposal:1025432

PI Name:Whitehead, Andrew

the ones in place at the Washington University Genome Center. What about the other entities involved in this project? I would have liked to see a separate statement of a release policy that applies to the entire proposal.

- The management plan does not spell out how often the PIs will confer on progress and planning; how conflicts will be resolved; how authorship will be determined etc.

Smaller comments:

- Numerous grammar mistakes and poor stylistic word choices are scattered throughout, which made reading this proposal more difficult than it should have been, given the sound science
- Many acronyms are not spelled out at first use (e.g., PCBs, PAHs, SVs)
- The common name for *F. heteroclitus* is not mentioned, as far as I can tell
- The background section is poorly referenced, i.e., the reader has to take at face value some of the claims that are being made
- Figures would have really helped comprehension, particularly for the strategies employed for sequencing/assembly/annotation and the preliminary results

What are the broader impacts of the proposed activity?

The Broader Impacts are acceptable, though not out of the ordinary, especially given the scope of the proposed research and the outstanding opportunities it could provide.

Here is an opportunity to bring an entirely new generation of young researchers into a growing field, yet all that is planned includes a handful of minority internships. Apparently, these students will be used as annotation helpers, which may not be very rewarding. While the goal of training graduate students and postdocs in *Fundulus* genomics through workshops etc. is laudable, it is not clear how these individuals will be selected (or whether they will simply be sent by their respective advisors).

Summary Statement

The proposal by Whitehead and collaborators is a revision of an earlier submission whose main goal is the whole genome sequencing of the small teleost fish *Fundulus heteroclitus*, which has become a useful model system in studies of natural variation, adaptation and environmental toxicology. Overall, this is a strong proposal by a strong team of first-rate collaborators. Enthusiasm is somewhat dampened by the hurried and occasionally sloppy writing, which results in a lack of clarity.

NATIONAL SCIENCE FOUNDATION
All Reviews (Jacket and PI copies)

Proposal:1025432

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 proposal will develop genetic tools necessary for understanding the genetic changes necessary for population-level differences in killifish, and widely distributed intertidal fish. In addition to allowing the principal investigators to address their specific hypothesis concerning differences in pollution tolerance, the genetic maps and sequencing in this proposal will allow this and other groups to examine the genetic basis of other adaptations. In this regard *Fundulus heteroclitus* is a wise choice of a 'model' species. It has already been used to demonstrate the underlying biochemistry of protein thermal adaptation, and is widely used as a model for osmoregulatory physiology. It is widely tolerant of environmental change, but has shown population level variation in a variety of traits. I am especially intrigued by the authors proposal to address the potential trade-offs in development of adaptive traits (e.g. pollution tolerance versus hypoxic tolerance); this idea of constraints and trade-offs has been addressed theoretically, but needs more empirical evidence.

The proposal is concise and clearly outlines the major goals of the study. It would have been strengthened by more concrete description (or data presentation) of the actual difference in 'tolerance' among populations. The authors state that this has many endpoints, but I wonder if they are confusing tolerance (which I take to mean capacity to survive or maintain homeostasis) with adaptive response. Including population level differences in phenotypic plasticity and reaction norms would address this issue. Nonetheless, the proposal is very well crafted and has no major weaknesses. The collaborators have clearly outlined their roles, and all appear capable of carrying out their specific goals.

What are the broader impacts of the proposed activity?

In addition to addressing their own hypotheses, the outlined studies will provide genetic tools that can be used by others to examine adaptative trait and other evolutionary questions in this species. Their is substantial training opportunities (including inclusions of underrepresented groups), that are made richer by the collaborative nature of the grant. Their is a clearly outlined training program and a plan for communication among collaborators.

Summary Statement

NATIONAL SCIENCE FOUNDATION
All Reviews (Jacket and PI copies)

Proposal:1025432

PI Name:Whitehead, Andrew

NATIONAL SCIENCE FOUNDATION
All Reviews (Jacket and PI copies)

Proposal:1025432

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?

This revised, six PI, five institution collaborative proposal seeks to test the hypothesis that convergent evolution of four killifish populations occurred through fixation of the same small number of ancestral polymorphisms, such that the same genes are responsible for tolerance to pollution in each population. A major aim is to sequence the killifish genome to provide a framework for interpreting other kinds of genomic data (QTLs, AFLPs, RAD markers, etc.), collection of which is either completed or ongoing in the various labs of the PIs.

There are many proposals these days looking to establish the genetic basis of one or more interesting traits in a wide variety of organisms, raising some question about why we need to do this in a large number of different systems. Here, however, I think the PIs succeed in making the case that the *Fundulus* system offers some unique opportunities, especially given the breadth of work already completed by a variety of investigators, many of whom are associated with this project. That said, I find the hypothesis testing framework rather general and therefore not all that interesting: is evolution repeatable, what is the relative role of protein versus regulatory change, and what is the relative roles of selection on standing variation versus de novo mutations have become familiar questions these days, the answers to all of which seem pretty obvious (i.e., all things are possible and will be found to occur in different combinations in different systems!). Despite recent reviews on the topic, I find the issue of regulatory versus structural change to be particularly uninteresting as a general question; much more interesting would be a clear a priori hypothesis about why one kind of change might be more likely in this particular situation (i.e., adaptation for resistance to environmental contamination).

