1. BACKGROUND AND GOALS

Introduction. Eukaryotes depend on gene products encoded in both the nuclear and mitochondrial (mt) genomes in order to maintain proper cellular energetics. Mutations in human mtDNA are known to cause neuromuscular and neurodegenerative "mitochondriopathies" and play a role in aging and complex human diseases such as cancer, diabetes, and heart disease[1-7]. However, sources of mt mutations and their physiological consequences remain understudied and controversial. A major obstacle in this field is the focus on vertebrate and particularly mammalian study species. Such species have uniformly high mt mutation rates and their mtDNA contains only a small fraction of the structural and functional diversity found across eukaryotes[8-17]. Invertebrates, plants, and other eukaryotes show wide variety in mt mutation rates. For example, some lineages have very low mt mutation rates[18,19] and others show wide variation among closely related species[11,20-22]. As in humans, mt variation can have large phenotypic effects in diverse eukaryotes[23]. Our lack of understanding of the underlying causes of mt mutations across eukaryotes and their physiological consequences prevents our ability to understand when mt mutations are predicted to affect human health and limits the possibility for mt interventions to rescue health or alter aging.

We propose to overcome this constraint in our knowledge of mt mutation by undertaking coordinated studies linking mt mutations, mt physiology, and organismal phenotypes in diverse eukaryotes, which will more fully uncover the range of mt mutational causes and effects. Examining wild, non-inbred organisms experiencing natural stress is currently a "shortcoming of many foundational research programs in cell and molecular biology"[24]. Evolution has also provided a natural laboratory to uncover the bounds of mt processes in aging and disease. Linking mt mutations and their physiological consequences in non-model species can build a foundation for translational therapies for disease and aging[2,4,5,25-31]. Therefore, exploring mt mutations in diverse eukaryotes may extend our understanding of the possibilities for mt mediated disorders in humans.

My lab is using a comparative, integrative approach to understand the causes and consequences of mt mutations through employing wet lab experiments and bioinformatic tools in diverse systems. This puts us in a unique position to make connections between mt mutations, stress, physiology, health, and aging.

To overcome the current bottlenecks in our understanding of mt mutations and their physiological effects we are addressing three challenges:

1) **What are the sources of mt mutations?** Mt mutations have been characterized in a limited diversity of eukaryotes. We are using high-fidelity sequencing approaches to examine mt mutations across several systems (Table 1), revealing when oxidative damage vs. replication errors cause mutations.
2) **How does mt physiology connect to mt mutations?** Few studies have measured mt mutations simultaneously with mt physiology. We are quantifying high-resolution respirometry, ROS production, and metabolic rate (MR) in species with variable mt mutation profiles and rates.
3) **How have mt mutations and physiology affected the evolution of aging?** It is unclear what roles mt mutations play in aging. We will characterize mitonuclear protein balance, mt mutations, and mt physiology in two systems where lifespan varies drastically among closely related species.

2. RECENT PROGRESS AND FUTURE RESEARCH

Much of my recent research has centered on molecular evolution and physiology of mitochondria. Below I focus the three challenges outlined above, preliminary and recently published results from my lab that address them, and our future plans to expand these studies.

Challenge 1: Identify the causes of mt mutations across eukaryotes

One ongoing debate in mt biology concerns the relative importance of oxidative damage vs. replication errors as the major causes of mt mutations. Theories such as the mt free radical theory of aging posit that oxidative damage leads to mt mutations and age-related pathologies[1,30,32-39]. Examining mtDNA mutations in real time can reveal the relative importance of these sources, but has proven difficult due to the rarity of these mutations (e.g., ~10^{-9} per site per generation) and high error rates of sequencing technologies (e.g., ~10^{-5} for standard Illumina), resulting in a high "noise" threshold[40,41]. Recently, advanced high-fidelity sequencing (error rates of ~10^{-11}) has mitigated this[42], and has been used to confirm that mt mutations increase with aging[43]. However, these studies found replication errors, not oxidative damage, as the dominant source of mt mutations[44-46]. Specifically, they found an abundance of C -> T transitions indicative of replication errors via deamination reactions, not G -> T transversions (8-oxoguanine or 8-oxoG variants) indicative of oxidative damage. Replication errors also appear responsible for mt mutations in neurodegenerative diseases, cancer, and other diseases[47-49]. However, reactive oxygen species (ROS) also increase with age and in disease[50].

