UNIVERSITY OF CALIFORNIA SAN DIEGO

Heat stress response and evolution of thermal tolerance in the copepod *Tigriopus californicus*

A Thesis submitted in partial satisfaction of the requirements for the degree of
Doctor of Philosophy

in

Marine Biology

By

Sumaetee Tangwancharoen

Committee in charge:

Professor Ronald Burton, Chair
Professor Eric Allen
Professor Amro Hamdoun
Professor James Posakony
Professor Martin Tresguerres

2018

The Dissertation of Sumaetee Tangwancharoen is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____
Chair

University of California San Diego
2018

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First of all, I would like to express my gratitude to my advisor, Ron Burton for his guidance and support throughout the years. Ron always believes in my potential. He always told me that I should be more excited with my results. Ron encouraged me to dream big. When I proposed to work on allele specific expression in HSPB1 gene, Ron instead suggested me to do the whole RNA sequencing. He always encourages all of his students to present at conference even when we have doubts in our own work sometimes. Those conference experiences are definitely valuable for me to grow as a scientist.

To two of my committees, Martin and Amro, thank you for letting me work in your labs and for writing me a recommendation that landed me a job. Chapter 2 and 3 in this dissertation expanded from a proposal I wrote in Martin's class. I thank Martin for such inspiration. I also want to thank Martin for his suggestion for me to try the immunohistochemistry project even though it did not go as plan, but I learned a valuable technique that I can use in the future. To Amro, I am grateful to be your TA for 3 years. I learned a lot from you. You are a great teacher. Amro has asked me how I am going to start my own research lab in the future and gave me some advice. I thank my remaining committees, Eric Allen and Jim Posakony for all of your insightful comments and suggestions on my research.

Aside from the committees, many people made this dissertation possible. The most important person is Gary Moy who is a cloning expert and a co-author in Chapter 3. Thank you for teaching me the cloning skills and many biochemistry tips and tricks. Gary's contributions made the HSPB1 story more complete. Next, I would like to thank Felipe Barreto whose work on RNAi I built upon and whose work on both *Tigriopus* genome and transcriptome references I

used in my experiments. Without Felipe's previous work, this dissertation would not be possible. I also want to thank Kristen Jepsen from the Institute for Genomic Medicine at UCSD who handled my RNA sequencing.

I want to acknowledge a couple people who help me on various projects that did not end up in this dissertation but still contributed to the bigger picture of my research project. I give special thanks to Brice Semmens who developed a Bayesian model to analyze the allele specific expression that will be used in further analysis of the RNA sequencing data. I also want to thank Garfield Kwan for your time and expertise on the microscope work.

I thank my Thai government scholarship from The Institute of the Promotion of Teaching Science and Technology for my financial support during my time abroad in the United States. I would like to thank government officials at the Office of Educational Affairs at Thai Embassy in Washington DC and staff at Scripps especially Gilbert Bretado, Marty Tullar, Shelley Weisel, who took care of my financial and administrative support. Thank you Sam Chin Hoang and Yvonne Bohan, both of you made Hubbs Hall a friendly work place.

My experiences in the Burton lab have been friendly and fun. Special thanks to Sumi Hunjan, the social glue in the Burton lab who make this lab feels more like a family. Thank you to all past and present Burton lab family members: Ricardo Pereira, Thiago Lima, Satomi Tsuboko-Ishii, Tim Healy, Lani Gleason, Tessa Pierce, Reggie Blackwell, Maggie Sefton, Amanda Parker, Lori Luers, Elena Duke, Mado Krick, Manoela Costa Brandão, Ana Luisa Ahern, and Mark Sherman for your insights and comments on my work and your friendship over the years. Finally, to Alice, thank you for being a great friend, office mate, and academic sibling. Thank you for bringing Agnes, the best therapy dog, to our office. We shared the experiences of ups and downs in grad school together. I wish you all the best. Thank you to all my

undergraduate volunteers Joseph Chang, Kiana Woodward, Wenpei Li, and Kimberly Moreno for helping me with experiments and taking care of my copepods over the years. I am very proud of all of you.

To my friends at Scripps, thank you for your friendship and fond memories. Special thanks to Mike Tift, Jess Blanton, Ryan Guillemette, Logan People, Josefin Stiller, Josh Stewart, Lynn Waterhouse, Sarah Lerch, Natasha Gallo, Lauren Linsmayer, and Maitryei Nagarkar.

I would like to thank my friends in the Thai community at UCSD. You guys made my time abroad felt more like home. Thanks to Khun Pornpatcharapong, my comrade and roommate for 5 years; Win Hongjamrassilp and Tonmai Lertvilai, my fellow Thai students at Scripps; Boss Chansangavej and Vit Vorapipat, my home brew boys; Mint Limsakul, Kim Udomprapasap, June Pruegsanusak and Boat Jirapan for all weekend dinners, potlucks, concerts, and boardgame nights over the years.

Also, I would like to thank to all my undergraduate mentors at Duke who groomed me to become a scientist from early on. To my undergraduate advisor, John Willis, thank you for allowing me to be a part of the Willis lab family. I know that you wrote an incredible recommendation letter for me. To previous graduate students and post-doc in the Willis lab, Arielle Cooley, Youngwha Lee, Dave Lowry, and Lex Flagel, thank you for teaching me all the lab skills. Finally, thank you Dan Rittschof and Catherine McClenllan who introduced me to the wonderful world of lives in the ocean at Duke Marine lab.

Last but not least, I would like to thank my family, mom dad and my 2 sister Pui and Pun, for all the love and support from afar. Thank you, mom and dad, for letting me choose my own path and do whatever I want to do.

Chapter 1, in full, was published in the Marine Ecological Progress Series: Tangwancharoen S, Burton RS (2014) Early life stages are not always the most sensitive: heat stress responses in the copepod *Tigriopus californicus*. Marine Ecology Progress Series. 517:75-83. The dissertation author was the primary investigator and author of this paper.

Chapter 2, in full, is being prepared for submission: Tangwancharoen S, Burton RS. F1 RNA-seq reveals regulatory divergence underlying thermal adaptation in the copepod *Tigriopus californicus*. The dissertation author was the primary investigator and author of this paper.

Chapter 3, in full, was published in Molecular Biology and Evolution: Tangwancharoen S, Moy GW, Burton RS (2018) Multiple modes of adaptation: regulatory and structural evolution in a small Heat Shock Protein gene. Molecular Biology and Evolution. 35: 2110-2119. The dissertation author was the primary investigator and author of this paper.

VITA

EDUCATION

2011 Bachelor of Science in Biology. Graduation with Distinction in Biology. *magna cum laude*. Marine Science and Conservation Leadership Certificate, Duke University

2017 Master of Science in Marine Biology, Scripps Institution of Oceanography UCSD

2018 Doctor of Philosophy in Marine Biology, Scripps Institution of Oceanography UCSD
Advisor Dr. Ronald Burton

PUBLICATIONS

Tangwancharoen, S., Moy, G.W., Burton, R.S. (2018) Multiple modes of adaptation: Regulatory and structural evolution in a Small Heat Shock Protein gene. Molecular Biology and Evolution. 35: 2110-2119.

Tangwancharoen, S., Burton, R.S. (2014) Early life stages are not always the most sensitive: heat stress responses in the copepod *Tigriopus californicus*. Marine Ecology Progress Series. 517:75-83.

FIELDS OF STUDY

Evolutionary genetics
Evolutionary physiology
Molecular ecology
Marine Biology

ABSTRACT OF THE DISSERTATION

Heat stress response and evolution of thermal tolerance in the copepod *Tigriopus californicus*

by

Sumaetee Tangwancharoen

Doctor of Philosophy in Marine Biology

University of California San Diego, 2018

Professor Ronald Burton, Chair

With the warming trend due to climate change, conservation of species requires knowledge in ecological and evolutionary aspects of thermal tolerance and adaptation. In this dissertation, I use tidepool copepod, *Tigriopus californicus*, as a model for studying both aspects of thermal tolerance. For ecological aspect, identifying the most sensitive life stage can help us predict future responses especially in marine organisms with complex life history. I examined different survivorship to acute heat stress among life stages and across populations of *T. californicus*. Results revealed early life stages of *T. californicus* survived acute heat stress at higher temperatures than adults in contrast to popular belief. However, heat stress during larval stage of *T. californicus* resulted in developmental delay. Survivorship in larval and juvenile stages across populations also conform with a pattern previously observed in adults with more

heat tolerant populations toward southern range of the species. In order to uncover the evolutionary basis underlying thermal tolerant in *T. californicus*, I examined allele specific expression in F1 hybrid between populations from San Diego (SD) and Santa Cruz (SC). RNA sequencing revealed regulatory divergence in several gene ontology categories that potentially contribute to thermal tolerance including, electron carrier genes, genes involved in muscle and cuticle assembly, genes involved in proteolysis and Heat Shock Protein (HSP) genes. Heat Shock Protein Beta 1 (HSPB1) is one of the highest expressed HSPs in response to heat stress. HSPB1 Allelic imbalance suggested divergence in *cis* regulatory element underlying heat stress induced expression. HSPB1 promotor sequencing revealed polymorphisms in the Heat Shock Elements (HSEs), a binding site for Heat Shock Transcription Factor (HSF), where heat tolerant southern populations contain 2 canonical HSEs while northern populations have substitutions in the conserved motif of HSEs. Allele specific expression in more F1 crosses confirmed biased expression favoring alleles from populations with 2 canonical HSEs. Functional assays comparing recombinant SD and SC HSPB1 demonstrated that SD HSPB1 has a better capacity for preventing protein aggregation and preserving enzymatic function under high temperature. Overall, results from this dissertation provide insights on both ecological and evolutionary perspectives of thermal tolerance.

