Introduction

Temperature is a fundamental environmental force shaping species abundance and distributions through its effects on biochemical reaction rates, metabolism, activity, and reproduction (Helmuth et al. 2005, Kingsolver 2009). Current models of global climate change project increases in both mean and variance in temperature (Diffenbaugh & Field 2013), which has the potential to drive species to extinction (Cahill et al. 2013; Urban et al. 2015) by pushing their performance towards lethal limits unless they can successfully acclimate, shift their geographic ranges, or adapt to chronic thermal stressors (Huey et al. 2012, Sunday et al. 2014; Deutsch et al. 2008). **The aim of my dissertation is to identify molecular mechanisms of acclimation and adaptation to the thermal environment in a clade of ecologically important North American forest ants.**

Research on mechanisms of temperature adaptation has focused on induced and constitutive levels of well-characterized heat shock proteins (Hsps; Lindquist & Craig 1988; Sørensen et al. 2003) that refold damaged proteins, part of the highly evolutionarily conserved heat shock response (Krebs & Feder 1997; Feder & Hofmann 1999; Garbuz et al. 2003; Jensen et al. 2009; Carmel et al. 2011; Calabria et al. 2012). In fruit flies, higher thermally tolerant individuals and species produce more Hsps (Kreb & Feder 1997; Feder & Hofmann 1999). In a number of other studies, however, Hsp production does not fully explain variation in upper thermal limits (Jensen et al. 2009; Rezende et al. 2010; Terblanche et al. 2011; Calabria et al. 2012), suggesting that additional mechanisms are likely to contribute to conferring heat resistance (Somero 2012).

Potential alternative defenses against heat stress include altering membrane fluidity (Sinensky 1973, Cooper et al. 2014), production of thermally stable metabolites (Singer and Lindquist 1998, Rivero et al. 2004), and shifts in the thermal properties of proteins (Somero 2010; Lockwood & Somero 2012). Proteins can maintain activity under elevated temperatures if they are more thermally stable, which may be brought about through the induction of thermally stable isoforms and/or evolutionary changes in amino acid sequences. Although such changes resist heat denaturation, however, protein stiffness comes at a cost of lower catalytic activity (Fields 2001; Somero 2004), and stability-conferring mutations not eliminated by purifying selection may accumulate at a relatively slow rate (Yang et al. 2000, Nei 2005). Thus, if proteome stability is the primary determinant of thermal tolerance at the whole-organism level, purifying selection on protein function and/or low plasticity in expressing different isoforms may constrain the extent and speed of temperature adaptation (Zeldovich et al. 2007; Somero 2010), and could explain the relatively narrow range of upper thermal limits in comparison to lower thermal limits demonstrated in previous studies (Sunday et al. 2011; Hoffmann et al. 2013).

Until recently, measuring proteomic stability was technically challenging (Savitski et al. 2014), making it exceedingly difficult to assess the importance of shifts in protein structure for thermal adaptation. However, new techniques in liquid chromatography-mass spectrometry (LC-MS/MS) can now quantify the stabilities of each identifiable protein in a biological mixture (Wühr et al. 2012), permitting characterization of protein stability at the level of individual proteins, molecular pathways, and up to the entire proteome. When applied within a comparative framework, this technique can potentially identify whether changes in protein stability have played an important role in adaptation to thermally stressful environments, as well as the key molecules and

pathways most responsive to thermal selection pressures. **I propose to extend my dissertation work in this exciting new direction by testing whether increases in proteome stability are associated with evolutionary increases in upper thermal limits across the North American clade of *Aphaenogaster*.**