Table 1. Model systems to be used in the proposed research

Species	Group	Notes
H. pluvialis	Algae	High carotenoid content; stress-tolerant
Eurycea spp.	Amphibians	Variable/elevated life spans; low MR
Silene spp.	Angiosperms	Variable mutation rates/genome sizes
B. atricristatus	Aves	High MR; speciation linked to mtDNA?
L. teres	Bivalves	DUI; highly divergent M vs. F mtDNA
N. vectensis	Cnidarians	Low mt substitution rates
T. dohrnii	Cnidarians	Low mt rates; biological immortality
H. rubra	Crustaceans	Hi mt substitution rates; red; stress-tolerant
S. cerevisiae	Fungi	Facultative anaerobe; stress-tolerant
B. dendrobatidis	Fungi	Parasitic lifestyle; stress-tolerant
M. musculus	Mammals	C -> T mt mutations; models available
C. elegans	Nematodes	G -> T mt mutations; models available
X. malinche	Teleosts	Speciation linked to mtDNA
Sebastes spp.	Teleosts	Variable/elevated life spans

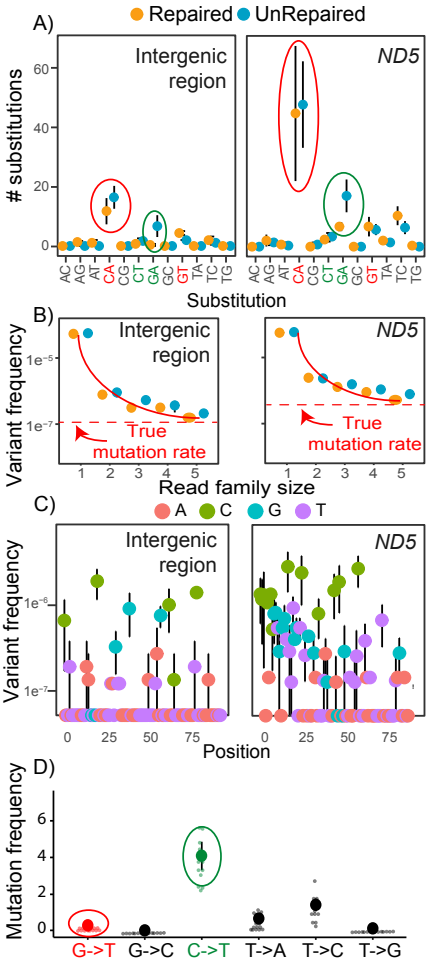

Fig. 1. Mt mutations in *C. elegans* (A-C) and the PolG mt mutator mouse (D). Mutations from replicative errors and oxidative damage are highlighted in green and red, respectively. Two regions of the mt genome were characterized in A-C (*n* = 3), while D) is based on the entire mt genome (*n* = 15).

Our understanding of the sources of mt mutations and their roles in disease is shaped almost entirely from mammalian studies[51]. The mtDNA mutator mouse ("PolG") has proven important in such studies. PolG mice have deficient exonuclease activity, leading to reduced proofreading, increased mt mutations, and premature aging[52]. However, it is still unclear whether increased oxidative damage characterizes PolG mice, or if interventions such as exercise and caloric restriction alleviate PolG phenotypes[53-59]. Moreover, Sod2-knockout mouse models with increased ROS have myriad mt deficiencies, but do not have increased mt mutations[50]. The few non-mammalian studies further complicate matters. For example, mt mutations do not appear to increase with aging in *Drosophila*[45]. Applying high-fidelity sequencing techniques in a broader sampling of eukaryotes may reveal taxa where oxidative damage is the primary source of mt mutations, offering new models for mt disease.

