Hi everyone. Welcome to our Friday afternoon seminar series from Department of Biological Sciences. I'm very happy to have Dr. Andrew Whitehead from UC Davis here today. He's from the Department of Environmental Toxicology where he studies the genomics of killifish (am I pronouncing that correctly?). And Dr. Whitehead, as it turns out, received his PhD from UC Davis as well. But then, spent as a postdoc in University of Miami before taking a faculty position at LSU in Baton Rouge, Louisiana, where among on things he worked on the Deepwater Horizon oil spill and house effects for--on fish and birds, I believe, right? And from there, he--his research interests were changing or were increasing in the genomics area. And so he saw an opening at Davis and took it. And so he's been at Davis for the last year and his research group is building now. So we're very happy to have him. It turns out, we have a birthday sort of thing going on where our birthdays are very close to one another, and our six years old are both having birthday parties tomorrow afternoon, right when it’s supposed to rain. So, there we go. Alright, so welcome.

Dr. Andrew Whitehead, Department of Environmental Toxicology, University of California, Davis>> Great. Thank you.
Well, thank you for the invitation to come visit you Chico. It's my first time here, and to come and share some of our research with you. So, what I--what we study in our lab is very generally, we're interested in some of the mechanistic basis of resistance or resilience to environmental stress. And not only that, but how those mechanisms of resistance and resilience vary among individuals, populations and species. So we kind of take a comparative approach to our science. And we're interested in lots of different types of stressors. Many of which are natural or human source, environmental stressors. So, Louisiana was actually a good place for starting my career in environmental stress biology.
I guess we moved there two weeks before Hurricane Katrina hit. And we left just after the Deepwater Horizon oil spill. But I love this. This is a cover article, "The New Orleans levee, who knows like the onion". You know, all that sort of satirical news thing. So the New Orleans levee is like Louisiana equivalent. Their tag line is, you know, we don't hold anything back. And this is a cover article just after the oil spill, you know, "Disaster recovery, key to Louisiana's future". And it says here, if you can't read it, marsh fires, oil disaster, chemical dumping, tropical cyclones and industrial accidents have made some of Louisiana, the epicenter of the international disaster industry. Governor Bobby Jindal predicts, disaster response and disaster cover will make up 40 percent up state economy by 2015. It's even higher than it's possible if they can break in--somehow break into the earthquake market. But California kind of has that covered. And so, maybe that's why we moved here, it's to breaking the earthquake market. But, you know, so Louisiana has, you know, suffered a lot of environmental damage. Many of the hands of people, you know, we've--it's the fastest disappearing land mass on earth. The South Louisiana marshes buried a football field every 35 minutes, goes underwater because of this, you know, channeling of the Mississippi river and that land-building sediment has been disappearing over the last number of decades.
So, South Louisiana is sinking. Also, there's habitat loss. You impose on top of that, you know, hurricane damage, so there's increase stress from natural events. And then, you've also got imposed on this human influences, you know, urbanization, and oil and gas industry, and all that goes along with that. So we're interested in what are the consequences of these kinds of stress exposures on resident species. How does that affecting organism's resistance or resilience? Which species are really resilient to this kind of environmental probation? And what are the mechanisms underlie that? And how those mechanisms have evolved under strong selection pressure for persistence of these habitats?
So we tend to think of what are the sorts of the potential long-term outcomes of this persistent environmental stress? And I figured that in three different ways, there are--you know, one consequence of this kind of exposure is, you know, a demographic decline until they're locally extinct, right? Either through compromised fitness or because, you know, everyone has moved out of there by immigration.
The second possible long-term outcome is there's a physiological competitive response. There's an acclimation whether that be the ability to metabolize stuff or to developmentally compensate for these kinds of environments. There's this idea of being able to physically or physiologically compensate in sort of ecological or physiological time.
And then there can also be compensation on evolutionary time scales where natural selection sorts among phenotypic and genotypic variants in populations evolved over time to gain greater resistance to environmental stress. So which of these outcomes happens depends on lots of different things.
It depends on the severity, the persistence of the stressor or lots of inherent traits of the stressor itself, plus it depends on lots of traits of resident species. It depends on what your niche breadth is. If you have the ability to move out of their and still persist in a local area, you know, that's inline with your migratory ability. And some species are just more plastic than others. They've got more metabolic capacity than others or behavioral flexibility. Population size matters as far as the adaptive potential of a species or of a population. Larger populations can harbor more genetic variation. And genetic variation being the substrate for natural selection makes larger population more likely to evolve by natural selection in a particular habitat. And then evolutionary history matters as well. If you happen to come from a lineage that, you know, each plant with toxic metabolites, then perhaps, you have the metabolic machinery that's going to enable you to also metabolize, you know, the hydrocarbons and oil. Or perhaps, prior exposure history to related chemicals has enhanced your ability or primed your physiology to deal with this new environment. So clearly, there are lots of things about the stressor itself and about the species that are potentially at risk that will determine whether they decline to zero, whether they are able to deal with it physiologically, or whether they can potentially adapt.
So today, you know, I'm going to first emphasize the point that genotypic and phenotypic variation is ubiquitous in nature. So hallmark of natural systems. And then I guess, I will make the point that variation is important. Many scientists is going to think of variation as a bit of a nuisance. And for some biological investigations, it is kind of nuisance. You know, you want a genetically homogeneously line to the study, you know, the biochemistry of, you know, pathway, X, Y or Z. But in understanding biodiversity, clearly, we want to be able to embrace the kind of variation that we find in nature. So variation, to me, is an opportunity to discover something about the mechanisms of action of particular stressors to generate markers, for--be able to track how organisms are dealing with stress in the environment and to learn something about the population dynamics and response to environmental stress. And then, I'll give an example from our own research in the lab where we have exploited evolutionary--this really cool and crazy adaptation story in killifish to offer insights of the mechanisms and the genomic targets of naturally evolved variable tolerance. In a species like humans, that's operon [phonetic] like humans, it's not the most line in the lab. This is a natural species that--it's got a lot of genetic variability and that interacts with pollutants in nature. And in certainly human health, we're interested in that sort of human environmental interaction as it relates to disease.
So it kind of goes without saying these days that phenotypic variation is a real hallmark of natural populations, you know, including humans. And it's this diversity, this phenotypic and diversity and the genetic diversity that can underlie it, that's really the substrate for evolution processes like natural selection. So humans, you know, you just look around the room, you see plenty of phenotypic diversity. Look around the world, you see an enormous amount of phenotypic diversity. And it's not only phenotype, it's not just interactions with the environment,
there is also a ton of genetic variability within human--within the human species as well. Maybe not as much as other species that we tend to think of, but there’s still plenty of genetic variation in there. So these are studies that have come up over the last couple of years that have shown that a lot of this genetic variation harbored within humans is partitioned across geography. So here, you know, we’ve got different genetic types coded with different colors. And you could see as you go across the globe, we’ve got different genetic groups that’s partitioned out by geography. And some of that--So here, we’ve got sort of global variation. But then also, you know, within Asia, within Europe, there's also a genetic variation that partitions across regions as well. That's not only on a global scale,
you also find this genetic variation that partitions out at local scales too. Different genetic, different people, living in different parts of the globe have different genetic backgrounds. And you can even see this at the scale of Europe. And this really blew me away when this came out. And John Novembre published this a couple of years ago and received even more, more of this as people come up with more and more detailed information about genetic structure in human populations. Here, what they did is they collected DNA from humans across Europe, you know, each of those little letters, these are different sample from a different geographic part of Europe. They're color-coded by the region that they were collected from. And then they genotype them at a thousands of markers across the genome. They just used an--a statistical ordination approach to sort of partition out variation in two-dimensional space. And that, we could picture later that, a map of Europe. So, this is just a two-dimensional partitioning of genetic variation mapped on top of Europe and the purple ones are the ones that came from that part of Europe. The red ones are the ones that came from Great Britain. I mean it's staggering. It's like you can't make up data that are better than this. So, it's really stunning to me that there's this much genetic variation that partitions with--across really fine geographic scales of humans. And some of this is important genetic variation too.
So, human genome has been around for a few years now in various nature papers saying, we’re finally done with it now. Now, this is the complete one now and so on and so on. But back around early 2000s, they published the human genome. And it's important to note that the entities for sequencing the human genome was not just to have a human genome. It was to have a reference to understand human genetic variation, right? So, is this variation among humans in susceptibility to disease or susceptibility to different drugs that we really want to understand? So, is this variation across species that this reference genome was made for? And sometimes that's easy to forget and this is kind have been slowly penetrating into different sciences
including, you know, toxicity testing. But around, you know, 2007 when the National Academy of Sciences, you know, published this report on toxicity testing, your modern toxicity testing, they really gave voice to the importance of understanding and getting a handle on biological variation. So, this report highlights that population-based data or key elements to this emerging paradigm in environmental toxicology which reflects as growing recognition that variation and genetic background, and among individuals and populations can have significant impacts on sets to be the disease, sets to be to environmental stress.
And here is a paper that came out last year showing that a lot of this variation is actually important. It isn't just sort of random genetic variation there, that some of it is biomedically important. So they found that--they discovered after sequencing, you know, 14,000 people across the genome, they found there's a lot of genetic variation in genes that were drug targets. They were targets for common drugs. So, a lot of this variation could be biomedically important.
And not only is there are lots of variation, but there's variation in that variation among human groups, OK? So, this is, you know, part--these shows along the Y axis, the number of sites that are variable. And these partitions denote by population so, you know, African-Americans are more variable than Asians, and Europeans than the Fins. So, not only is there variation within a group, but there's variation in that variation across groups too that can be medically important.
So, variation is ubiquitous in outbred species. Genetic variation is ubiquitous and a lot of it has phenotypic consequences. This variation varies across populations. So, evolutionary history, what group--what genetic group you have affiliation with in history matters for understanding the partitioning of functional variation (drug/disease susceptibility) in natural populations. Comparative approaches provide insights into the origin and basis of biological novelties and universals.