General Introduction

Local adaptation to different environments has been a major research topic in evolutionary biology. Temperature is a major driver for local adaptation since it influences every level of biological organization from species distribution down to molecular kinetics (Hochachka and Somero 2002; Kingsolver 2009). Organisms are rarely in thermal equilibria with the environments where they inhabit (Angilletta 2009). Populations of species that distribute across heterogeneous habitats must be locally adapted to their local thermal regimes. All organisms must be adapted to the range where their environmental temperatures fluctuate and must respond to temperature fluctuations during their life time which sometimes exceed optimal ranges and lead to stress. With the projected increase of mean global temperature and frequency of extreme events due to climate change (Pachauri, et al. 2014), it became necessary to understand how organisms are going to respond and evolve to changes in their thermal environments. Understanding the mechanisms underlying thermal tolerance can allow us to assess future impacts of climate change and contribute to species conservation efforts.

At the molecular level, high temperature can disrupt protein folding and function which leads to proteotoxic stress (Somero 1995; Feder and Hofmann 1999; Gidalevitz, et al. 2011; Morimoto 2011). To cope with stress and damage, cells will induce a response commonly referred as Heat Shock Response (HSR). HSR involves expression of various types of genes including Heat Shock Proteins (HSPs), molecular chaperones that assist protein folding and protect cells from protein misfolding damages (Lindquist 1986; Parsell and Lindquist 1993; Feder and Hofmann 1999). HSR is largely transcriptionally regulated by Heat Shock Transcription Factor (HSF) (Wu 1995; Akerfelt, et al. 2010; Weake and Workman 2010). The mechanisms of HSF transcriptional induction have been figured out in *Drosophila melanogaster*

HSP70 (Lis, et al. 2000). It has been widely accepted that evolution of HSP gene expression plays a key role in adaptation to thermal environments (Sørensen , et al. 2003; Hofmann 2005; Sørensen 2010; Tomanek 2010; Chen, et al. 2018). HSP expression levels have been linked to species or populations' distribution across different climatic variations (Tomanek and Somero 1999; Hofmann 2005; Mizrahi, et al. 2016). Even HSP sequences also correlate with geographic distributions (Frydenberg, et al. 2003; Graham, et al. 2012). Despite the evidences for the importance of HSPs in the evolution of thermal tolerance, the link between adaptive phenotypes and their genetic bases remains poorly understood.

In this dissertation, I am using tidepool copepod *Tigriopus californicus* as a model to study HSR and the molecular mechanisms underlying heat stress responses and thermal adaptation. Many intertidal species including *T. californicus* are great systems for studying thermal tolerance because they experience extreme temperature fluctuations frequently. For example, temperatures in tidepools can fluctuate as much as 20 °C within a day (Kelly, et al. 2012). Studies have also shown that many intertidal organisms have already live close to their upper thermal limit (Somero 2010; Tomanek 2010). *T. californicus* system also provides an excellent system for studying thermal adaptation. The species occupies a broad geographic range extending from central Baja California, Mexico (Ganz and Burton 1995), to southeast Alaska (Dethier 1980). Isolated populations along rocky outcrops also show high level of genetic divergence suggesting no gene flow between populations (Burton and Lee 1994; Edmands 2001; Peterson, et al. 2013). Studies have revealed evidences of local adaptation to thermal regimes at the phenotypic level through survivorship (Willett 2010; Kelly, et al. 2012; Leong, et al. 2018) and difference in transcriptional response following heat stress (Schoville, et al. 2012; Lima and Willett 2017).

A recently developed allele specific expression technique has allow us to tease apart divergence in the genetic basis of transcriptional regulations (McManus, et al. 2010). Several studies have adopted the technique to study the genetic basis underlying adaptation to ecological environments (Lovell, et al. 2016; Mack, et al. 2016; Gould, et al. 2018). Together with recently published genome (Barreto, et al. 2018) and transcriptomic resources of several *T. californicus* populations (Schoville, et al. 2012; Barreto, et al. 2015; Pereira, et al. 2016), these allowed me to study the genetic basis of heat shock response and thermal adaptation in *T. californicus* system.

The main goal of this dissertation is to expand knowledge on the genetic basis underlying thermal adaptation with the focused goal on the evolution in the Heat Shock Response transcriptional regulation. In this thesis, I used comparative approaches by comparing locally adapted populations of *T. californicus* at different levels including DNA sequence, mRNA expression, protein function, and heat tolerance phenotypes in order to identify the potential causes underlying differences in thermal tolerance phenotypes.

In chapter 1, the objectives were to expand the existing knowledge on acute heat stress effects and heat tolerance phenotypes among populations of our study system, *T. californicus*. Chapter 1 examined heat tolerance phenotypes in different life stages of *T. californicus* and prolonged effects of acute heat stress during larval stage of *T. californicus*.

In chapter 2, I explored patterns of divergence in *cis* and *trans* regulatory elements, and patterns of inheritance gene expression phenotypes. The goals were to address the genetic basis underlying thermal adaptation by analyzing the genome-wide transcriptional responses to heat stress and allele specific expression of F1 hybrids and their parental populations, San Diego (SD) and Santa Cruz (SC), which have different heat tolerance phenotypes.

Chapter 3 focuses on the role of HSPB1, a gene previously found as one of the highest induction in response to heat stress in *T. californicus*, in the evolution of heat tolerance. Experiments in chapter 3 include both transcriptional regulation and protein function aspects of HSPB1 by comparing promotor sequences and allele specific expression from more *T. californicus* populations and protein functional assays on HSPB1 from SD and SC population.

Together, results from all chapters will provide insights into life history impacts of heat stress and genetic basis underlying thermal adaptation. The work of this dissertation will have broader impacts on the fields of thermal physiology, molecular ecology and evolutionary biology.

References

Akerfelt M, Morimoto RI, Sistonen L. 2010. Heat shock factors: integrators of cell stress, development and lifespan. Nature Reviews Molecular Cell Biology 11:545-555.

Angilletta MJ. 2009. Thermal Adaptation: A Theoretical and Empirical Synthesis. Thermal Adaptation: a Theoretical and Empirical Synthesis:1-290.

Barreto FS, Pereira RJ, Burton RS. 2015. Hybrid Dysfunction and Physiological Compensation in Gene Expression. Molecular Biology and Evolution 32:613-622.

Barreto FS, Watson ET, Lima TG, Willett CS, Edmands S, Li WZ, Burton RS. 2018. Genomic signatures of mitonuclear coevolution across populations of Tigriopus californicus. Nature Ecology & Evolution 2:1250-1257.

Burton RS, Lee BN. 1994. Nuclear and mitochondrial gene genealogies and allozyme polymorphism across a major phylogeographic break in the copepod *Tigriopus californicus*. Proceedings of the National Academy of Sciences of the United States of America 91:5197-5201.

Chen B, Feder ME, Kang L. 2018. Evolution of heat-shock protein expression underlying adaptive responses to environmental stress. Molecular Ecology 27:3040-3054.

Dethier MN. 1980. Tidepools as refuges: predation and the limits of the Harpacticoid copepod *Tigriopus californicus* (Baker). Journal of Experimental Marine Biology and Ecology 42:99-111.

Edmands S. 2001. Phylogeography of the intertidal copepod *Tigriopus californicus* reveals substantially reduced population differentiation at northern latitudes. Molecular Ecology 10:1743-1750.

Feder ME, Hofmann GE. 1999. Heat-shock proteins, molecular chaperones, and the stress response: Evolutionary and ecological physiology. Annual Review of Physiology 61:243-282.

Frydenberg J, Hoffmann AA, Loeschcke V. 2003. DNA sequence variation and latitudinal associations in hsp23, hsp26 and hsp27 from natural populations of *Drosophila melanogaster*. Molecular Ecology 12:2025-2032.

Ganz HH, Burton RS. 1995. Genetic differentiation and reproductive incompatibility among Baja California populations of the copepod *Tigriopus californicus*. Marine Biology 123:821-827.

Gidalevitz T, Prahlad V, Morimoto RI. 2011. The Stress of Protein Misfolding: From Single Cells to Multicellular Organisms. Cold Spring Harbor Perspectives in Biology 3.

Gould BA, Chen Y, Lowry DB. 2018. Gene Regulatory Divergence Between Locally Adapted Ecotypes in Their Native Habitats. Molecular ecology.

Graham AM, Merrill JD, McGaugh SE, Noor MAF. 2012. Geographic Selection in the Small Heat Shock Gene Complex Differentiating Populations of *Drosophila pseudoobscura*. Journal of Heredity 103:400-407.

Hochachka PW, Somero GN. 2002. Biochemical adaptation: mechanism and process in physiological evolution. Biochemical adaptation: mechanism and process in physiological evolution.:i-xi, 1-466.

Hofmann GE. 2005. Patterns of Hsp gene expression in ectothermic marine organisms on small to large biogeographic scales. Integrative and Comparative Biology 45:247-255.

Kelly MW, Sanford E, Grosberg RK. 2012. Limited potential for adaptation to climate change in a broadly distributed marine crustacean. Proceedings of the Royal Society B-Biological Sciences 279:349-356.

Kingsolver JG. 2009. The Well-Temperatured Biologist. American Naturalist 174:755-768.

Leong W, Sun PY, Edmands S. 2018. Latitudinal Clines in Temperature and Salinity Tolerance in Tidepool Copepods. Journal of Heredity 109:71-77.

Lima TG, Willett CS. 2017. Locally adapted populations of a copepod can evolve different gene expression patterns under the same environmental pressures. Ecology and Evolution 7:4312-4325.