Preliminary results. We have begun applying high-fidelity sequencing to determine if replication errors are a universal cause of mt mutations. Preliminary data using maximum depth sequencing (MDS)[60] on the model nematode *C. elegans* revealed that 8-oxoG mutations are by far the most abundant type of mutation in two regions of the mt genome, while C -> T transitions are largely absent after correcting for DNA damage introduced during DNA extraction (Fig. 1A). MDS also allowed us to quantify variation in the true mt mutation rate within the mt genome (Fig. 1B). This may allow us to further pinpoint the causes of mt mutations, as replication errors may be more common near origins of replication, while oxidative damage should affect all regions of the mt genome similarly. MDS targets a small (~100 bp) region of the genome for deep-sequencing and analysis, also allowing us to estimate rates of all mutations at every position in the region (Fig. 1C). We are currently combining MDS results in *C. elegans* with another approach, duplex sequencing. Although it has lower resolution, duplex sequencing is able to identify sequencing errors vs. true mutations bioinformatically[41,61,62], and we are applying it across the entire mt genome in *C. elegans*. We have also combined sequencing approaches with a new mt genome enrichment method[54] to characterize mutation rates and spectra in the PolG mouse, which should show mutations due to replication errors and serve as a control for our novel findings in *C. elegans*. As expected, C -> T transitions dominate the PolG mt mutational landscape (Fig. 1D). Our preliminary results in *C. elegans* offer new insights compared to previous work using mutation accumulation lines, including estimating a lower overall mt mutation rate[63,64]. Our work in the PolG mouse improves on previous, under-replicated studies, supporting low mutation rates in the D-loop and a role for selection on mt mutations in the germline[52,54,65].

These results highlight our ability to characterize mt mutations in diverse systems and suggest sources of mt mutations can vary among species.

Future research. *Developing the mt mutation atlas.* We will use high-fidelity sequencing to characterize mt mutation rates and spectra in diverse eukaryotes (Table 1). Duplex sequencing will be used to first survey the entire mt genome at a relatively low resolution. Resulting data will be used to target specific regions of interest at a higher resolution with MDS (3-6 regions per species, including using a set of "universal" MDS primers to target the same 3 regions in all species)[60]. The systems we have chosen to begin this survey (Table 1) have properties suggesting causes and rates of mt mutations may differ among them, including: variable mt substitution rates, high tolerance of environmental stress and ROS, variable mt inheritance, and variable MR. Eventually, we hope to generate a 'mt mutation atlas' for eukaryotes. Such a resource would allow researchers to find common characteristics of taxa with ROS-mediated mutations, target specific taxa as models for human mt disease, and identify the most common sources and regions for mt mutations.

Characterize mt mutation rates and spectra under different environments and in mouse models. Although *C. elegans* appears to normally accumulate mt mutations via oxidative damage (Fig. 1A), other taxa, including humans, may only accumulate G -> T mt mutations under environmental stress or elevated ROS levels. Moreover, it is important to mechanistically link mutation sources to the types of mutations detected via sequencing. Therefore, we will use initial screens of mutation spectra to identify 4 taxa that have different spectra and sources of mt mutations (e.g., low vs. high rates and caused by replication vs. ROS). For these taxa we will examine how mt mutation spectra changes under different environmental stressors and exposure to the pro-oxidant *tert*-butyl hydroperoxide (t-BHP)[66,67]. The exact stressors and exposures will vary depending on the taxa chosen, but temperature will likely be investigated given our previous studies (see below) and the role of temperature across eukaryotes in mediating ROS and mt function[68-86]. Finally, we will investigate rates and spectra of mt mutations in two mouse models under normal and pro-oxidant environments: the PolG mt mutator mouse and the Sod2-knockout mouse. A previous study did not show increased mt mutations in the Sod2-knockout mouse[50], but did not investigate mutational spectra as we propose here.

Challenge 2. Link mt mutations with mt physiology and organismal health

Many studies of mt mutations do not simultaneously examine mt physiology. Linking mt mutations with physiology and performance is critical to determine the causes and consequences of mt mutations. Individuals within a species can show large variation in the proportion of mutations caused by oxidative stress (e.g., error bars in Fig. 1A) and overall mutation rates[87]. Examining ROS production, mt respiration profiles, and the activities of relevant enzymes alongside mt mutation rates/spectra across species and environments may reveal why oxidative stress causes mt mutations in some systems, but not others. It is also important to examine organismal performance alongside mt mutations and mt physiology, because mt mutations and altered mt physiology may result in organismal dysfunction in some systems but be ameliorated in others.

Preliminary results. We have begun examining the consequences of mt mutations for molecular evolution, mt function, and organismal phenotypes. The angiosperm genus *Silene* is a particular useful model for investigating the consequences of mt mutations in a natural context because closely related species show extreme variation in mt substitution rates[20,88], likely reflective of mt mutation rates: some species have slow mt rates typical of other model angiosperms (e.g., *Arabidopsis*) while others have undergone recent, dramatic accelerations, making their rates similar to those in mammalian mt genomes. I previously showed that fast mt evolution is coupled with fast evolution in mt-interacting nuclear genes (N-mt genes), providing some of the most compelling evidence for mitonuclear coevolution in any system[10,89-103]. In other words, mt mutations may select for complementary nuclear changes.