So, in our lab, we take a comparative approach. We compare these different groups that have different genetic backgrounds and different histories to provide insights into the origin of, you know, these novelties between populations, these novel susceptibilities if you will, but also to understand what's universal, what doesn't vary across populations. That's presumably under strong selective constraint to not vary.
So, the conceptual model that we follow in our lab is, you know, starts with that recognition that individuals, populations and species, they vary in such when introduced to stressors. And these variants can be exploited to really explain something about the mechanisms of action of stressors. What's the molecular biology in the signaling pathways that unfold in response to the stress? And we use this system to explore the propensity of individuals, populations or species to maintain homeostasis and to persist in stressful habitats. So how might you approach studying the importance of genetic variation?
Well, laboratory mutant models have been very powerful in biomedical sciences. We're able to generate a strain that's got some sort of mutation that makes it susceptible to disease X, Y or Z, or susceptible to drug X, Y or Z, or has some phenotype that helps you study the genetic basis of it. So laboratory mutants have been a very powerful model in biological sciences and
wild animal models are also starting to gain some attraction as far as being exploited as evolutionary mutant models for studying disease.
So this is a really nice paper that was published a few years ago arguing for the use of this evolutionary mutant models to study human disease. So using ice fish to study things like anemia and osteopenia, using K-fish that have lost their eyes, all because they live in caves, to understand something about the mechanisms that underlie retinal degeneration and so on and so on. So why use evolutionary mutant models?
Well, there--these evolutionary models are outbred. Isn't some inbred strain that's been synthesized in the lab, these are things that evolve in nature and they're genetically variable like humans are. And the types of mutations that you find in natural evolutionary mutants are more representative of the types of mutations that you find among human populations, OK? So mutations that you find in nature are less likely to obliterate gene function. So you're--And you're also going to capture the types of mutations that have environmentally dependent penetrants and capture mutations that affect many traits or capture mutations that are of polygenic or that cause traits that are polygenic.
So the adoption of wild animal models, evolutionary mutants is adapted by some because, you know, these variants that are sort of by natural selection are by definition important. If something is evolving by natural selection, it means that it's important for something. And so evolutionary analysis in comparing closer related species that vary in these traits can distinguish variants that are important in this way. Variants that are sorted by natural selection are therefore important.
So this is our conceptual model.