Lindquist S. 1986. The Heat-Shock Response. Annual Review of Biochemistry 55:1151-1191.

Lis JT, Mason P, Peng J, Price DH, Werner J. 2000. P-TEFb kinase recruitment and function at heat shock loci. Genes & Development 14:792-803.

Lovell JT, Schwartz S, Lowry DB, Shakirov EV, Bonnette JE, Weng XY, Wang M, Johnson J, Sreedasyam A, Plott C, et al. 2016. Drought responsive gene expression regulatory divergence between upland and lowland ecotypes of a perennial C-4 grass. Genome Research 26:510-518.

Mack KL, Campbell P, Nachman MW. 2016. Gene regulation and speciation in house mice. Genome Research 26:451-461.

McManus CJ, Coolon JD, Duff MO, Eipper-Mains J, Graveley BR, Wittkopp PJ. 2010. Regulatory divergence in *Drosophila* revealed by mRNA-seq. Genome Research 20:816-825.

Mizrahi T, Goldenberg S, Heller J, Arad Z. 2016. Geographic variation in thermal tolerance and strategies of heat shock protein expression in the land snail Theba pisana in relation to genetic structure. Cell Stress and Chaperones 21:219-238.

Morimoto RI editor. Cold Spring Harbor symposia on quantitative biology. 2011.

Pachauri RK, Allen MR, Barros VR, Broome J, Cramer W, Christ R, Church JA, Clarke L, Dahe Q, Dasgupta P. 2014. Climate change 2014: synthesis report. Contribution of Working Groups I, II and III to the fifth assessment report of the Intergovernmental Panel on Climate Change: IPCC.

Parsell DA, Lindquist S. 1993. The function of Heat-Shock Proteins in stress tolerance - degredation and reactivation of damaged proteins. Annual Review of Genetics 27:437-496.

Pereira RJ, Barreto FS, Pierce NT, Carneiro M, Burton RS. 2016. Transcriptome-wide patterns of divergence during allopatric evolution. Molecular Ecology 25:1478-1493.

Peterson DL, Kubow KB, Connolly MJ, Kaplan LR, Wetkowski MM, Leong W, Phillips BC, Edmands S. 2013. Reproductive and phylogenetic divergence of tidepool copepod populations across a narrow geographical boundary in Baja California. Journal of Biogeography 40:1664-1675.

Schoville SD, Barreto FS, Moy GW, Wolff A, Burton RS. 2012. Investigating the molecular basis of local adaptation to thermal stress: population differences in gene expression across the transcriptome of the copepod *Tigriopus californicus*. BMC Evolutionary Biology 12.

Somero G. 2010. The physiology of climate change: how potentials for acclimatization and genetic adaptation will determine 'winners' and 'losers'. Journal of Experimental Biology 213:912-920.

Somero GN. 1995. Proteins and temperature. Annual Review of Physiology 57:43-68.

Sørensen JG. 2010. Application of heat shock protein expression for detecting natural adaptation and exposure to stress in natural populations. Current Zoology 56.

Sørensen JG, Kristensen TN, Loeschcke V. 2003. The evolutionary and ecological role of heat shock proteins. Ecology Letters 6:1025-1037.

Tomanek L. 2010. Variation in the heat shock response and its implication for predicting the effect of global climate change on species' biogeographical distribution ranges and metabolic costs. Journal of Experimental Biology 213:971-979.

Tomanek L, Somero GN. 1999. Evolutionary and acclimation-induced variation in the heat-shock responses of congeneric marine snails (genus *Tegula*) from different thermal habitats: Implications for limits of thermotolerance and biogeography. Journal of Experimental Biology 202:2925-2936.

Weake VM, Workman JL. 2010. Inducible gene expression: diverse regulatory mechanisms. Nature Reviews Genetics 11:426-437.

Willett CS. 2010. Potential fitness trade-offs for thermal tolerance in the intertidal copepod *Tigriopus californicus*. Evolution 64:2521-2534.

Wu C. 1995. Heat shock transcription factors: Structure and regulation. Annual Review of Cell and Developmental Biology 11:441-469.

Chapter 1

Early life stages are not always the most sensitive: heat stress responses in the copepod *Tigriopus californicus*

Sumaetee Tangwancharoen, Ronald S Burton

Abstract

Because of their complex life histories, different life stages of many marine invertebrates may be exposed to varying environmental challenges. Ultimately, the life stage that is least tolerant of its environment will determine the species' abundance and distribution. The intertidal copepod *Tigriopus californicus* lives in high intertidal pools along the Pacific coast of North America. Unlike most other invertebrates, the different life stages of *T. californicus* all share the same tidepool habitat. To determine physiological tolerances of various life-history stages of this species, we examined responses to acute heat stress in nauplii, copepodids, and adults from six populations along a latitudinal gradient. Results show that early developmental stages (nauplii and copepodids) are generally more tolerant than adults. Our results contrast with the widely accepted generalization that larval forms are more sensitive to physical stressors than adults. As previously observed in adults, nauplii and copepodids from southern populations survive higher temperatures than those from northern populations. Acute heat stress was found to delay development but did not affect adult size. We hypothesize that variation in the thermal tolerance of early life stages among intertidal species reflects ecological differences in larval habitats: where larvae remain in the intertidal and experience the same high temperatures as adults, selection will favor high larval thermal tolerance, while in species with planktonic larvae, the buffered temperature regime of the water column might relax such selection and thermal tolerance will be highest in the more exposed intertidal adults.

Introduction

The life cycles of many marine invertebrates include developmental stages that can vary greatly in all aspects of ecology. Some life stages inhabit completely different environments, as the case where sedentary benthic adults broadcast spawn and the offspring undergo development as planktonic embryos and larvae (Miller et al. 2013). Given such changes in organismal ecology over a species' life cycle, various life stages may experience substantially different environmental extremes. Consequently, natural selection for tolerances to those extremes may also vary across life stages and potentially result in stage-specific tolerances to stress. Although there have been numerous studies of the physiological tolerances to heat stress of many invertebrates, data across all life stages is available for relatively few marine species, including some mollusks (e.g.,Verween et al. 2007), ascidians (Pineda et al. 2012) and arthropods (Miller et al. 2013). Most of these studies have concluded that early life stages are generally less tolerant of abiotic environmental extremes than adults of the same species. However, recent studies show that this pattern is not universal; for example, early life stages of the mollusk *Crepidula fornicata* are more tolerant of heat stress than adults (Diederich & Pechenik 2013).

Because larval tolerances to thermal stress can differ from those of conspecific adults, consideration of all life stages is necessary for understanding how species persist in their local environments. For example, although geographic ranges of many species are typically attributed to adult thermal tolerances, Walther et al. (2013) recently pointed out that thermal sensitivity of larval stages can act as a physiological barrier for dispersal. They found that larvae of different gooseneck barnacle (*Pollicipes elegans*) populations have different thermal maximum ranges and these reflect the environment of their respective habitats. Larvae from cooler regions north and south of the equator cannot disperse across the equator where temperature is higher, leading to

genetic differentiation of populations (Plough & Marko 2014). Similar patterns have also been found in other intertidal species such as mussels (Hilbish et al. 2000) and limpets (Koufopanou et al. 1999).

In addition to inducing mortality, heat stress may also affect other fitness parameters, such as rates of growth and larval development (Rothlisberg 1979, Preston 1985, Bryars & Havenhand 2006, Bermudes & Ritar 2008, Roberts et al. 2012). Although increasing temperature can speed up development over certain temperature ranges, higher (stressful) temperatures can also slow developmental rates (Roberts et al. 2012). Stressful temperatures can be defined as the temperature range that reduces the performance of an organism (Pörtner & Knust 2007). Reduction in performance can include slower growth and development or it can be as severe as mortality. For meroplankton, reduced developmental rates lead to longer larval duration, which in turn increases the risk of starvation and predation (Morgan 1995). Understanding the effects of thermal stress across all life stages will allow us to better predict the fate of populations and species in a warming global climate.

The copepod *Tigriopus californicus* provides an excellent system for studying thermal adaptation. *T. californicus* has a broad geographic range extending from central Baja California, Mexico (Ganz & Burton 1995, Peterson et al. 2013), to southeast Alaska (Dethier 1980) and several studies have revealed extensive genetic differentiation among populations across this range (Burton & Lee 1994, Edmands 2001, Willett & Ladner 2009). Since the geographic range of *T. californicus* spans a broad climatic gradient, responses to thermal stress across populations of this species have been the focus of recent research (Edmands & Deimler 2004, Willett 2010, Kelly et al. 2012, Schoville et al. 2012, Kelly et al. 2013, Pereira et al. 2014). These studies show a clear pattern with adult copepods from southern populations surviving higher

temperatures than do adult copepods from northern populations, an observation consistent with local adaption of populations to their respective habitats. Notably, local adaptation is most evident in population differences in survivorship following acute (short exposures of ~60 min duration) thermal stress (e.g., Willett 2010). Ecologically this is consistent with the fact that average temperatures of tide pools in California are generally moderate, ranging from approximately 10 – 20 ^{o}C, but thermal maxima can reach over 40 ^{o}C occasionally in summer months (Kelly et al. 2012). Because these maximum temperatures are only achieved during relatively brief periods in the middle of hot, sunny days, natural selection will favor mechanisms for short-term tolerance to thermal extremes and these extremes are more frequent in the southern portion of the geographic range.