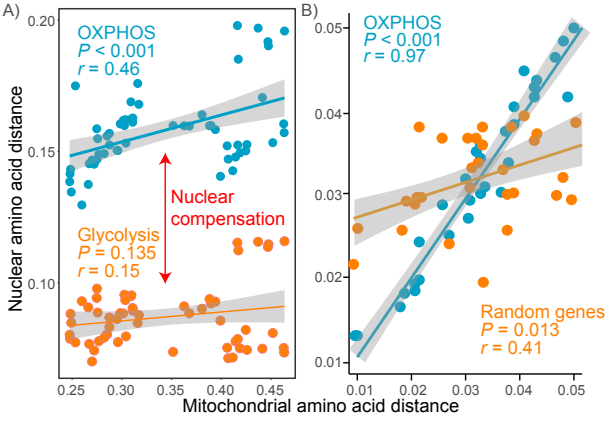

Fig. 2. Mitonuclear coevolution revealed through evolutionary rate correlation between mt and N-mt genes (blue), but not control nuclear genes (orange). A) mammals and B) bivalves. Higher rates in N-mt genes of mammals supports nuclear compensation, which is not observed in bivalves. Each point represents a species.

We are beginning to extend the rationale used with *Silene* to other eukaryotic systems to examine variation in mitonuclear coevolution. Under mitonuclear coevolution, rates of mt evolution should be correlated with rates of evolution in N-mt genes across many distantly related species showing a continuum of mt evolution rates. Such evolutionary rate correlation (ERC) studies are only beginning to be explored (e.g., mitonuclear ERC was found in arthropods[103]). We have recently applied this technique to mammals (Fig. 2A) and bivalve mollusks (Fig. 2B –

Piccinini et al., under review at *Molecular Biology and Evolution*). In both groups ERCs are strong between mt and N-mt genes but are diminished or absent between mt genes and non-interacting nuclear-encoded genes. Bivalves are an interesting model for mt studies because several lineages have independently acquired "doubly uniparental inheritance" (DUI) of mitochondria[104-107], which is in contrast to the strictly uniparental inheritance found in most eukaryotes.

Many assume mitonuclear coevolution takes a specific form: deleterious mt mutations accumulate due to inefficient selection in the non-recombining mt genome, and compensatory nuclear changes act to maintain mt function[10,99,100,108-112]. However, studies specifically investigating nuclear compensation in contrast to other forms of mitonuclear coevolution are lacking[99], and our results from bivalves are not consistent with nuclear compensation (i.e., N-mt genes do not show overall increased rates compared to other nuclear genes, as they do in mammals – Fig. 2). We are currently using structural stability modeling and phylogenetic information to more specifically test for nuclear compensation in mammals. *Our work untangling mitonuclear coevolution provides a molecular evolution complement to understanding the physiological consequences of mt mutations.*

It is unknown how acceleration in mt mutation rates may affect mt function. To address this, I first developed a methodology for high-resolution respirometry in isolated plant mitochondria (Fig. 3A). Applying this protocol to *Silene* revealed that "fast" species have relatively normal levels of maximum mt respiration, despite accumulating many amino acid substitutions (Fig. 3C, right panel)[113]. My lab then made detailed, direct comparisons of mt function in fast vs. slow species[114]. We found that "normal" levels of mt respiration in fast species are maintained by increased reliance on the nuclear-encoded alternative oxidase (AOX) and accessory NADHs (Fig. 3C), both of which mitigate ROS in stress responses[115-119]. We tentatively hypothesize that mt mutation accumulation in fast species caused elevated ROS levels, which triggered increased reliance on AOX and accessory NADHs. *This result suggests mt mutations, changes in mt physiology, and ROS production may be linked in some lineages or environmental conditions.*