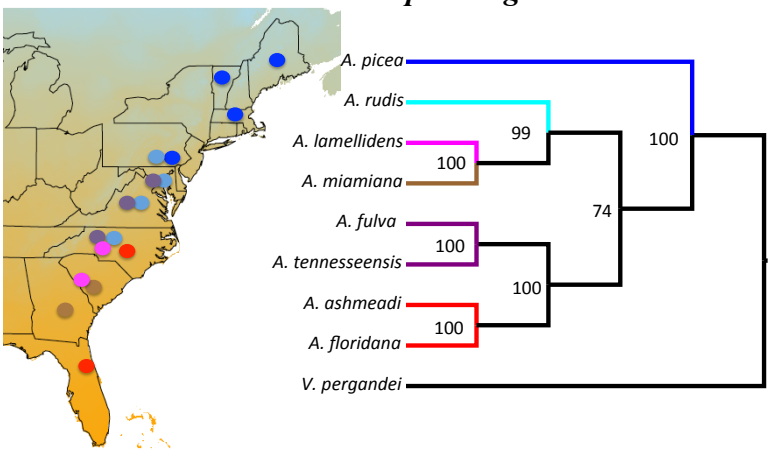

Figure 1. The geographical and phylogenetic structuring of common woodland ants (*Aphaenogaster*) along the Eastern United States. Diversification patterns (simplified phylogeny; *right*) are mapped onto the Eastern United States (*left*), where the different colors on the phylogeny and map correspond to sampling sites where colonies of each species were collected. To reconstruct phylogenetic relationships, we first assembled a 12,362 SNP matrix generated from double restriction enzyme assisted digestion-sequencing (ddRAD-seq; Peterson et al. 2012). This matrix was analyzed in a maximum likelihood framework in RAxML 8 (Stamatakis 2014) and group support was evaluated with 100 fast bootstrap replicates.

Dissertation Work to Date:

I have characterized intra- and inter-specific variation in upper thermal limits, quantified as the temperature at which the individual can no longer right itself (CT_{max}, Lutterschmidt & Hutchison, 1997), in *Aphaenogaster*, a group of forest ants occupying thermally diverse habitats from Florida to Maine (Fig. 1; Demarco & Cognato 2015). The speed at which increasing temperature is applied, or ramping speed, determines the extent to which heat shock protein production can be up-regulated in response to accumulating protein denaturation (Sørensen et al. 2013). I quantified CT_{max} along with Hsp mRNA expression at both a fast (1°C/min) and slow (0.1 °C/min) ramp speed; although the heat shock response is fully activated at the slow speed, heating under the fast ramp is too rapid for significant up-regulation to occur and only constitutively expressed, basal levels of heat shock protein would be expected to contribute to thermal tolerance (Fig. 2).

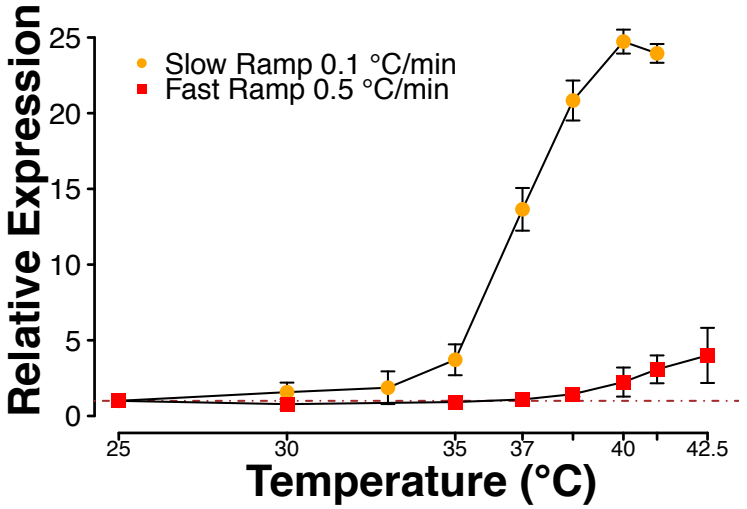

Figure 2. The induction of *hsp70* under a fast (red) and slow ramp(orange) reflecting the reliance on constitutive or induced defenses. Relative expression was calculated using the ΔΔCT method (Livak & Schmittgen 2001) using 18s rRNA as the internal reference.

I found evidence for adaptive variation in CT_{max} under both ramping speeds, but in both cases, patterns of heat shock protein expression failed to fully explain enhanced upper thermal limits. When two species, the cooler-climate *A. picea*, and the warmer-climate *A. rudis*, were assayed in a fast ramping protocol after being housed under common garden conditions for two months, CT_{max} was significantly higher in *A. rudis* than in *A. picea* and was strongly predicted in *A. rudis* at the population level by mean annual temperature (Fig. 3). When colonies were assayed for basal levels of three major heat shock proteins (*hsp40*, *hsp70*, and *hsp83*), however, neither taxonomic identity nor mean annual temperature (°C) was significantly associated with their expression (multiple linear regression, all n.s., $p > 0.05$), although short-term variation in rearing temperature did lead to elevated basal levels of *hsp70* (multiple linear regression, $r^2=0.42$, $p<0.05$), suggesting that it is involved in the acclimation response.