To date, no studies have examined thermal stress responses among the larval stages of *T. californicus*. If pre-reproductive stages of *T. californicus* are more sensitive to heat stress than adults, the species could be more vulnerable to increasing temperature than previously predicted. However, an interesting aspect of the *T. californicus* system is that, in contrast with most other marine invertebrates with free-swimming larvae, all life stages of *T. californicus* occupy the same tide pool habitats and are subjected to the same extreme thermal regimes. Consequently, although the larvae of many marine invertebrates experience the relatively buffered thermal environment of the pelagic water column, *T. californicus* larvae do not. We hypothesize that larval thermal sensitivities will reflect the thermal environments to which they are exposed. Under this hypothesis, *T. californicus* larvae are expected to have higher thermal tolerances than larvae that develop in the water column.

In this experiment, we addressed three main questions: (1) Do different life stages of *T. californicus* (nauplius, copepodid, and adult) vary in tolerance to thermal stress? (2) Does heat

tolerance in nauplii and copepodids vary among populations? (3) Does acute thermal stress have any long term effects on larval development and adult size? In the absence of extensive migration, populations can only persist where all life stages are tolerant to local conditions; hence understanding patterns of thermal sensitivity in different life stages and populations may provide us with insight into how *T. californicus* populations persist in stressful environments.

Materials and Methods

Laboratory culture

Six populations of *Tigriopus californicus* were collected from high intertidal rock pools along the coast of California and Baja California (Figure 1-1). Each stock population was maintained in the laboratory in 4 or more 400 ml beakers filled with 250 ml of filtered (5 μm) sea water at 20 °C with a 12-hour light-dark photoperiod. Each beaker contained several hundreds to thousands of individuals, reflecting the density range observed in natural populations. Beakers containing animals from the same population were intermixed monthly to prevent inbreeding in laboratory cultures. The animals were fed ground *Spirulina* wafers and were acclimated to laboratory conditions for at least one generation (approximately four weeks) prior to experimental treatments. Salinity was monitored with a refractometer and maintained at 35 ppt.

Heat stress experiment

All three life stages of *T. californicus* were obtained separately for heat stress experiments. To obtain nauplii, gravid females with mature egg sacs (indicated by orange-red color and visible naupliar eye spots under a stereomicroscope) were collected from stock

populations. Egg sacs were separated from females using needles and each clutch was placed in filtered seawater and maintained at 20 °C overnight. On the following day, hatched nauplii were counted and divided equally into two treatments, control and heat stress. Except for rare undeveloped (and perhaps unfertilized) eggs, 100% hatching rates were observed in these experiments. Clutches producing less than 20 nauplii were pooled into one sample before separating into control and heat stress treatment so that each control and treatment replicate had at least 10 individuals. To obtain copepodids, egg clutches were obtained the same way as the nauplii, but the larvae were kept at 20 °C until the majority of each clutch metamorphosed into copepodids, approximately eight days after hatching. The numbers of individuals from each clutch were counted and all individuals were used in the heat stress experiment. Since all hatchlings from the same egg clutch were full-sibs, all nauplii or copepodids from the same clutch or pool of clutches were considered one replicate. Adult samples were collected from outbred stock populations with each experimental replicate containing 10 individuals.

To conduct heat stress experiments, copepods of the same life stage were put in 2 ml filtered seawater in 15 ml Falcon tubes. Each replicate tube had at least 10 individual copepods. The number of copepods used in each life stage and treatment ranged from 10 to 299 individuals. For control treatments, animals were maintained at 20^{o} C with a 12-hour light-dark period. The number of replicates used in the control treatments is shown in Table 1-1. From our pilot experiment, only the naupliar stage had variable survivorship. Almost all copepodids and adults survived after 3 days under the control treatment at 20 ^{o}C. Therefore, we only conducted 20 ^{o}C control treatment for the naupliar stage. For short-term (acute) heat stress experiments, tubes were submerged in a water bath at temperatures ranging from 34 to 38° C for one hour. This range of temperatures has been observed along the California coast in *T. californicus* habitats

(Willett 2010, Kelly et al. 2012) and is known to cause differential mortality in adults from different populations (Willett 2010, Kelly et al. 2012, Pereira et al. 2014). The numbers of replicates for different life stages and populations are shown in supplement Table A1-1. During the experiment, the temperature in the experimental tubes increased by more than 10 ^{o}C within first five minutes and reached the target temperature within 10 minutes. Following stress treatments, all animals were returned to 20° C with a 12-hour light-dark period. Ground *Spirulina* wafer was added to each tube as food after the stress treatment. Survivors were counted three days after the heat stress.

To perform statistical analysis, survivorship proportions were arcsine transformed. Data were analyzed using Analysis of Variance (ANOVA) with temperatures as predictor and life stage and population as covariates. Tukey-Kramer Honest Significant Difference (HSD) tests were used as a post-hoc analysis to compare the effects among different life stages and populations (analyses were carried out using JMP, version 11). Copepod survivorship proportions of the control and heat stress (34 ^{o}C) treatments were compared using Wilcoxon signed-rank test.

Heat stress effects on developmental rate and adult size

San Diego (SD) nauplii from both control and heat stress treatments at 36 °C were monitored after the three day survivorship count to test for effects of short-term heat stress on development. For each treatment sample, the date the first copepodid appeared was recorded to test for an effect of heat stress on developmental rate; because nauplii and early stage copepodids cannot be sexed, no sex-specific analysis of these results could be made. After the copepodids matured into adults, body lengths of both adult males and females were measured under a stereomicroscope. Mean body length of adult copepods (n = 124) from control and heat stress

treatments were compared. Data on developmental time from nauplius to copepodid and adult size were analyzed using Wilcoxon's tests JMP, version 11).

Results

Heat tolerance across populations and life stages

Even under benign control temperatures (20 °C), nauplii from all populations showed some mortality during the first three days after hatching. Survival rates of nauplii at 34°C were similar to those of the control samples from the same clutches (Table 1-1), indicating that this heat shock did not result in mortality of nauplii. In contrast, copepodids and adult copepods showed little or no mortality under control conditions, and like nauplii, experienced no mortality following exposure to mild heat stress (at 34 °C) (Table 1-1, Figure 1-2A). Although survivorship of nauplii from all six populations at 34 °C and 35 °C was lower than the survivorship of copepodids and adults at the same temperature, this appears to simply reflect the mortality of nauplii under the control treatment. Survivorship of nauplii was generally higher than other life stages at temperatures higher than 35 °C. Adults from all populations did not survive acute heat stress at 37 °C (Figure 1-3B). However, some nauplii and copepodids from the southern populations survived at 37 °C. At 36 °C both nauplii and copepodids from the northern populations survived in higher proportions than adults of the same populations. Nauplii also generally have higher average survivorship than copepodids at the temperatures where each population was at its maximum thermal limit (37 °C in southern populations and 36 °C in northern populations), except for the Abalone Cove (AB) population where average survivorship

of copepodids was the highest among all life stages at 37 °C (Figure 1-3B). These results show that pre-reproductive stages of *T. californicus* are more heat tolerant than adults.

Significant differences in survivorship were observed between northern (Pescadero, PES and Santa Cruz, SCN) and southern (La Bufadora Mexico, BUF; San Diego, SD; Bird Rock, BR and Abalone Cove, AB) populations (Table 1-2, Table 1-3B). Southern populations exposed to acute heat stress at 36 °C showed higher survivorship than northern populations in all three life stages (Figure 1-3). At 37 °C, difference in survivorship between northern and southern populations was the most significant in the copepodid stage. Some copepodids from southern populations survived while copepodids from northern populations experienced 100% mortality.

Heat stress effects on developmental rate and adult size

A total of 58 clutches of SD nauplii, from control and one hour heat stress at 36 °C, were monitored to determine the effects of stress on developmental rate. Of these, 11 clutches (124 individuals) were measured for adult sizes. Acute heat stress significantly affected developmental rate in *T. californicus*; nauplii that experienced acute stress took longer to develop into copepodids than nauplii that were raised at 20 °C (Figure 1-4, mean 8.80 days vs. mean 8.03 days respectively; Wilcoxon's test: $Z = 3.10$, $p = 0.002$). No differences in adult size of either sex were observed between control and heat stress treatments (Wilcoxon's tests -- males: $n = 74$, $Z = 0.11$, $p = 0.91$; females: $n = 50$, $Z = 0.08$, $p = 0.94$).

Discussion

After accounting for mortality under control laboratory conditions, we found that nauplii of the tidepool copepod *T. californicus* were generally more tolerant of acute heat stress than

copepodid or adult life stages. Nauplii survived heat stress at extreme temperatures (36 ^{o}C for northern populations, PES and SCN, and 37 ^{o}C for southern populations, AB, BR, SD and BUF) in greater proportions than did adults. Although the statistical analysis suggests that copepodid and adult are not significantly different from each other (Table 1-3A), some copepodids from all 4 southern populations survived heat stress at 37 ^{o}C while the adults did not (Figure 1-3B). These results indicate that both pre-reproductive stages (nauplii and copepodids) of *T. californicus* are more heat tolerant than reproductive adults.

Our results contrast with the widely held generalization that larvae of marine invertebrates are more sensitive to environmental stresses than adults. Several previous studies have suggested that larval stages of marine species are typically more sensitive to heat stress than adults (Verween et al. 2007, Pineda et al. 2012, Miller et al. 2013). Such differences might reflect the gradient of thermal environments encountered by different life stages. Intertidal habitats are among the most physically challenging of all marine environments, and a diverse fauna has evolved adaptations to cope with rapid fluctuations in temperature, salinity and other abiotic variables. A significant proportion of those species, however, produce planktonic larvae that develop in the water column, an environment that is substantially buffered from abiotic variation. Consequently, selection for tolerance to these stresses may be relaxed for these planktonic life stages. Miller et al. (2013) use this line of reasoning in explaining the fact that the zoeal larval stage of a porcelain crab species (*Petrolisthes cinctipes*) has the lowest LT50 (temperature which results in 50% mortality) of all life stages (i.e., zoea are the most sensitive to thermal stress); because eggs and early embryos are brooded and late larval stages recruit back to the intertidal, only zoea escape exposure to thermal extremes.