We are also investigating how environmental stress affects mt function. Thermal variation drives mt physiology because mt respiration increases passively with increased temperatures[120-123] and mitochondria can generate heat through OXPHOS uncoupling[124-127]. Not surprisingly, warm vs. cold-adapted mt genomes have been identified across animals[70,72,128-133]. We recently explored mt function in mayfly larvae from cold, high elevation streams vs. thermally variable, low elevation streams[77]. We found that maximum mt respiration (along with other characteristics) was more sensitive to warm acclimation in the cold-adapted populations (Fig. 3D). We also developed a new method to quantify thermal acclimation in respiration and found that acclimation only partially compensates for passive changes with temperature across animals[134]. *These findings point to temperature as a major moderator of mt physiology and possibly mt mutations.*

In addition to linking mt mutations to molecular evolution and mt physiology, we have also begun establishing connections to organismal metrics of health. Recently, we examined whether responses to temperature are coordinated across mt enzymes, mt respiration, and MR. We found that across 235 pairwise comparisons of thermal responses between levels reported in the literature, there was significant discordance, sometimes with different levels of biological organization showing completely opposite responses to the same range of temperatures (Fig. 4)[135]. As another example of examining responses across biological levels, we recently characterized responses to salinity in related species of shrimp from extreme habitats[136,137]. We found disparate responses to salinity when examining gene expression, protein abundance, tissue remodeling,

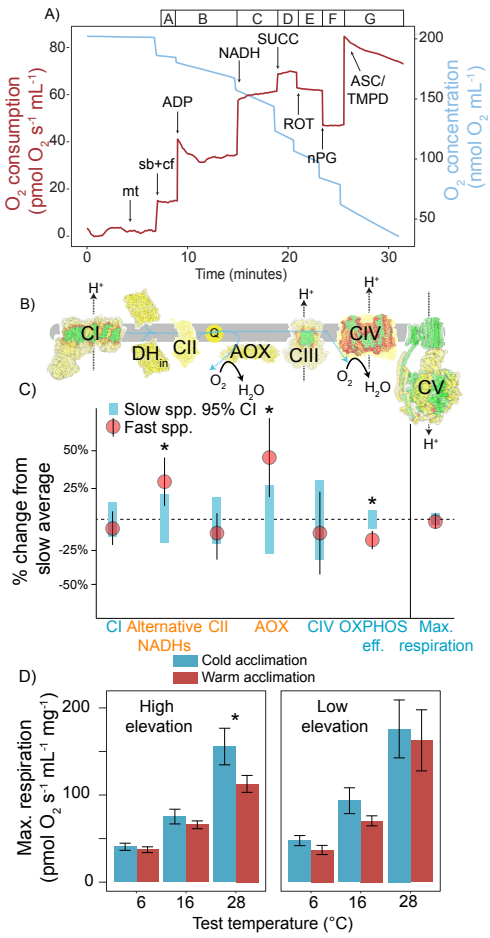

Fig. 3. A) Protocol for high-resolution mt respirometry in isolated plant mitochondria[113], B) OXPHOS machinery in plants (with mt-encoded residues in green, nuclear ones in yellow, and contact ones in red), C) Fast *Silene* spp. OXPHOS flux control parameters (red points showing 95% CIs) in relation to slow spp. (blue bars showing 95% CI) – blue text indicates complexes with mt-encoded components, orange are entirely nuc-encoded (n = 50)[114], D) Cold-adapted high elevation populations show more thermally sensitive mt function (interaction $P < 0.001$, n = 10)[77].

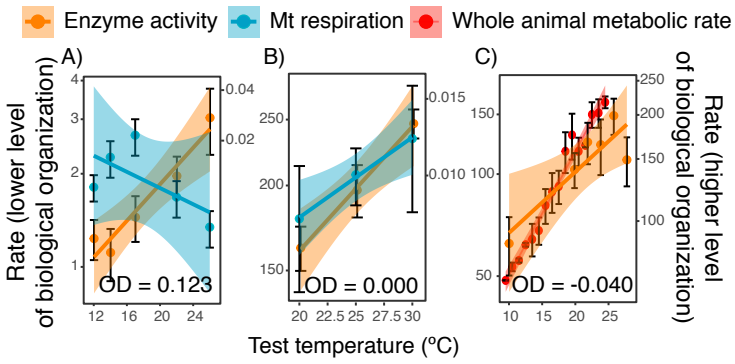

Fig. 4. Mt responses to temperature across levels of biological organization in A) *Oreochromis*[173], B) *Crocodylus*[174], and C) *Limnodynastes*[175]. These 3 responses show the range of organizational disagreement (OD) in our dataset[135]: A) the most discordant OD, B) the most congruent OD, and C) average OD.