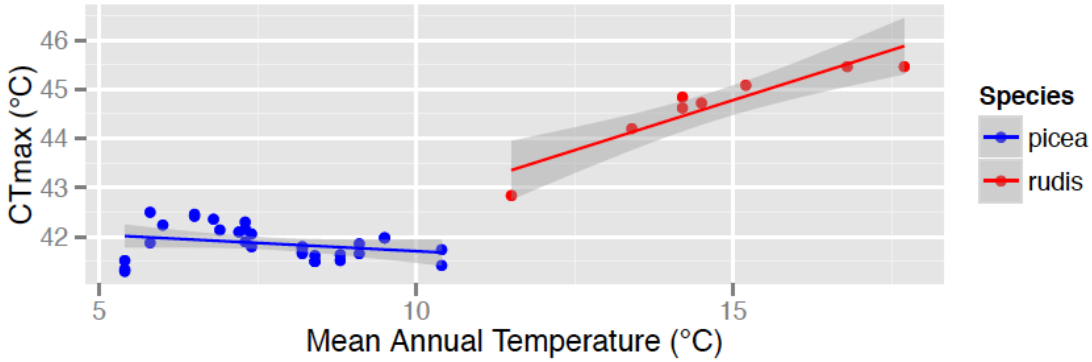

Figure 3. The relationship between heat tolerance (CT_{max} under fast ramp) and mean annual temperature (°C) for populations of *Aphaenogaster picea* (blue) and *A. rudis* (red).

Under the slow ramping protocol, I found no intraspecific or interspecific relationships between CT_{max} and mean annual temperature, either in these two focal species or across the broader *Aphaenogaster* clade (linear regression, $r^2=0.57$, $p=0.14$). However, two species that occur in open habitats, where nests are likely to experience greater chronic thermal stress due to direct solar radiation, displayed significantly higher

slow-ramp CT_{max} values than the rest of the clade (Phylogenetic ANOVA: $F_{1,78} = 62.51$, $p<0.05$, Fig. 4). Under a slow ramp, Hsp70 gene expression follows a sigmoidal-shaped curve which is best modeled with a Boltzmann function (Fig. 5, left panel), from which I can extract the extent of up-regulation (max), the critical temperature at which the gene is activated (Tm), and the expression velocity once activated (slope). If elevated upper thermal limits were directly conferred by the presence of Hsps, then higher CT_{max} should be associated with higher max, higher expression slope, and/or lower Tm values. Preliminary analyses for 10 out of 65 colonies, however, show a different pattern: species with higher CT_{max} have significantly higher critical temperatures (Tm; linear regression, $r^2=0.40$, $p<0.05$; Fig. 5, right panel) and slower expression velocities (slope; linear regression, $r^2=0.19$, $p<0.05$; Fig. 5, right panel), indicating that Hsps in these species are not activated until temperatures become more extreme. This result implies that mechanisms other than Hsp induction function to maintain physiological processes in the face of high temperatures in these species.

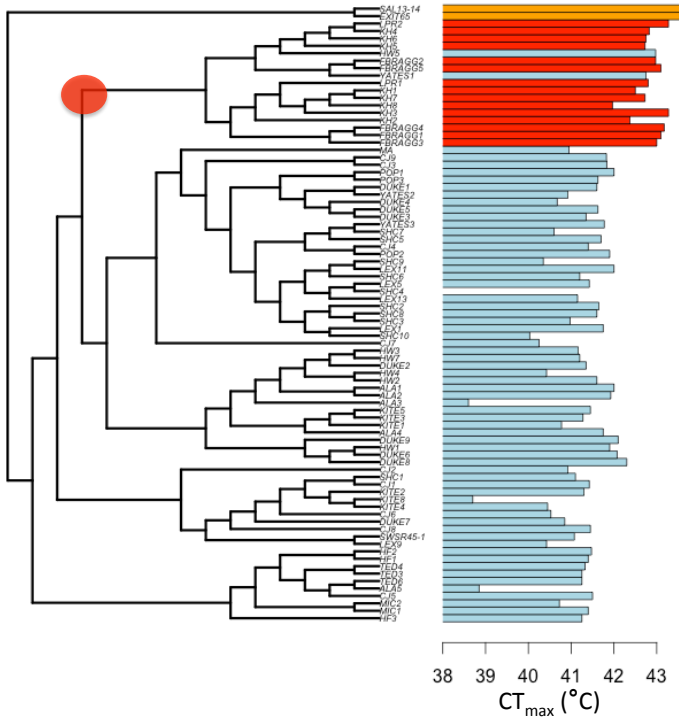

Figure 4. Slow-ramp CT_{max} mapped against phylogenetic relationships for colonies sampled from nine species of *Aphanogaster* (3-20 colonies measured per species). Colonies collected in deciduous forests are colored in blue, whereas colonies collected from more open flat woods are colored in red (barplot, right). Orange bars represent values for the outgroup (*Verromessor pergandei*). The node colored in red shows an evolutionary shift into more a thermally stressful environment (open habitat).