Model/Concept/Paradigm:

Individuals/populations/species vary in sensitivity to stressors, and such variance can be exploited to explain mechanisms of action of stressors, and to explore the propensity of individuals/populations/species to maintain homeostasis and to persist.
And our question is how do genomes function and how do they vary in that function to enable physiological resilience and to ultimately enable success in one environment, perhaps not another? How do these functional genomes evolve?
So our model animal is this cute little guy. We study a number of species within the genus Fundulus. And the reason we study killifish is a few fold. One is they're real superstars when it comes to sort of physiological tolerance to environmental stress. You know, they've got a range from Nova Scotia. At least this one specie is from Nova Scotia down to Florida along the Atlantic Coast. It's a huge temperature range. They could tolerate big swings in temperature. You can find them from sea water to fresh water habitat in an estuary, so a huge range of salinities. You can take the oxygen and suck it out of the water and they don't care, they just come out to the surface and start gulping air, you know. So these are pretty tough little guys, but they're wusses when it comes to human pollution. Among fish, they're actually pretty sensitive. So what's really cool for us is that not only are they tough, but they also vary in that toughness. There are some species that are tougher than others. So they live in really diverse ecological niches. Now, we can work with them in the lab. We know a lot about them 'cause people have been studying them for a while. And we've got a growing genomics tool kit. We just sequence their genome and we've got a lot of tools, microbiology tools that help us do some fun biology. So, the one species that we really focused on Fundulus heteroclitus, they live in estuaries. And estuaries are stressful places, right?
Because there is natural stressors. You know, salinity fluctuates on a tidal basis, on a daily basis and on a periodic basis as there are storm surge or there's a big rainfall event that deluges things out. So salinity varies. They're shallow water habitats. So temperature swings could be huge. So lots of natural environmental variation, there's hypoxia, a periodic hypoxia and all that kind of stuff. Estuaries are also stressful because of, you know, human activities.
So humans tend to live close to estuaries. There's a lot of human activity around estuaries. So you all know what this is, the largest marine oil spill in history that happened just off the coast of Louisiana, the largest estuary in North America. And really had some impacts in some of these marsh areas and we've done some research along those lines, but I won't be talking about that today. So we've got these, you know, periodic stressors, you know, happen from accidents.
And then we've also got more persistent human cause stressors in ecosystems, you know, from urbanization or from the dumping of industrial pollutants that are highly, really persistent in the environment that don't go away. So humans who live close to estuaries and we put our stuff in estuaries and that can cause problems.
Estuaries: anthropogenic stressors
So what we do in our lab is more specifically compare the physiological genomics.
So we study genomic phenomena that emerge at physiological time scales like compensatory responses or acclimation responses or toxicity responses to some change in the environment. And that gives us some insight into the mechanisms that enable physiological resilience to fluctuating environments.
But we’re also interested in genomic phenomena that emerge on evolutionary time scales, you know, like phenomena that—like natural selection—excuse me, adaptive processes. And that gives us some insight into the nature of physiological resilience in different habitats, in different niches.
So, you know, for these kinds of experiments, we did sort of physiological challenge experiments in the lab and field. And--But we do that in a comparative context. So we do those challenge experiments by comparing different species that vary in their history and vary in their sensitivity to environmental stress.
So we use genome expression profiling in many of our studies to offer sort of this top-down systems level insight into the mechanisms that underlie these adjustments to the environment, OK? So, we're interested in--at a global scale which genes are getting turned on and turn off in the cell in response to some environmental change. And then we compare those responses to the environment to offer insight of the mechanisms that contribute to all of these variations that we see in nature. Excuse me. OK. So I told you that killifish are pretty resilient when it comes to natural stressors.
But when you compare them to other species for example in their sensitivity to dioxin or dioxin-like compounds, they're at the sensitive end of the continuum. So here is a bunch of fish species, killifish with the black bars, other fish with the gray bars and, you know, dioxin, the lethal concentration and along a long scale there. And you could see that among fish, there's about an order of magnitude or so, a little bit more range of variations among fish species and their tolerance to this sort of model chemical. And killifish are at the sensitive end of things. It takes less chemical to kill them than most other fish species, OK?
But if you go along the East Coast and you go to some of these really polluted estuaries, you know, where there are--it's a bit of long legacy of pollution of PAHs and dioxins and PCBs, you've tend to find populations of killifish that look perfectly fine that are living there. If it's appropriate habitat as far as, you know, partially submerged salt marsh, you're going to find killifish there. And it's not Blinky, the fish with three eyes. I mean these are normal looking fish with, you know, healthy looking populations, not riddled with disease or anything like that. And not only do you find them in Bridgeport which is a superfund site, but you find them at many superfund sites along the East Coast.
These are sites that are federally designated by the EPAs, priority clean up sites 'cause they're so badly polluted. So we've been studying this populations, and for the context of this talk, I'm going to show you--we've been comparing these populations from polluted sites which are in red and we've been pairing them.
with other populations from clean sites which we've listed in blue here. Most of the data I'm going to show are from these--those six northern populations and I've left out Elizabeth River stuff 'cause it's still--we're just processing those data right now. So, what we first did and others have done this stuff too, it's known--have been known for a while that all of these populations vary in their sensitivity to these pollutants,