and survivorship across species[138]. One species, *Halocaridina rubra*, is of particular interest to us due to its high mt divergence among populations, presumably reflecting elevated mt mutation rates[139-141]. We recently described differences in red, carotenoid-based coloration among *H. rubra* populations[142]. Carotenoids are hypothesized to play roles in ROS responses[33,143,144] and are associated with increased immune response and perceived health in humans[145-147]. We are currently exploring how mt mutations, environmental stress, ROS, and red coloration are connected in this species. *Overall, this work highlights that mt mutations must be examined at multiple levels of biological organization to accurately understand their consequences.*

Future research. *Identify which aspects of mt physiology predict mt mutations.* Based on the mt mutation parameters quantified in Challenge 1, we will select a subset of species that have variable sources and rates of mt mutation, likely the same systems that will be investigated under different environments in Challenge 1. Based on preliminary evidence (e.g., Fig. 1), tentatively appealing systems include: *C. elegans* (G -> T mutations), *Lampsilis teres* (low metabolic rates and slow mt substation rates), *H. rubra* (high metabolic rates and mt substation rates), and *Haematococcus pluvialis* (slow mt substitution rates). For these species, ROS production in isolated mitochondria will be quantified during distinct mt respiratory states using an O2K FluoRespiometer[148], which circumvents some common technical pitfalls[149]. Rates of ROS production, maximal mt respiration, the OXPHOS efficiency flux control factor[77,113,150], and other OXPHOS characteristics will be correlated with mt mutation rates and rates of 8-oxoG mutations across and within species. A positive correlation between ROS and 8-oxoG mutations would be consistent with oxidative damage directly causing mt mutations. A positive correlation with C -> T mutations might suggest ROS indirectly mutates mtDNA through replication errors, supporting recent findings that oxidative stress reduces mt polymerase fidelity, but does not directly cause mt mutations[151]. Finally, we will use RNAi to knockdown mt superoxide dismutase activity in these four systems, which should increase ROS. We will then characterize mt mutations (as in Challenge 1) and the mt physiology parameters outlined here to provide a mechanistic link between ROS and mt mutations in different systems.

Determine fitness and health effects of mt mutations. MR is a fundamental, easily measured property of animals linking their physiology to fitness, health, and ecology[152-155]. We will therefore measure resting metabolic rates (RMR) in most of the animal systems in Table 1. In two systems (*H. rubra* and *Xiphophorus malinche*), we will also measure maximum MR allowing us to calculate aerobic scope (MMR - RMR). Aerobic scope represents the excess energy available to the organism for performance and is a general metric of health[78,131,155-161]. If mt mutations lead to decreased performance, we predict animals with high mt mutation rates (or high 8-oxoG rates), increased ROS levels, and diminished mt respiratory capacity/efficiency should also show elevated RMR and/or decreased aerobic scope.

Challenge 3. Characterize mt dynamics in aging

A primary goal of my research program is to test the hypothesis that mitochondria are a nexus for the evolution of key eukaryotic innovations[99,100,108,162-165]. Many ideas center on the role of mitochondria in aging[36,166]. However, most biomedical studies of mt mutations and aging focus on mouse or other mammalian models. Natural selection has produced species with incredibly divergent lifespans, but few studies have characterized mt mutations or mt function in these systems. Examining these processes could provide useful insights into aging mechanisms or possible therapies in humans.

Preliminary results. Work based on experiments in *C. elegans* and mammalian cells described an imbalance in mitonuclear protein stoichiometry as a conserved mechanism of longevity that acts via the mt unfolded protein response[167]. Multiple methods used to increase lifespan in these systems resulted in the same imbalance: when lifespan is increased mt protein abundance is reduced. PolG mice also show altered protein levels that are somewhat normalized with exercise[59]. We have previously found that mt genes are always

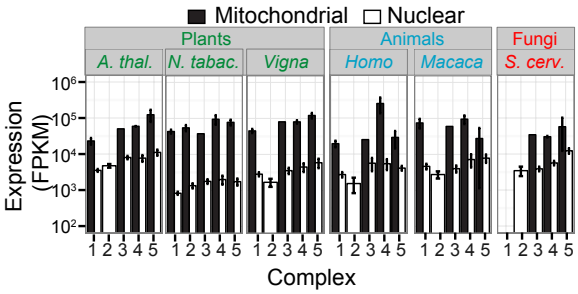

Fig. 5. Mt OXPHOS genes are expressed at higher levels than their nuclear-encoded counterparts across eukaryotes[10].

expressed higher than their nuclear counterparts across plants, animals, and fungi (Fig. 5)[10]. In other words, the expression of a gene is determined by its genomic identity, not its function. However, others have shown complex relationships between mitonuclear RNA and protein levels[168].