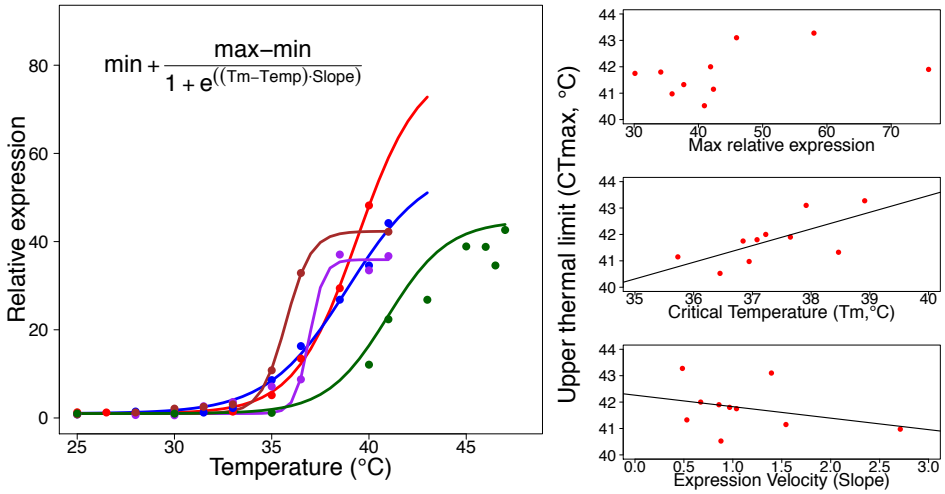

Figure 5. Preliminary *hsp70* expression shows that species vary in their stress response. For each species, points are fitted with the Boltzmann function (left panel) to determine whether species shift their max, critical temperature (Tm), and velocity (slope) expression patterns. Upper thermal limits of different species show a significant positive relationship with the critical temperature (Tm) and a negative relationship with expression velocity (slope).

Proposed DDIG Project

Background and relevance to dissertation

My dissertation work suggests that understanding the evolution of thermal tolerance will require consideration of more molecular mechanisms than molecular chaperones. One alternative that has been relatively understudied is a biochemical shift toward enhanced proteome stability (Somero 2004). Fine-scale studies of a few proteins have identified amino acid sequence changes that stabilize enzymes (Miyazaki et al. 2000; Fields 2001; Lockwood & Somero 2012) and expression of more stable isoforms (Hochachka & Somero 1968; Somero 1995) under high temperatures, which vary with heat tolerances. Whether these results scale up to the level of entire proteomes, however, remains to be determined: there have been no comprehensive empirical tests of the extent the proteome has been shaped by selection, either systematically or over a subset of functional classes.

Although there are few empirical data for whole proteomes, computational approaches suggest a number of potential molecular changes that may contribute to shifts in heat tolerance (Chakravarty & Varadarajan 2000; Karshikoff & Ladenstein 2001; Knight 2004; Dill et al. 2011). Comparisons of predicted proteomes (inferred from genome sequences) between mesophilic and thermophilic bacteria suggest that thermophiles have evolved greater proteome stability by shifting toward smaller average protein sizes to minimize the costs of folding large proteins and by investing in more polar residues to enrich for stabilizing non-covalent bonds (Chakravarty & Varadarajan 2000; Karshikoff & Ladenstein 2001). It should be noted, however, that because empirical measurements of each protein are lacking in these studies, it is not clear whether structural changes are causally related to protein stability, and these comparisons did not account for phylogenetic relationships (Felsenstein 1985).