but here are some of our data. So just for the sake of what I'm going to show you now, the, you know, red data are going to be from the polluted populations and blue data are going to be from the reference populations. And the shape of the symbol unites our geographic pairs of polluted reference, OK? So here is the circle pair, this pair up here. And this is increasing log doses of our model toxic and this is a PCB. And this is--These are their survivorship curves. And this is pretty stunning, right? This is not a subtle difference between populations. I'll remind you, this is a log scale. So our--It takes about three to four orders of magnitude, thousands of times more chemical to kill a fish from a polluted site than a fish from a reference site. And I should mention that these are fish that have been raised in a common clean environment for at least two generations. So the differences between populations are genetic, not induced by their polluted environment. They've been raised in common clean environment for at least a couple of generations. So that's our circle pair.
Here is the square pair, same thing survivorship curves all set by a couple of orders of magnitude.
Here is the triangle pair. Pretty much wherever you find a population living in a polluted site, they're extremely tolerant to these types of organic pollutants. So not only a survivorship offset, but a lot of the sublethal endpoints are also offset.
So, OK, well, we see this girl, these are developing embryos and at lower concentrations, these chemicals cause developmental abnormalities like heart malformations or hemorrhaging in the tail. And so, when you look at those
developmental abnormalities where they increase number means more abnormal. You can see that those are also offset by a couple or some magnitude, so lethal and sublethal effects. There are huge differences between populations from polluted and clean habitats.
So let's go back to this figure here. Here is a range of tolerances across all fish. Let's superimpose on that, our fish from superfund sites.
But I hear jaws hitting the floor, right? I mean that's crazy. So there's more variation in tolerance harbored within a single species than there is across all species. And that's evolved by natural selection in a few dozen generations. It's evolved repeatedly, it's not a subtle phenotype and it's evolved quickly. So we want to know, how do they do that?
So we take a comparative evolutionary toxicogenomics approach.
And our question is, what is the mechanism of derived tolerance that enables these guys to explore this pretty extreme niche? These are extremophiles, right?
And not only that, but is it the same mechanism each time this evolves. So we know from the population genetics that the polluted-site guys are more closely related to fish from nearby clean sites than they are to each other. So it appears that this tolerance has evolved multiple times. It wasn't just one derivation of tolerance and then spread it to polluted sites. It was in the multiple independent evolutions of this tolerance. So the same mechanism, do they skin the cat the same way, or the--each disease population comes up with the same or different genetic mechanistic solutions to these?
So I went back to this dose response experiment and we are interested in discovering what are they doing differently? What are these tolerant guys doing differently when they're challenged with these chemicals compared to the sensitive guys? And what's remarkable to me is a study that was published by a colleague of mine just before we started doing this study, showing that gene expression across the whole genome was identical under sort of constant clean conditions between these populations. There wasn't some big differences in gene expression under constant conditions. So we decided to challenge these organisms. Maybe there will be some differences in how the genome is regulated that you only see when you challenge them with chemical, OK?
So we look at what the tolerant and sensitive guys were doing differently at this common dose of 200 milligrams per liter. And so what we did is we look at the 220 and 200 dose for the sensitive guys, and we looked at that 200, 2,000 and 20,000 doses for the tolerant guys. So we call this our common dose that we experiment, control to the 200. And then we compared their effects match dose range which is the blue panel and the red panel, right? So the blue panel captures the sensitive populations, sort of a highest no effect dose and lowest effect dose when you look at how that superimposes upon their curves. And then, you know, we looked at a dose range that's offset by a couple or some magnitude up for the tolerant guys that captures the same effects range, right? It captures their highest no effect level and their lowest effect level.
So we have this question, we've got clearly this parallel adaptation that each of these populations have arrived to the same sort of tolerant phenotype. So our question is, is it the same mechanism, or have they each arrived different mechanisms? So, we did a common garden experiment where we pulled all these populations into a common clean environment, raising for a couple of generations to minimize any environmentally induced differences. And then challenge them with PCBs and looked at which genes were responsive to this challenge across the whole transcriptome, across all the genes in the genome, or as many as we could study at once. And so, if you think of their response, you know, in this principal components phase where you sort of take a response and you can plot out that response in a two-dimensional plane, you know, we had a couple of different predictions. If it's a unique mechanism of tolerance, here is—we expect all the sensitive guys to be responding in the same way. But if each tolerant population responds in a different way than when we plot that response in this principal component space,
we should see each of these populations doing something differently.
In contrast, if it's a common mechanism of tolerance, we should see all of our tolerant populations doing something different than the sensitive ones,
but each converging on the same difference. So that's our very basic hypothesis. So what do the data say?
So here, we do our dose response, we see all these sensitive populations responding, you know, they're sort of clustering together.
Here's our first generation of tolerant guys raised in a lab. They're all converging on an indistinguishable response and
here it is in the second generation, the red ones. So, this is telling us that the tolerant guys are doing something very different from the sensitive guys, but they're doing the same different thing and that it's heritable, that the F1s [phonetic]. Therefore, one generation removed from few collected animals are doing the same thing as animals that are two generations removed from the field. OK. So that's at the common dose. We see this big difference between tolerant and sensitive guys, right?
Now, let’s compare what we see between effects match doses. When we crank up the dose of the tolerant guys still doing something really different or what’s going on there, OK?
So we looked at that. So here's a--our trajectories of dose response through a principal components space, you know, from our low dose, our control at the base of the arrow. And this is just summarizing two-dimensional space about like 3,000 genes are doing in terms of being turned on or turned off in their transcription. Low dose, middle dose, high dose, you see that all the sensitive populations are doing pretty much the same thing, yeah, across this dose range. Now, let's crank up the dose towards the magnitude and see what the tolerant guys are doing.
PCA: Interaction @ effects-matched dose

PC1 81%

PC2 8%

Sensitive pop.