Although the above argument holds for a great many intertidal invertebrates, not all have planktonic development. Although all life stages of *T. californicus* are able to actively swim, genetic analyses clearly demonstrate a remarkable lack of gene flow between *T. californicus* populations inhabiting neighboring rock outcrops, indicating that dispersal in the plankton is extremely rare (Burton & Feldman 1981, Edmands 2001, Willett & Ladner 2009) and all life stages typically occur in rocky tidepools. Consequently, all life stages of *T. californicus* experience the same highly variable thermal regime and are therefore expected to be adapted to the same temperature range. In fact, we found that pre-reproductive stages of *T. californicus* are more heat tolerant than adults; this may reflect the fact that mortality in pre-reproductive stages has a stronger selective effect on the population because it results in no offspring. Therefore, in systems where all life stages occupy similar temperature ranges, pre-reproductive stages need to be at least as tolerant to abiotic stresses as adults. In this regard, the finding of high thermal tolerance in *T. californicus* larvae parallels the recent findings of Diederich and Pechenik (2013) on the brooding gastropod *Crepidula fornicata* whose embryos also experience the same thermal regime as adults.

Recent studies have revealed significant differences in thermal tolerance in adult *T. californicus* from different geographic populations. Willett (2010) and Kelly et al. (2012) have shown that populations of *T. californicus* differ in lethal maximum temperatures, with southern populations consistently surviving higher acute temperature stresses than northern populations. Both studies conclude that each population is adapted to its respective local temperature range. Kelly et al. (2012) have shown that temperatures in tide pools where *T. californicus* during spring and summer months in California range from 8 ^{o}C to more than 40 ^{o}C. Results from the present study are consistent with previous studies: adults from southern populations of *T.*

californicus can tolerate higher temperatures than those sampled from northern populations. Here we found that this pattern also holds for early life stages. Both nauplii and copepodids of the southern populations can tolerate higher temperatures than the same life stages from northern populations. Again, we interpret these differences in thermal tolerance among populations of *T. californicus* as evidence of adaptation to local environments.

Since all life stages of *T. californicus* developed in a 20 °C "common garden" environment for a full generation prior to thermal stress exposure, the differences in thermal tolerance among life stages may largely reflect genetic effects. With increasingly stressful environments, the least heat tolerant stage will determine population persistence in the future environment. Knowing the pattern of heat stress survivorship among life stages allows us to better predict how populations of *T. californicus* will fare in response to increasing temperature. For example, our results indicate that nauplii and copepodids of the southern California populations can survive at 37 °C stress but adults cannot. Frequent heat stresses above 37 °C could threaten populations with extirpation because the adults do not survive and fail to reproduce. Kelly et al. (2012) suggest that populations of *T. californicus* have limited potential to adapt to increasingly stressful temperatures; consequently, local extirpation might result from increases in environmental temperature. However, Pereira et al. (2014) found that hybrid lineages between neighboring southern California populations, SD and BR, frequently have thermal tolerances exceeding either parental population. Although migration between populations is known to be very low, these results suggest a potentially important evolutionary scenario where rare dispersal and hybridization could result in further adaptation to environmental extremes.

As discussed above, comparisons of stress tolerance across life stages in our experiments met with one unexpected complication: nauplii (but not copepodids or adults) show significant mortality during lab rearing at the 20°C control temperature. This mortality appears to be independent of heat stress since levels of mortality among nauplii do not increase significantly at temperatures up to 34°C. The cause of this baseline level of mortality among nauplii is unknown, but is likely due to both natural causes and experimental handling (required to count and transfer nauplii between culture and experimental vessels). Despite this background level of mortality, nauplii still showed higher survivorship than all adults and most copepodids under acute high temperature stress (37 °C in BUF, SD, BR, and AB and 36 °C in SCN and PES), indicating that nauplii are, in fact, the most thermally tolerant of the three life stages. Survivorship of copepodids is also higher than that observed among adults of the same population, implying that adult *T. californicus* might be the most thermally sensitive life stage.

Although larvae and juveniles of *T. californicus* can survive higher temperatures than adults, acute heat stress may have other fitness consequences. Numerous studies have shown that high temperatures increase metabolic rate and lead to faster development in many marine species (Roberts et al. 2012, Runcie et al. 2012). Indeeed Egloff (1966) found that high temperature greatly accelerates larval development in *T. californicus* over the tested range of 15-25 °C. In contrast, here we report that acute heat stress resulted in slower larval development (i.e., increased naupliar duration). This suggests that when larvae experience heat stress, they likely divert energy and resources from growth and development to stress responses, resulting in slower larval development. Prolonged larval duration as a result of acute thermal stress could be a factor affecting the population dynamics of *T. californicus*; Altermatt et al. (2012) suggested that the rapid temperature changes and increased risk of desiccation in smaller tidepools contribute to

increased rates of local extinction of *T. californicus*. Our results are consistent with this suggestion since more frequent acute heat stress in smaller pools will lead to increased mortality of all life stages and slower larval development, resulting in both slower population growth and more frequent local extinction. Given the short duration of our experimental treatment (i.e., 1 hour), the observed significant increase in larval duration is remarkable.

Finally, although larval duration was increased, a single bout of acute heat stress in the naupliar stage did not have a significant effect on adult size. Given that heat stress in the natural environment probably recurs on successive days during periods of warm weather and that stressful temperatures may frequently last more than one hour, additional testing is needed to definitively ascertain heat stress effects on adult size.

In conclusion, our results indicate that early life stages of *T. californicus* have higher thermal tolerance than adults from the same population. Although there are too few studies to make a strong generalization, we hypothesize that high thermal tolerance in larvae will be found in species where all life stages experience the same thermal environment. This is the case for brooding species, like *Crepidula fornicata* (Diederich & Pechenik 2013), and for species like *T. californicus* where larvae develop in the same tidepools inhabited by adults. Second, as previously observed in adult *T. californicus* populations, differences in larval thermal tolerance parallel the latitudinal gradient in environmental temperature; larvae from northern populations are more sensitive to high temperatures than those from southern populations, a pattern consistent with adaptation of populations to local environment variation. Finally, acute heat stress was found to affect fitness traits other than survivorship; even a single one-hour heat stress resulted in decreased developmental rates that could ultimately translate into reduced population growth rates.

Acknowledgements

We thank four anonymous reviewers who provided valuable suggestions and comments on the manuscript. Felipe Barreto, Ricardo Pereira, and Lani Gleason collected the *T. californicus* populations used in this experiment. Eric Allen, Martin Tresguerres, Felipe Barreto, Ricardo Pereira and Lani Gleason provided helpful comments on earlier versions of the manuscript. This research as supported in part by NSF (DEB 1051057 to RSB) and the Thai government for stipend and tuition support to ST.

Chapter 1, in full, was published in the Marine Ecological Progress Series: Tangwancharoen S, Burton RS (2014) Early life stages are not always the most sensitive: heat stress responses in the copepod *Tigriopus californicus*. Marine Ecology Progress Series. 517:75-83. The dissertation author was the primary investigator and author of this paper.

Table 1-1: Numbers of replicates from different life stages and populations used in control treatment and average survivorship proportions ± SE of nauplii from all populations at 20° C. p values show the results from Wilcoxon signed-rank test of the survivorship proportions between control treatment 20 °C and heat stress treatment 34 °C of the same population.

Stage	Population	Number of replicates	Number of clutches	Proportion survivorship	Wilcoxon test p value
Nauplius	PES	32	39	0.74±0.04	0.67
	SCN	29	29	0.56±0.05	0.91
	AB	27	34	0.76±0.12	0.21
	BR	39	52	0.69±0.04	0.71
	SD	38	61	0.72±0.04	0.57
	BUF	28	32	0.78±0.03	0.25
Copepodid	SD	10	10	0.93±0.03	0.65
Adult	SD	5	n/a	1±0	n/a

Table 1-2: ANOVA table showing the effects of temperature, life stage, and population factors on arcsine transformed survivorship proportions of *T. californicus* after acute heat stress

	Degrees of freedom	Sum of squares	F ratio	p value
Temperature	1	58.95	393.40	<0.0001
Life stage	2	7.16	23.90	<0.0001
Population	5	9.67	12.91	<0.0001
Residual	431	64.58		
Total	439	134.98		

Table 1-3: Results (p values) from Tukey-Kramer post hoc pairwise comparisons among the effects of life stages (A) and populations (B) on survivorship proportions

A

	nauplius	copepodid	adult
nauplius			
copepodid	<0.0001		
adult	<0.0001	0.92	

B

	PES	SCN	AB	BR	SD	BUF
PES						
SCN	0.87					
AB	0.0049	<0.0001				
BR	<0.0001	<0.0001	0.98			
SD	<0.0001	<0.0001	0.96	1		
BUF	0.0088	0.0001	1	0.96	0.93	