Proteome-wide thermal profiling:

Recent innovations in mass spectrometry enable new ways to characterize and quantify the proteome (Savitski et al. 2014). Typically, standard techniques of liquid

chromatography-mass spectrometry/mass spectrometry (LC-MS/MS) can identify peptides by matching patterns of mass to charge (m/z) to a database, but do not allow for quantitative comparisons across mass spectrometry runs (Ong et al. 2002). To overcome this obstacle, tandem mass tags (TMT) have been developed to uniquely label up to 10 different samples which can be mixed together in one experimental run to allow for simultaneous identification and relative quantification within and among LC-MS/MS experiments given the correct controls (Savitski et al. 2014). To date, only one study has used TMT tagging, in a biomedical context, to quantify protein stability (Savitski et al. 2014). In collaboration with proteomics specialists Drs. Bryan Ballif and Ying Wai Lam, at the Vermont Genetics Network Proteomics facility, **I propose to use the TMT labeling method and their new LC-MS/MS instrument (Q-Exactive Hybrid Quadupole-Orbitrap Mass Spectrometry) to quantify the stability profile of the proteome across *Aphaenogaster* species to test whether evolutionary changes in stability are associated with shifts in Hsp induction and upper thermal limits.**

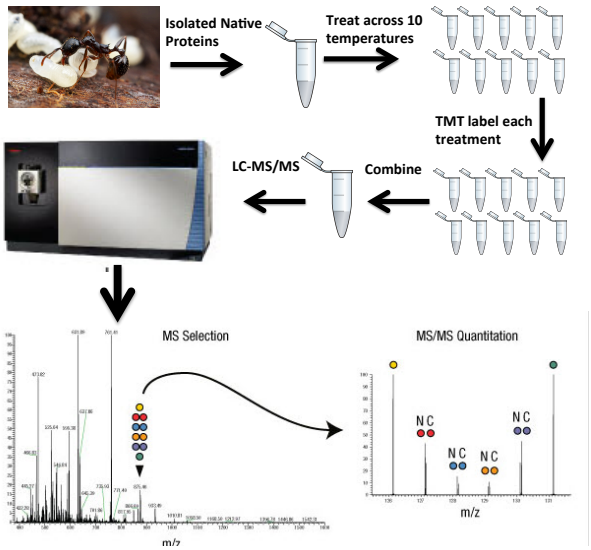

Figure 6. Modified schematic from ThermoFisher Scientific of LC-MS/MS approach in quantifying proteome stability. For a given species, native proteins will be isolated and treated across 10 temperature treatments ranging from 25-65 °C. Then, samples will be labeled with TMT reagent (after electrophoresis and trypsinization) and all treatments will be combined for LC-MS/MS analysis. Because each TMT reagent uniquely shifts the mass spectra, peptides from each treatment can be identified and quantified by comparing relative abundances. LC-MS/MS analyses will be conducted at the Vermont Genetics Network Proteomics Facility (http://vgn.uvm.edu/proteomics).

DDIG Question 1: Do more thermally tolerant species have more stable proteomes?

Predictions and Interpretations

Because proteins (enzymes) mainly drive biochemical reactions whose efficacy depends on their thermal stabilities, more heat tolerant species should have higher average proteome stabilities than less tolerant species. Alternatively, stability may be most critical for only a subset of the proteome, suggesting that relatively few amino acid changes may be required for temperature adaptation. This set of proteins may be enriched (relative to the protein set with no relationship) in processes such as metabolism, oxidative stress, and growth functions (Knight et al. 2004).

Experimental Approach and Data Analysis

To characterize thermal stability across the proteome of each of the 12 ant species, I will implement the TMT labeling method to quantify the stability of identified proteins using mass spectrometry (LC-MS/MS). Each species will include three biological (colony level) replicates; roughly 1mg of native proteins will be isolated for each colony and the isolate will be aliquotted evenly into 10 temperature treatments (100 μg per treatment) from 25°C-65°C for 10 minutes to denature those proteins that are unstable at the target temperature (Figure 6). The remaining non-denatured proteins will then be collected from each treatment through centrifugation and the samples will be labeled uniquely with TMT tags. Proteins from all temperature treatments will be combined together for identification and quantification with LC-MS/MS. Comparing mass spectra to known protein databases from the 10 published ant genomes (Wurm et al. 2009) will individually identify proteins, including splice variants (Figure 6). I expect to be able to unambiguously identify 3,000-6,000 proteins, representing roughly 25-50% of the set of genes in a typical ant genome(Gadau et al. 2012). For each protein identified, the amount of protein remaining at each temperature relative to the 25 °C treatment represents the fraction in the 'folded' state at that temperature. Finally, a logistic curve fit to the relationship between folded fraction and temperature will provide an estimate of the melting point (Tm), which is a measurement of thermal stability. I will test for a relationship between the average melting point across the proteome and each of the two CT_{max} measures (fast and slow ramp) using a phylogenetically controlled regression analysis (Felsenstein 1985), with enough power to detect positive relationships given 12 species and high phylogenetic signal (lambda = 0.73) in CT_{max} (Boettiger et al. 2012). To identify the subset of proteins associated with heat tolerance, I will perform independent evolutionary regression analyses for each protein (Felsenstein 1985) and then correct for multiple comparisons, followed by an enrichment analysis (Huang et al. 2008) to determine whether specific functional groups are targeted by selection.