BI
Flax
SH

Whitehead et al.,
They’re doing the exact same thing. It's just you require thousand times of dose to do it. And this is not because these animals are excluding the chemicals from their body. The body burdened of chemicals that you find in these embryos is just the same whether they are tolerant or sensitive animal. So here it is in a different way of looking at it.
The seed now, we're actually able to see individual genes. These are the genes that show a difference in their expression between populations. They're doing a--They have a different dose response here between populations and all of those differences between populations, partitions held by whether you're from a polluted site or clean site, whether you're tolerant or sensitive. So each column is a treatment average of the expression level of many replicate animals. Each row is a different gene and the color of the cell indicates whether the gene is upregulated yellow, or downregulated blue compared to the control which is black. So we could see a pile of genes that get turned on with dose in our sensitive populations, but are not turning on with that same dose in our tolerant populations.
But again, crank it up a couple of orders of magnitude and they converge on the exact same response, OK?
So, you know, this hypothesis is correct. So I know you're sitting there, going, "Come on, why can't you tell us something about what the genes are? We want to know something about the mechanism." OK, I'm getting there. OK, so we actually know a fair bit, at least some of the molecular biology of the initiating events in this cell for this class of chemical. So for things like PCBs and dioxins and PAHs,
when they get in the cell, they bind to the aryl hydrocarbon receptor, the cytosolic transcription factor. And once the ligand binds to that transcription factor— I guess I need to stand over here—
yet, it migrates to the nucleus, associates with some other proteins, binds to response elements in the genome and
Canonical dioxin/PCB signaling pathway

dioxins/PCBs → AHR (to nucleus)

HSP90

Cytoplasm
Canonical dioxin/PCB signaling pathway

dioxins/PCBs to nucleus

AHR

cytoplasm

HSP90

ARNT
turns on the expression of a bunch of genes. I'm not going to go into detail what these genes are. We've got phase one and phase two metabolism genes. We've got other genes in there too. A whole bunch of genes get activated by this ligand-activated aryl hydrocarbon receptor.
These are the genes that make up this figure.

[ Pause ]

So what this tells us is that this aryl hydrocarbon receptor signaling pathways profoundly desensitize in the tolerant animals. It's not an ACA [phonetic] mutation. We can activate this pathway. It just takes, you know, 100,000, you know, 100 to 1,000 to 10,000 times the dose to make that happen. So we get this highly correlated dose or something. It's awesome by a couple hours of magnitude, not only for the AHR genes of these upregulated genes, but also for a pile of genes that are downregulated too.
So it's--By studying these genes, they can tell us something about the mechanisms of action of these chemicals, but also some of the mechanism of tolerance. So a bunch of these downregulated genes were starting to try and put together this black box between activation of this AHR pathway and the developmental toxicity that you ultimately see. It's a bit of a black box right now. You know, people know a lot of those initiating events of those genes that are turned on, but between that and, you know, a screwed up heart is pretty much unknown. So we're trying to follow up with some of these studies.
It appears that a lot of these downregulated genes form a network that they interact with each other either in expression or protein-protein interactions. And this network is made up of genes that are involved in cardiomyopathies, you know, when you go back to literature and some angiogenesis genes. And these in particular, this downregulated genes are part of that last heat map are genes that are involved in the development of cardiac sarcomeres. So it looks like misregulation of this sarcomere development genes are part of this toxic phenotype of screwed up cardiovascular system. So we're trying to shed a little bit of light on this black box that occurs between initiation and developmental toxicity. But to be honest, my passion is more one of all differences between populations.
So, we've learned something about the mechanism of toxicity. I mean, we already know that each receptor is important.
Evolved pollution tolerance

**Mechanism of Toxicity:**

Activate AHR pathway
But we see that this—leads to altered expression of cardiomyocyte genes leading to essentially to cardiac sarcomere misassembly. And there's some other hypothesis that we have from the transcript dome. It's in what we know in the literature of the microbiology. These pathways, perhaps should--through uncoupling of nitric oxide synthase signaling,
and if that leads or at least is correlated with the development of cardiovascular toxicity. We've also something—
learned something about the mechanism of this really cool, wild evolutionary story where that
seems to evolve as global blockade of the AHR or signaling pathway,
but this blockade is leaky. When you crank up the dose high enough, then the IEL [phonetic] starts to turn on. And then that—When that turns on, when that AHR pathway turns on is when you start seeing toxicity in even these tolerant animals. OK. So we've got this cool evolutionary story of parallel adaptation. We asked if this is a convergent mechanism and the transcript domex [phonetic] tells us that, yes, it's a common mechanism that share across each of these populations that have drive tolerance.
But what's the genetic basis of this? One of the genes that are targets of natural selection here? It turns out that when you go to literature, you know, every study where somebody has discovered the genetic variant, that's responsible for variable sensitivity to dioxins or dioxin-like compounds. And you can find that in,
you know, rat strains that are resistant, you know, engineered rat strains or in different bird species or in tomcat [phonetic] living in a Hudson River, that each time that the genetic basis has been discovered,
it's a mutation in the aryl hydrocarbon receptor. This kind of make sense, you know, it's right the top of that signaling pathway that has something to do with the binding affinity if the chemical is reduced and so it's not activated. So you think that that would be a pretty good hypothesis for what's going on a killifish.
And so, first, we’ve looked at these colleagues of mine, Marcon [assumed spelling] and his group and others, have sequenced the AH receptor, sensible hypothesis.
And it just doesn't appear to be the case in killifish that there--that the variation, the genetic variation to aryl hydrocarbon receptor, there aren't any fixed differences between populations. There's not that much functional variation. There's some genetic variation, but doesn't appear to be functional. So it doesn't appear to be the AH receptor. It could still be some weirdness with AHR. There are multiple copies in killifish, but we just don't know yet.
I'm going to skip through this for now 'cause it's mostly hand waving. I've heard some mechanisms that we think might be going on. Well, OK, I'll go through it.
Pathway crosstalk interference?