Future research. *Does mitonuclear protein balance predict lifespan?* We will use transcriptomics and proteomics to characterize mitonuclear RNA and protein stoichiometry in two lineages containing species with divergent lifespans: rockfishes[169] and cave salamanders[170]. Species in these lineages vary in lifespan by over an order of magnitude and include some extremely long-lived species. This will allow us to test whether the mitonuclear imbalance identified in lab systems has been selected on in natural populations. We will also examine expression of N-mt ribosomal genes across fishes using publicly available RNA-seq data to ask whether maximum lifespan (gathered using the fishbase database) is negatively correlated with gene expression across species, as it is across strains of mice and *C. elegans*[167].

Characterize mt mutations and function in closely related lineages with divergent lifespans. To determine if ROS production, mt mutation rates, and mt physiology vary with lifespan, we will examine these processes in the rockfish and cave salamander systems. Previous work in rockfishes found a strong correlation between mt substitution rates and longevity[169], but did not examine mt mutation rates, spectra, or mt physiology. Theories of aging that implicate mt mutations predict that longer lived species would accumulate fewer mt mutations, show lower ROS levels, and lower mt respiration. We will test these predictions using the approaches outlined above. We will also examine the mechanistic link between mt mutation rates/spectra and mitonuclear protein imbalance by quantifying mt mutations in Mrps5 knockdown *C. elegans* (using RNAi and methods described in Challenge 1). We predict lower mt mutation rates and fewer 8-oxoG mutations with Mrps5 knockdown.

3. INNOVATION, QUALIFICATIONS, AND POTENTIAL PITFALLS

Innovation. The proposed research uses a combination of physiological, biochemical, and bioinformatics approaches. This research will be carried out across a diverse range of eukaryotes (Table 1), which are seldom considered in these types of studies. In recent work, we have begun to link mt mutations to molecular evolution, mt physiology, and organismal performance. Future research will further explore these themes to address the three challenges outlined above.

Qualifications. As a new investigator in this competitive field, I face several challenges. I am broadly trained as an evolutionary biologist and am likely unfamiliar to molecular bio-scientists examining aging, health, and mt function. However, my lab is well-positioned to undertake the proposed research. First, we have expertise in mt physiology and biochemistry, and have developed methods to simultaneously quantify respiration, ROS production, and membrane potential in isolated mitochondria across eukaryotes. Second, we use sequencing-based approaches to detect and characterize mt mutations, including generating new bioinformatics pipelines to analyze such data. Third, we are leaders in mitonuclear ecology[99,100,108,171,172], which links mt processes to broad themes in evolution. Most importantly, I quickly form collaborations with relevant experts. We will reach out to established specialists at my home institution and elsewhere to help guide us in these efforts.

Potential difficulties and alternative approaches. A strength of the proposed work is its use of diverse eukaryotic systems (Table 1), which may seem ambitious. While it is necessary to explore a broad swath of eukaryotic diversity to populate a “mt mutation atlas”, applying deep-sequencing and mt physiology methods may prove challenging in these systems. In this case, we will rely on more traditional approaches: for estimating mutation rates/spectra we will propagate mutation-accumulation lines and sequence them with more conventional next-gen methods. This will only be feasible for taxa with short generation times. For other taxa, we will use the 4-fold degenerate synonymous substitution rate compared to a close relative as a proxy for the mt mutation rate. For mt physiology, we will use established enzyme assays for mt protein complexes and horseradish peroxidase-based assays for ROS production.

One theoretical shortcoming of the proposed work is the lack of focus on human cell lines or mammalian systems, making direct inferences on human health difficult. However, our work is envisioned as a foundation for later studies where patterns and mechanisms uncovered here can be evaluated in dedicated models of human disease and aging. The proposed work will reveal the *potential* for mt mutations to shape mt function and organismal traits, which is of interest in human disease and aging.