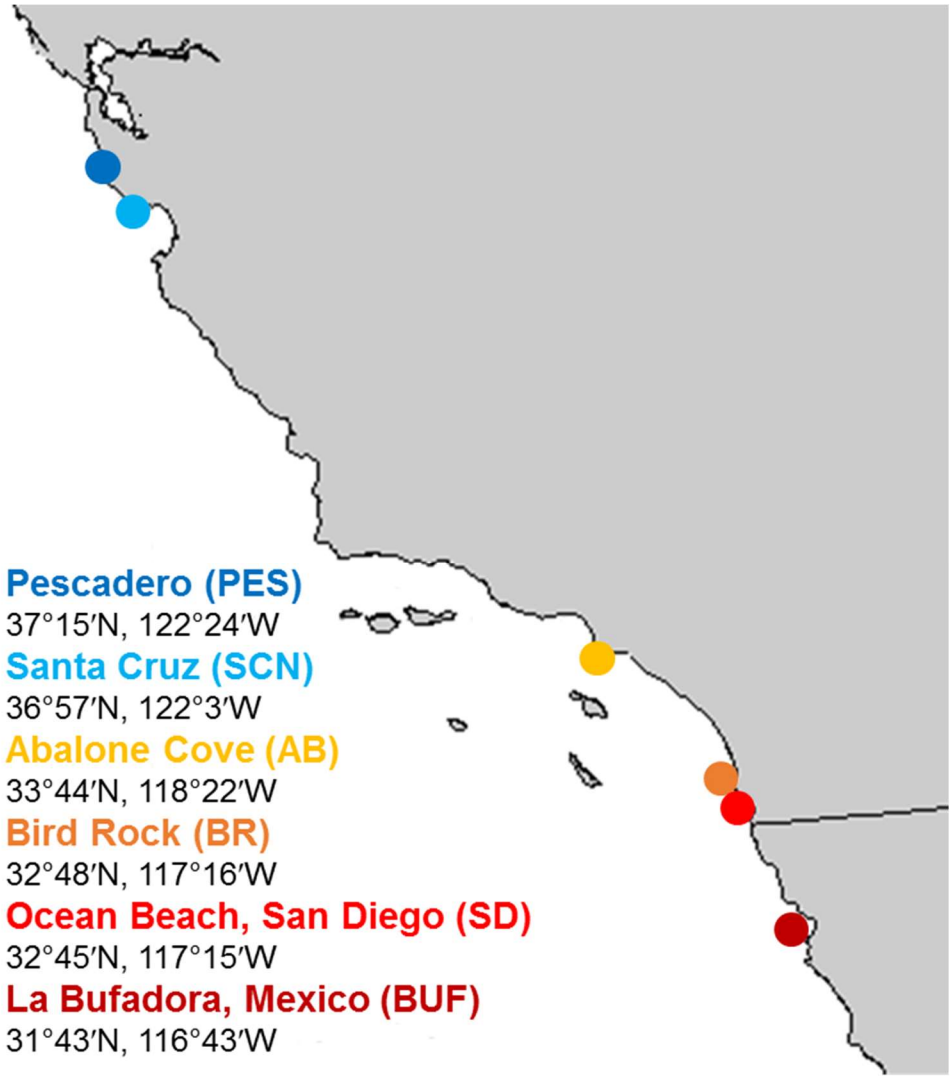


Figure 1-1: Locations (Pacific Coast, North America) of *Tigriopus californicus* populations used in the experiment.

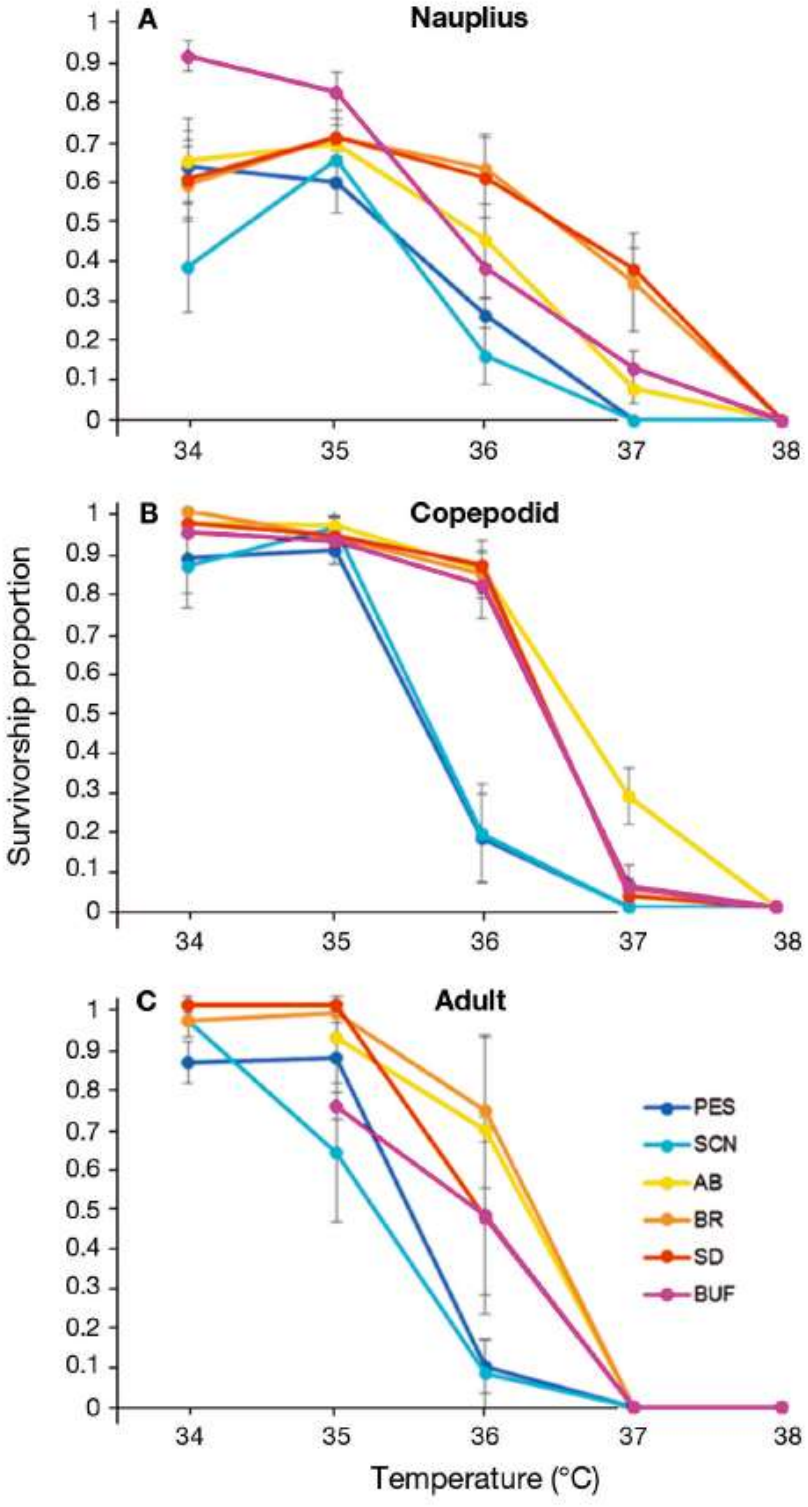


Figure 1-2: Average (±1 SE) survivorship proportions of 6 populations and 3 life stages of *Tigriopus californicus* following acute heat stress. Life stage: (A) nauplius, (B) copepodid, and (C) adult.

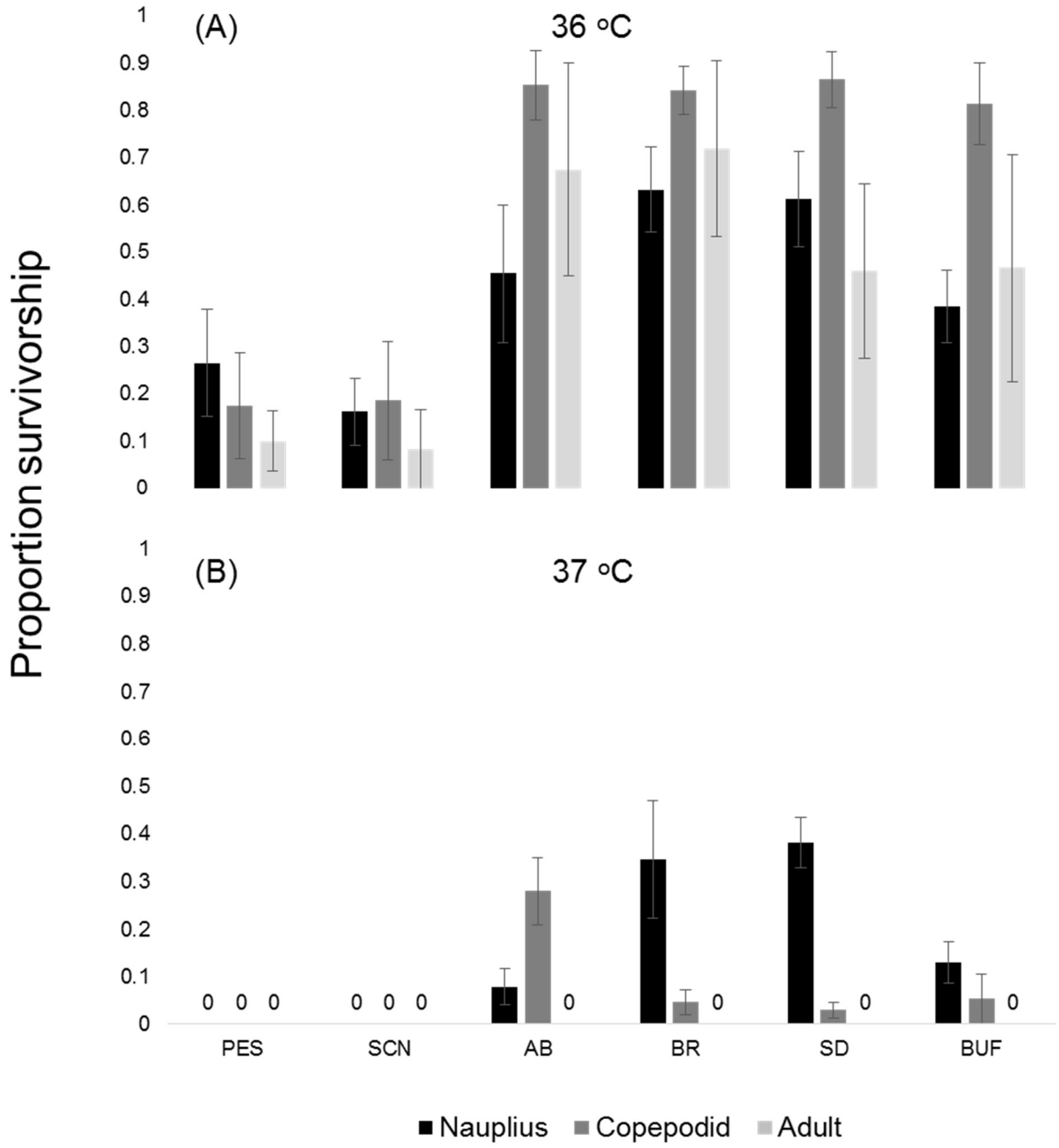


Figure 1-3: Average (±1 SE) survivorship proportions after acute heat stress of 3 *Tigriopus californicus* life stages at (A) 36 and (B) 37°C.