ER → NOS → AHR → ARNT → HIF-1α

TTN → nitric oxide

MYH10

Cardiac remodeling

"Protective" expression in tolerant population only
Evolved pollution tolerance: NEXT?...

WHAT ARE THE GENOMIC VARIANTS THAT ENABLE RESISTANCE/RESILIENCE?

➔ TOP-DOWN COMPARATIVE POPULATION GENOME SEQUENCING APPROACH
So here's another study we did, similar to this, I showed you. This is—These are individuals here, each bar represents an individual's phenotype. The S, it means a sensitive population, T, is the tolerant population. The number is increasing log doses control and so on and so on. There are five replicate animals for treatment. And you could see that with increasing dose, this is an index of developmental abnormality, but the sensitive guys get pretty screwed up at high doses. But at those same high doses, the tolerant guys, is it starts to show up here in there, but they're far fewer screwed up animals, right?
And so there's some gene expression that's correlated with dose, right? And we see all these genes that are upregulated here that aren't regulated here. But this is a really weird pattern that we noticed, OK? So there are some dose responsive genes, yes, OK, that are different between populations.
But here's a really weird pattern that took awhile to discover, but I think it's important. Here is a gene, nitric oxide synthase that does not show a dose response in the tolerant population. It's not correlated with dose, but it is highly correlated with whether you're screwed up or not, right? So it's not regulated with dose in the sensitive population, but here it is in control, no, it's not, you know, base line levels. Here's low doses. It's cranked way up, way up, way up. Oh, you're screwed up, it's off. You're normal, it's on. Screwed up, off; normal, on; screwed up, off; normal, on; screwed up, off; normal on; screwed up off. So it's not dose--there's no dose treatment effect because the phenotype is too variable at high doses. But there is a really, really strong phenotype association. So they call this expression pattern predictive or protective, that if you have a NOS upregulated, you're OK. If it's not upregulated, you're kind of host or you're screwed up. Now, we use a different term in the paper, but you know what I mean. OK.
So we call that response protected when we—and there's NOS there. And we see a whole bunch of genes that are showing this protective response, that show this correlative response with NOS that's highly correlated with phenotype.
And it so happens that NOS crosstalks with pathways associated with the aryl hydrocarbon receptors signaling pathway. I'm not going to go into these into much detail. This is where the hand waving comes in. It doesn't appear in the AHR, but it's the AHR pathway. So something is--something's messing with the AHR signaling pathway and how that's responding. And how NOS interacts with AHR signaling pathway might be a one clue to what's going on. But we don't know. We're sort of waving our hands at this point.
So what are the genomic variants that enable this extreme phenotype, this resistance? We're taking the brute force approach now. We don't know, so we're going to look at everything. So we're taking this top-down comparative population genome sequencing approach. So we just--The first step in that was sequencing a reference genome for our animal.
So we finished that about a year ago. And for those of you that are interested in fish genomics, it's the best annotated fish genome out there. It kicks the butt of zebrafish and other sort of, you know, offense the zebrafish people in the room. But--And that's not bragging for myself. My colleagues, the informatics colleagues that I worked with have done an outstanding job at assembling and annotating the genome. So it's really great fish genome. It's not publicly available yet, but if you want access to it, just send me an e-mail and I'll give you access to it. So there's a reference genome out there, but like I said before, reference doesn't tell you much. We're using reference as an anchor to understand something about population variations.
So what we're going ahead doing now is resequencing multiple populations. So in resequencing genomes of about 400 animals, 50 animals from each of these tolerant and sensitive populations including the southern one—southern pair that I haven't talked about here. So 50 individuals from each of these geographic pairs of tolerant-sensitive populations, OK? And I should mention that these tolerant and sensitive—that these tolerant animals don't come from a single genetic background. I mentioned before that killifish are pretty genetically variable.
and as you go along sort of the East Coast from, you know, New England, you know, down to Florida, there's a lot of genetic variation that's partitioned along that latitudinal gradient. We've got this sort of northern clayed up here. So again, the color of these bars indicates sort of your genetic type. And we see this transition right around New Jersey, right around here when you go from one genetic type to another genetic type.
And it turns out that our pollution tolerant guys, these guys here are from this genetic background and these guys here are from this really different genetic background. So again, we've got a lot of complicating factors here, but we've got a lot of genetic variation that we're working with. And what we're interested in is what component of that genetic variation is important for sensitivity to environmental pollutants.
OK. So, a lot of sequencing, a fire hose of data about to turn on. Literally, this--it's a little bit scary what's coming down the pipe. But the idea is once we got all of this data, we map it all of the genome, we call where--we discover where variants are in the genome, where mutations are in the genome, among individuals and among populations. Then what you can do is you can use these population genetic algorithms to screen the genome to go across chromosomes and going to go across all the chromosomes to discover regions of the genome that are too variable to explain by chance alone. That are too variable, too different between these geographic pairs of populations to be explained by neutral processes.
Comparative population genome sequencing