DDIG Question 2:What biochemical properties are associated with enhanced protein stability in more heat tolerant species?

Predictions and Interpretations

Of the proteins that are likely important for temperature adaptation (DDIG Q1), I will test for the relative importance of expressing different isoforms and different biochemical properties such as bonding properties, sequence length, and structural motifs that are associated with enhanced proteome stability. If more heat tolerant species minimize the costs of folding larger proteins, then there should be a negative relationship between stability and sequence length (Fields 2001). Alternatively, if more heat tolerant species invest more in stabilizing bonding properties, then there should be a positive relationship between the abundance of charged, polar, and/or hydrophobic amino residues and protein stability (Karshikoff & Ladenstein 2001). Furthermore, significant positive relationships for polar amino acids reflects a strategy for more stabilizing non-covalent interactions, whereas, significant positive relationships for hydrophobic residues identifies a strategy for stabilization through the hydrophobic effect.

Experimental Approach and Data Analysis

To determine the types of biochemical/physical properties that underlie the stability of proteins, I will determine the proportion of variance in protein stability that can be explained by length and amino acid composition in a phylogenetically corrected multiple regression analysis (Felsenstein 1985). Since there will be multiple statistical

analysis for many proteins, significant relationships will be corrected for multiple comparisons. To determine the role of splice variation, I will determine the relationship between upper thermal limits and the degree of bias toward more thermally stable alternative isoforms with a phylogenetically corrected analysis of covariance (ANCOVA; Felsenstein 1985; Huang et al. 2008).

Summary, Synthesis, and Intellectual Merit

The goal of my thesis work is to understand patterns and mechanisms of temperature adaptation in a clade of common woodland ants. I have found that both temperature and habitat type have shaped ant thermal limits, but in somewhat different ways: while mean annual temperature is associated with inherent resistance to heat stress, as measured with a fast ramping protocol, habitat appears to drive evolution of the induced response, reflecting its more chronic effects on conditions within the nest. This decoupling of basal and acclimation adaptation deviates from patterns found in other systems such as fruit flies, where these two aspects of thermal tolerance are genetically correlated (van Heerwaarden & Sgro 2013), suggesting that ants represent a promising system for uncovering novel mechanisms of response to heat stress. Furthermore, Hsp expression does not appear to be the primary mechanism underlying these patterns, warranting further investigation of other possible mechanisms associated with upper thermal limits. DDIG funding will allow me to expand beyond the "usual suspects" of thermal adaptation, and explore the role of the proteome in driving temperature adaptation. As a result, I will be able to more effectively link potential mechanisms of thermal adaptation in this system with their specific physiological and ecological consequences.

The outcome of my dissertation and proposed DDIG project will also contribute to conservation practices and policies in two ways. First, many macro-evolutionary physiologists work under the assumption that basal and acclimation responses are correlated when assessing species responses to climate shifts (e.g., Deutsch et al. 2008, Sunday et al. 2010), which may underestimate or overestimate species vulnerabilities depending on which measure is used to calculate CT_{max}. Second, the collective or functionally relevant thermal stabilities of the proteome may provide more precise early indication for when species become stressed due to temperature increases. This is particularly useful for long-lived taxonomic groups for which direct fitness measurements cannot be readily quantified. In fact, the proposed methods in this DDIG proposal may be used to quantitatively compare protein abundances and other properties in field animals, which are likely to be sensitive to environmental conditions and thus can inform the efficacy of conservation policies and practices in real time.

Broader Impacts

Funding from a DDIG will provide resources for me to push my development as a scientist and educator in a number of ways. First, funding will improve my scientific training by gaining cutting-edge, critical skills such as mass spectrometry and bioinformatics to answer pressing questions, especially in an interdisciplinary framework. Second, I will train the next generation of biologists to become independent thinkers by mentoring undergraduate researchers, including those from underrepresented groups. Lastly, undergraduate researchers and I will present these findings to other members of the scientific community within the University of Vermont (Student Research Conference), at national and international conferences, and through publications.