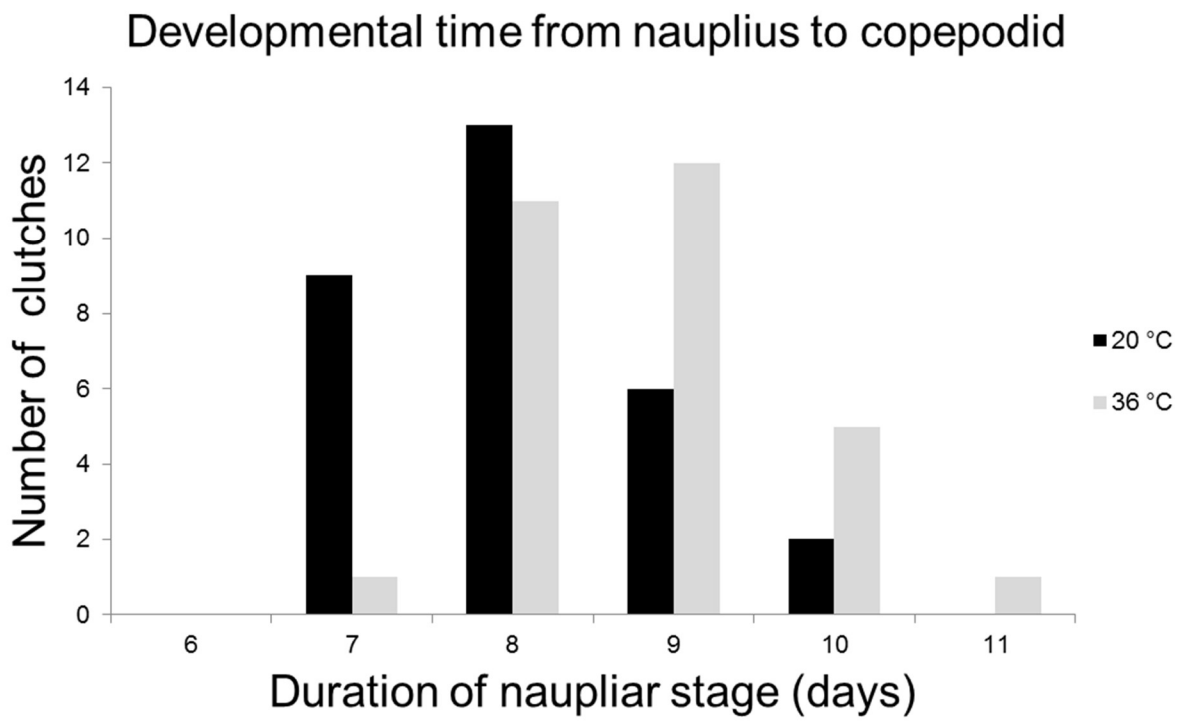


Figure 1-4: Effects of acute heat stress on larval development in *T. californicus*: The histograms show frequency of the number of days the copepods from San Diego population stayed in the naupliar stage before they metamorphosed into copepodids for both nauplii that were raised at 20 °C and nauplii that experienced acute heat stress at 36 °C.

Appendix

Table A1-1: Numbers of replicates, clutches, and total individuals used at each heat stress treatment.

A) Nauplius

Population	Replicates					Clutches					Individuals				
Temperature (°C):	34	35	36	37	38	34	35	36	37	38	34	35	36	37	38
PES	11	9	9	4	2	13	12	10	5	2	139	116	140	49	55
SCN	10	8	10	4	2	10	8	10	4	2	187	155	209	95	55
AB	5	6	7	8	2	7	6	8	12	2	53	72	106	110	42
BR	10	10	10	10	2	14	15	11	13	2	133	120	153	138	19
SD	10	10	10	10	2	14	18	12	18	3	126	133	140	139	30
BUF	6	6	5	9	2	8	6	7	9	2	76	85	87	115	23

B) Copepodid

Population	Replicates					Clutches					Individuals				
Temperature (°C):	34	35	36	37	38	34	35	36	37	38	34	35	36	37	38
PES	2	4	6	6	4	2	5	8	7	4	60	63	178	104	123
SCN	2	4	8	8	3	2	4	8	8	3	112	95	231	299	89
AB	2	3	7	7	4	2	3	9	8	4	41	95	102	128	72
BR	1	3	8	8	4	2	5	9	10	6	10	42	124	129	63
SD	2	4	12	8	4	2	5	15	10	4	38	47	189	140	86
BUF	2	4	6	6	4	2	4	6	6	4	38	58	138	124	51

C) Adult

Population	Replicates					Individuals				
Temperature (°C):	34	35	36	37	38	34	35	36	37	38
PES	5	4	5	1	1	50	40	50	10	10
SCN	5	5	6	1	1	50	50	60	10	10
AB	0	3	4	1	1	0	30	40	10	10
BR	5	5	5	1	1	50	50	50	10	10
SD	5	5	5	1	1	50	50	50	10	10
BUF	0	3	3	1	1	0	30	30	10	10

References

Altermatt F, Bieger A, Morgan SG (2012) Habitat characteristics and metapopulation dynamics of the copepod *Tigriopus californicus*. Marine Ecology Progress Series 468:85

Bermudes M, Ritar AJ (2008) Response of early stage spiny lobster *Jasus edwardsii* phyllosoma larvae to changes in temperature and photoperiod. Aquaculture 281:63-69

Bryars SR, Havenhand JN (2006) Effects of constant and varying temperatures on the development of blue swimmer crab (*Portunus pelagicus*) larvae: Laboratory observations and field predictions for temperate coastal waters. Journal of Experimental Marine Biology and Ecology 329:218-229

Burton RS, Feldman MW (1981) Population genetics of Tigriopus californicus. II. Differentiation among neighboring populations. Evolution:1192-1205

Burton RS, Lee B-N (1994) Nuclear and mitochondrial gene genealogies and allozyme polymorphism across a major phylogeographic break in the copepod *Tigriopus californicus*. Proceedings of the National Academy of Sciences USA 91:5197-5201

Dethier MN (1980) Tidepools as refuges: Predation and the limits of the harpacticoid copepod *Tigriopus californicus* (Baker). Journal of Experimental Marine Biology and Ecology 42:99-111

Diederich CM, Pechenik JA (2013) Thermal tolerance of *Crepidula fornicata* (Gastropoda) life history stages from intertidal and subtidal subpopulations. Marine Ecology Progress Series 486:173-187

Edmands S (2001) Phylogeography of the intertidal copepod *Tigriopus californicus* reveals substantially reduced population differentiation at northern latitudes. Molecular Ecology 10:1743-1750

Edmands S, Deimler JK (2004) Local adaptation, intrinsic coadaptation and the effects of environmental stress on interpopulation hybrids in the copepod *Tigriopus californicus*. Journal of Experimental Marine Biology and Ecology 303:183-196

Egloff DA (1966) Ecological aspects of sex ratio and reproduction in experimental and field populations of the marine copepod *Tigriopus californicus*. Stanford University

Ganz H, Burton R (1995) Genetic differentiation and reproductive incompatibility among Baja California populations of the copepod *Tigriopus californicus*. Marine Biology 123:821-827

Hilbish TJ, Mullinax A, Dolven SI, Meyer A, Koehn RK, Rawson PD (2000) Origin of the antitropical distribution pattern in marine mussels (*Mytilus* spp.): routes and timing of transequatorial migration. Marine Biology 136:69-77

JMP®, Version 11. 2013. SAS Institute Inc., Cary, NC, 1989-2007.