→ Sequence a reference genome
→ Re-sequence multiple populations

- 4 pairs of TOLERANT and SENSITIVE populations
- 2 different genetic backgrounds (pleistocene history)
- 50 individuals per population
These aren't our data. I wish it was, and maybe our data will show this, sometimes seen. But this, it might--some of you might be familiar with the stickleback story. But stickleback, these fishes that have repeatedly evolved from a marine type into a fresh water type. And sort of they're--they've done these studies where they're screening the genomes, looking for regions of the genome that are adaptive to fresh water. And so they're a couple years ahead of us, but this is kind of what the data look like. As you scan across the chromosome, you're looking for these peaks of divergence between tolerant and sensitive populations that only selection or something that's not neutral can force this strong divergence between populations. So that's kind of what the analysis is going to look like down the road. And what's cool about the killifish story is we've got multiple populations that have each undergone a similar selection pressure. So not only do we expect there to be, you know, big peaks of divergence because it's been extreme selection pressure and these are lethal habitats if you're not--you don't have the right genetic type. So not only should there be a big fixed differences between populations, but if these regions of fixed differences overlap on the same region for each of these pairs that have independently evolve tolerance, then that's going to really scream out in the data. It's probably not going to be simple. I really hope it is. But it's--They're presumably other mutations that are important for life in polluted environments. But certainly, the ability to not see the chemicals and not to respond to them is important. And presumably, there are other pathogen stressors and other things in these habitats that's different than where you came from that you might have to evolve and responds to.
So, we take this top-down population genome sequencing approach, and then we, you know, screen the genome for these variants that are adaptive. So it's just taken out--it's somewhere in the genome. And we've got the tools now to screen that whole genome to look for that needle in the haystack where it's just a few years ago, we just didn't have the technology or the ability to do that. So it's a really exciting time to be in biology, in particular, my area of research because you're back to like the 18th century naturalist where, you know, there--genomes are these landscapes, they're these ecosystems. They're these--They have these emergent properties. And we've--we're now sort of getting these first glimpses, not into just a reference genome, but into population genomes. And how those vary in different habitats and how those genomes unfold during development and during environmental stress to enable these pretty cool phenotypes, right? So, we think that because this tolerance has evolved so quickly and involves a pretty dramatic phenotype, it's not a subtle two-fold difference in tolerance, right? It's a big difference and it's evolved quickly. You know, you do the math on that. These sites have been polluted since like the 1940s. These animals' generation time is a couple of years, sort of talking a few dozen generations that this has evolved, evolved very quickly. So it almost certainly has to be a genetic variant that preexisted in the populations at some appreciable frequency. And then when this population found itself in a polluted habitat, that just really increased in frequency very quickly. It almost certainly has to be a preexisting mutation rather than a new mutation that occurred after the onset of selection pressure. So, I will eat my shorts if it's otherwise, but I better not say that out loud.
So if--Then the question is, if these adaptive variants preexisted environmental pollution, what the hell are they doing there? What are pollution adaptive alleles, alleles that are adaptive to chemicals that don't belong in the environment, that word in the environment prior to us arriving disasters? What are they doing in natural populations? Your guess is as good as mine, as mine at this point. But once we discover the genetic variants that are responsible, then we can come up with better hypothesis and we'll target a hypothesis as to what they might be doing there as far as knowing what that gene does, the one that's responsible for mutation.
So, we've got all these great tools now, these great technologies to be able to generate genomic references. But in my opinion, that's kind of boring.
There's really no such thing as a reference genome, just like there's no such thing as a type animal, like you see it in museums. There aren't types and they're important as far as their historical importance. But no museum collects a type animal. They have a type animal, but then they collect as many animals as they can get their hands on to understand something about phenotypic variations in species. And that's where we're at with genomics now. You know, we were at the point a few years ago to maybe be able to scrounge enough millions of dollars to sequence our reference genome. Now, we could sequence our reference genome for about 20 to 50,000 grand depending on what quality you want. And then you can resequence genomes for about the same amount of money. So, the types of data that you can collect to discover things about this sort of unexplored landscape of the genome is really exciting right now.
So in some rate, you know, we use ecological and comparative genomics to understand something about how genomes interact with the environment, how genomes integrate cues from the environment in a functional way and ultimately how they’re shaped by the environment over evolutionary time.
And what does this do for us? Well, it helps contribute to our understanding of the mechanistic basis for organism-environment interactions at sort of a global level. You know, we're able to—you know, the term omics [phonetic] means holistic. You know, we're looking at the whole thing, the whole genome, the whole transcript and the whole—all the proteins, all the metabolites. So not only do we understand something about the mechanistic basis of organism-environment interactions, but it also contributes an
evolutionary and ecological contrast or natural history basis for understanding natural genomic variation. Now, we've known for a while. We've known since the 1960s that there's a lot of genetic variation within populations more than was thought realistic. So what's the importance of this variation? How much of that variation is important for life in difficult habitats?
And I would argue that evolutionary mutant models, these naturally evolved mutant models can help us identify the types of functional variants that are likely to exist in natural populations and be important as a substrate for evolutionary change.
So with that, thank you for your attention. There are people in my lab that do all the work and I get to talk about it. And lots of really important collaborators at the EPA and various universities, and we've been fortunate in our funding. So thank you for your attention.
Applause

Yes.

>> Right, nice talk. It reminds me some of the--between these old streets that branched on and accommodate what we're studying in fern and grass adaptations and [inaudible].

>> Yeah, yeah.