In general, however, this seems like an excellent system for evolutionary analyses and one in which there is a great deal of ongoing research and I think it make sense to support a whole genome effort. I am confident that the genome will both strengthen ongoing analyses described in the rest of the proposal as well as generate new research opportunities.

There was only one objective that I found unrealistic and that is the potential to identify the actual nucleotide level changes responsible for adaptation. Also, even with fine scale mapping, it may be difficult to confidently discriminate between regulatory and structural effects without significant functional assays (e.g., a fixed amino acid change in a candidate gene may simply have been swept to fixation along with a cis-regulatory mutation (qPCR on candidate genes, however, may provide some insight on this question). If the functional sites can not be identified, this also raises questions about the

NATIONAL SCIENCE FOUNDATION
All Reviews (Jacket and PI copies)

Proposal:1025432

PI Name:Whitehead, Andrew

potential for testing for the presence of the 'adaptive allele' in the ancestral (sensitive) populations.

I also have a suggestion that I hope the PIs will seriously consider: Rather than use individual fish for the eight additional populations, I think it would be more productive to do bulked DNAs from 20-30 or more unrelated fish from each population û with 25X coverage per population, most individual reads would from different fish/alleles and this would provide an outstanding opportunity to get at least rudimentary allele frequency information from across the genome. Patterns of nucleotide diversity, population divergence, and linkage could then be used in population genetic tests for detecting signatures of selection.

Minor comment: '17% of metabolic genes have evolved adaptive changes in gene expression' û would be more accurate to say 'evolved differences between populations have resulted in differential expression of 17% of metabolic genes' û all of these genes did not need to evolve in order to produce this effect.

What are the broader impacts of the proposed activity?

Broader impacts relate primarily to plans for encouraging and managing new collaborations (primarily among the PIs) and providing important genomic resources for an emerging model species. Community-based manual annotation will be supported through a training workshop with 40 participants.

Summary Statement

I like this proposal. Less opportunity for functional assays than in some other systems, but excellent opportunities for analysis of ecologically relevant traits and a strong existing research community.

NATIONAL SCIENCE FOUNDATION
All Reviews (Jacket and PI copies)

Proposal:1025432

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 work will provide a genome-wide overview of loci involved in inherited tolerance to environmental contaminants in killifish. Rather than focusing on specific mechanisms, e.g. altered xenobiotic metabolism or reduced penetration of organic compounds, this proposal will shed light on all loci involved, at the same time, in different populations. This wide scope is a very strong aspect of the proposal. For a long time, animal physiology has stayed with specific pathways and mechanisms that underly adaptive responses. Such detailed approaches have provided many interesting mechanisms but may prevent one from noticing interactions between different defence mechanisms or adaptations. The genome-wide approach taken in this proposal will survey, in a holistic way, all aspects of the adaptive phenotype. This is the more appropriate since where we are only beginning to explore the still unknown territory of evolutionary adaptation to the environment.

What are the broader impacts of the proposed activity?

Evolution is known to be fraught with convergence, classical examples being the body shape of aquatic animals and the lens eye. Most of these examples relate to similarity in morphology. Convergence on the physiological level is less expected because there are often many different biochemical pathways that can lead to the same adaptive phenotype. Killifish (*Fundulus heteroclitus*) populations have evolved tolerance to environmental contaminants several times and independently. The quickness and repeatability of these evolutionary changes suggest that a similar mechanism is involved, that can be traced back to a limited number of polymorphisms in the ancestral population. It therefore represents a unique case of investigating convergent evolutionary change and test an important, fundamental hypothesis of evolutionary theory. The research will also address the question whether stress tolerance evolution is due to changes in coding sequences or changes in cis-regulation of key proteins involved in tolerance. Given the extensive inter-individual gene expression variation already documented in *Fundulus*, the latter possibility seems the most likely, however, some of the candidate loci involved such as Cyp-genes, are known to consist of large protein families and structural variation may also be important. This proposal is, as far as I know, the only project in which genome sequencing and genome-wide polymorphism mapping are used to analyse evolution of tolerance to environmental contaminants.

Summary Statement

NATIONAL SCIENCE FOUNDATION
All Reviews (Jacket and PI copies)

Proposal:1025432

PI Name:Whitehead, Andrew

The proposal uses a number of innovative techniques, such as Illumina sequencing for whole genome assembly and restriction site-associated DNA polymorphisms for ultra high-throughput genotyping. The research strategy focuses on a high-quality complete reference genome plus genome scans for a series of populations showing different degrees of tolerance to environmental contaminants. The technological aspects of the proposal, although extremely demanding, are very professionally elaborated in the proposal. The groups involved will use the latest bioinformatics tools to analyse the data, and given their extensive experience in the field, can be trusted to turn this project into a success. I was highly impressed by the technical quality of this proposal.