>> --which kind of really makes like that a question. You shot at this 63 pairs of population and then more, you know, nicely sort of converging adaptations, perfectly reasonable and patterned [phonetic] on natural selection and become--also in the genomics stuff, you've shown toward the end that it seemed like they wanted sort of a--like this is a similar genetic background or genomic background. So I guess what I was wondering is from those data or perhaps the data coming down the pipe, if you could--if you already have estimates of gene clone among this populations. So--

>> Yeah.

>> --you know, perhaps this, you know, mutations or series of mutations each occur independently in their population or maybe once in a row--
Yeah. Yeah, there are estimates of gene flow between these pairs of populations. Between those six northern populations, the FSTs range from zero to around--not big, like maybe 0.03. So there's--there is some genetic variation that's partitioned between populations, but there are times since shared ancestry is relatively recent. It's hard to put up molecular clock on that because we've been able to do--we've cycled--sequenced some nuclear mitochondrial genes for variety of purposes. It's a little bit hard to apply molecular clock to them 'cause there are no fossils to calibrate it and stuff. But yeah, these populations--what's nice is that genome-wide FSTs are pretty low. So that means that, it's sort of the best case scenario for trying to detect selection, right? Because there's--if FSTs are high, then there's lots of the genome. It's going to be variable, it's going to be hard to see that trees for the forest, right? So we're hopeful there. But certainly, between north and south, FSTs are much bigger, like 0.1 greater. So that's why it's important to have a--for the southern population, to have its own reference. Yeah. But when you look at--so FST is between geographic pairs are always lower than between phenotypic pairs. So the--Between each of the tolerant populations, the FST is always greater than between each of the tolerant populations, geographic normal sensitive pair. So that's the evidence that we're using to kind of say, well, it's happen to evolve--had to evolve quickly. And migration rates, I mean, it's a--it doesn't look very far in the map, but it's pretty far to a killifish from each of those sides. And over the course of a few dozen generations, just their migration rate is not--can't do that.
But maybe some fisherman and tourism, where there are big fish, making them-- taking them from polluted sites and planting them in another polluted sites. I mean, I don’t know, I hope not. Yeah?

>> So, I was interested in transcriptome of the resistant versus the sensitive populations were kind of responding in the same fashion. And so, and you kind of ruled out the AHR, aryl hydrocarbon receptor for the candidate, but how much of the regulatory portion of that gene could you analyze--

>> The whole gene.

>> The whole gene?

>> Yeah.

>> In terms--I mean, most of the variation that we gathered now is not in the gene itself--

>> Right.

>> --but in the regulatory portion.
>> Yes.

>> So in this, clearly, what's going on here, the regulatory of the mutation--

>> That AH receptor is--can be transcriptionally regulated by these chemicals. There is a feedback there. I mean some of the genes that repress AH receptor or that activate this, especially AH receptor are a part of that--those response elements in the genome. But you don't need upregulation of the AH receptor itself to cause that genomic response. And it appears that AH receptor, its expression level is not different between tolerant and sensitive populations. And it's not differentially regulated.

>> Even in the sense of--Even in the stressors part--

>> Right.

>> --you stress them, it still doesn't?
Right. There's not a difference between populations and how it's activated. Now, there were some allele frequency differences in the AH receptor coding sequence between sensitive and tolerant populations. And so they took those ones and kind of said there's--there aren't big--the big differences that you'd expect in frequency. But there are some differences in frequencies. So let's take the type that's most abundant in polluted sites and the one that's most abundant in sensitive sites and see if their ability to activate the genome is different, and they weren't. So it's not a close book on it, but yeah, each sort of candidate approach for the AH receptor has kind of come up. But it doesn't--yeah, nothing convincing there yet. But what's--you know, fisher--weird fish have, you know, there are a couple of whole genome duplications. So various fish have lost various paralogs and killifish seem to have retained up to four paralogs of the AH receptor. And not all of them have been sequenced. We know from at least zebrafish that some of those AH receptors are associated with toxicity and some of them are not. So there might be something weird going on there, that's still AH receptor. We just haven't sequenced the right one yet. Yeah?

I was curious. For the southern population, it seemed to me that [inaudible] genotype. Have you done any preliminary, those interesting in some those?

Yeah.

And also is there a difference in the dioxin toxicity of these sites?
Yeah. What's weird is—so, here's something weird, OK? So, the northern sites are predominantly PCBs. There's other crap in the sediment, but it's mostly PCBs. There are some PAHs and some metals and stuff. The southern one, the Elizabeth River one is a creosote site. It's mostly PAHs, some PCBs and dioxin. So the northern ones, the guys that evolved in PCB habitats, they are resistant to PCBs and they are also pretty resistant to PAHs, which kind of makes sense because both of those classic chemicals act to the AH receptor. The southern guys that evolved at the PAH site, they're pretty tolerant to PAHs. They're wicked tolerant to PCBs. You can't kill them. You know, we dialed up even another order magnitude like, I don't know, even another order magnitude behind the highest dose that were starting to kill the tolerant northern ones, and the southern guys weren't blinking. So there are differences, certainly, differences between habitats. They're all united in the fact that they're polluted with things that activate the AH receptor. But there are different classes of chemical that are predominant in each of those sites, their different genetic backgrounds. There are some differences in phenotypes as far as PAH evolved, killifish being more tolerant to PCBs than PCB evolved killifish. So there's some weirdness going on there that we hope— I think that it's going to be a really rich dataset when we get both of those. We do have the experiments. We do have some data from PCB-exposed southern guys and it's just not ready to show yet, and where it's kind of in the middle of the analysis. And yes, so in a couple months, get back to me.

All right, there's no other question. But hey, come on down.

Cool. Thanks.