Kelly MW, Grosberg RK, Sanford E (2013) Trade-offs, geography, and limits to thermal adaptation in a tide pool copepod. The American Naturalist 181:846-854

Kelly MW, Sanford E, Grosberg RK (2012) Limited potential for adaptation to climate change in a broadly distributed marine crustacean. Proceedings of the Royal Society B-Biological Sciences 279:349-356

Koufopanou V, Reid DG, Ridgway SA, Thomas RH (1999) A molecular phylogeny of the patellid limpets (Gastropoda: Patellidae) and its implications for the origins of their antitropical distribution. Molecular Phylogenetics and Evolution 11:138-156

Miller N, Paganini A, Stillman J (2013) Differential thermal tolerance and energetic trajectories during ontogeny in porcelain crabs, genus Petrolisthes. J Therm Biol 38:79-85

Morgan S (1995) Life and death in the plankton: larval mortality and adaptation. Ecology of marine invertebrate larvae:279-321

Pereira RJ, Barreto FS, Burton RS (2014) Ecological novelty by hybridization: experimental evidence for increased thermal tolerance by transgressive segregation in *Tigriopus californicus*. Evolution 68:204-215

Peterson DL, Kubow KB, Connolly MJ, Kaplan LR, Wetkowski MM, Leong W, Phillips BC, Edmands S (2013) Reproductive and phylogenetic divergence of tidepool copepod populations across a narrow geographical boundary in Baja California. Journal of Biogeography 40:1664-1675

Pineda MC, McQuaid CD, Turon X, López-Legentil S, Ordóñez V, Rius M (2012) Tough adults, frail babies: an analysis of stress sensitivity across early life-history stages of widely introduced marine invertebrates. PLoS One 7:e46672

Plough LV, Marko PB (2014) Characterization of microsatellite loci and repeat density in the gooseneck barnacle, *Pollicipes elegans*, using next generation sequencing. Journal of Heredity 105:136-142

Preston N (1985) The effects of temperature and salinity on survival and growth of larval *Penaeus plebejus*, *Metapenaeus macleayi* and *M. bennettae*. National Prawn Seminar 2, Cleveland

Pörtner HO, Knust R (2007) Climate change affects marine fishes through the oxygen limitation of thermal tolerance. Science 315:95-97

Roberts SD, Dixon CD, Andreacchio L (2012) Temperature dependent larval duration and survival of the western king prawn, *Penaeus (Melicertus) latisulcatus* Kishinouye, from Spencer Gulf, South Australia. Journal of Experimental Marine Biology and Ecology 411:14-22

Rothlisberg PC (1979) Combined effects of temperature and salinity on the survival and growth of the larvae of *Pandalus jordani* (Decapoda: Pandalidae). Marine Biology 54:125-134

Runcie DE, Garfield DA, Babbitt CC, Wygoda JA, Mukherjee S, Wray GA (2012) Genetics of gene expression responses to temperature stress in a sea urchin gene network. Molecular Ecology 21:4547-4562

Schoville SD, Barreto FS, Moy GW, Wolff A, Burton RS (2012) Investigating the molecular basis of local adaptation to thermal stress: population differences in gene expression across the transcriptome of the copepod *Tigriopus californicus*. BMC Evolutionary Biology 12

Verween A, Vincx M, Degraer S (2007) The effect of temperature and salinity on the survival of *Mytilopsis leucophaeata* larvae (Mollusca, Bivalvia): The search for environmental limits. Journal of Experimental Marine Biology and Ecology 348:111-120

Walther K, Crickenberger SE, Marchant S, Marko PB, Moran AL (2013) Thermal tolerance of larvae of *Pollicipes elegans*, a marine species with an antitropical distribution. Marine Biology 160:2723-2732

Willett CS (2010) Potential fitness trade-offs for thermal tolerance in the intertidal copepod *Tigriopus californicus*. Evolution 64:2521-2534

Willett CS, Ladner JT (2009) Investigations of fine-scale phylogeography in *Tigriopus californicus* reveal historical patterns of population divergence. BMC Evolutionary Biology 9

Chapter 2:

F1 RNA-seq reveals regulatory divergence underlying thermal adaptation in the copepod *Tigriopus californicus.*

Abstract

Divergence in transcriptional expression has become widely accepted as a key mechanism for phenotypic evolution and local adaptation. Thermal tolerance studies have identified differential gene expression in response to heat stress, especially among Heat Shock Protein genes (HSPs), as a key factor explaining thermal adaptation in many systems. Yet, the genetic mechanism underlying differential gene expression underlying thermal adaptation remains poorly understood. Here we used two locally adapted *Tigriopus californicus* populations (heat tolerant San Diego, SD, and less tolerant Santa Cruz, SC) and their F1 hybrids to study evolution of transcriptional regulation by examining allele specific expression in response to heat stress. Overall, heat stress treatment elevated the number of genes with evidence for regulatory divergence between populations. The number of genes with *cis* regulatory divergence nearly doubled under heat stress conditions compared to no-stress controls. Genes displaying SD-dominant expression phenotypes increased in number in response to heat stress. These included genes involving in muscle assembly, cuticle formation, and chitin metabolic process. In response to heat stress, HSP genes showed the highest expression and the highest fold-change compared to controls. Among HSP genes, *cis* regulatory divergence resulted in elevated expression of the heat tolerant SD allele and higher total expression in SD population more frequently than HSPs that have expression bias toward SC population. In addition to HSPs, genes involved in electron transport revealed regulatory divergence under heat stress condition. Overall, this study provides insights on the genetic basis underlying evolution of thermal tolerance and local adaptation.

Introduction

Temperature is a major driving force of natural selection (Hochachka and Somero 2002; Kingsolver 2009). Populations of a species that are distributed across heterogeneous habitats will experience different temperature ranges and fluctuations, resulting in selection for differences in thermal tolerance across the species range. Organismal responses to acute thermal stress have been widely studied and often involve transcriptional responses (Lindquist 1986; Feder and Hofmann 1999). Among the many genes that respond to heat stress, increased transcription of Heat Shock Proteins (HSPs) has been documented across diverse taxa (Teranishi and Stillman 2007; Lang, et al. 2009; Lockwood, et al. 2010; Schoville, et al. 2012). Many comparative studies have shown that species or populations with different thermal tolerance express HSPs differently. Heat tolerant organisms either express more HSPs via "front-loading" in anticipation of heat stress (Barshis, et al. 2013; Gleason and Burton 2015) or via increased induction of HSP expression in response to heat stress (Tomanek and Somero 1999; Schoville, et al. 2012). Although HSPs have many cellular roles and consist of proteins from multiple families, their main role is to mitigate cellular damages from protein unfolding during heat stress. Different HSPs can bind to unfolded proteins to prevent aggregation, to help proteins fold back to their proper configurations, or to target unfolded proteins for proteolysis (Lindquist 1986; Parsell and Lindquist 1993; Feder and Hofmann 1999).

Numerous studies have shown that evolution in HSP gene expression underlies thermal tolerance across populations and species (Chen, et al. 2018) but only a few studies have been able to identify underlying causal changes at the genetic level (Lerman, et al. 2003; Lerman and Feder 2005). Aside from thermal adaptation, it is now widely accepted that evolution in

transcriptional regulation is a key mechanism of phenotypic evolution (Tautz 2000; Wray 2007; Wittkopp and Kalay 2012).

A simple model of gene expression explains transcriptional regulation as a product of interactions between *cis*-regulatory elements (eg. DNA sequence where a transcription factor binds) and *trans*-regulatory elements (eg. a transcription factor)(Wray, et al. 2003). Mutations that result in alteration in gene expression can occur in either *cis* or *trans* regulatory elements. Recent studies have developed a method to disentangle the genome-wide effects of *cis* and *trans* regulatory elements by looking at allele specific expression in F1 hybrids (Cowles, et al. 2002; McManus, et al. 2010). This method has led to a better understanding of the mechanisms underlying the evolution of gene expression phenotypes. Since F1 hybrids have both *cis*-alleles in the same cellular environment, the differences in allele specific expression are only *cis*-effects. The effects of *trans*-acting elements can be inferred by comparing F1 allele specific expression to the parental populations. The levels of gene expression in parental populations are the combination of both *cis* and *trans* regulatory elements (Wittkopp 2010; Coolon and Wittkopp 2013). Recent studies have adopted this this method to study the evolution of gene regulation in response to temperature stress (Chen, et al. 2015) and other different environmental conditions (Naranjo, et al. 2015; Lovell, et al. 2016; Mack, et al. 2016; Gould, et al. 2018).

Here we used the tidepool copepod *T. californicus* as a model to study population differences in gene regulation following thermal stress. Populations of *T. californicus* occur on isolated rocky outcrops along the Pacific coast of North America with no evidence of gene flow (Willett and Ladner 2009; Peterson, et al. 2013). This results in high level of genetic divergence between populations; transcriptome studies found a median coding sequence divergence between

San Diego (SD) and Santa Cruz (SC) populations of approximately 2% (Barreto, Pereira, et al. 2015). Yet, viable hybrids can be established and maintained under laboratory conditions.

As an intertidal organism, *T. californicus* often experiences acute heat stress due to temperature fluctuations in tidepools (Kelly, et al. 2012). Populations of *T. californicus* exhibit differential survivorship in response to acute heat stress, with a clear latitudinal cline suggesting local adaptation (Willett 2010; Kelly, et al. 2012; Tangwancharoen and Burton 2014; Leong, et al. 2018). SD and SC populations also differ in gene expression, particularly at HSP genes in response to acute heat stress, with most HSP genes showing greater upregulation in the more heat tolerant SD population compared to the less tolerant SC population (Schoville, et al. 2012). Because many genes participate in the response to thermal stress, isolating the effect of single highly upregulated genes requires a direct experimental approach. Barreto et al (2015) used RNAi to knock down the expression of Heat Shock Protein Beta 1 (HSPB1), one of the most upregulated genes after heat stress; knockdown resulted in decreased survivorship after heat stress, suggesting that even single HSP locus can have a major impact on thermal tolerance. These studies imply that divergence of transcriptional regulation is a key mechanism underlying evolution of thermal tolerance in *T. californicus*.

The goal of this study was to explore genome-wide basis of thermal adaptation focusing on the divergence in transcriptional regulation of acute heat stress responses between two locally adapted *T. californicus* populations (SD and SC populations). This was achieved by examining F1 and parental transcriptomes under both control and acute heat stress treatments. To identify both *cis* and *trans* regulatory divergence and modes of inheritance, I investigated allele specific expression in F1 hybrids and parental transcriptomes. I expected to identify regulatory

divergence under acute heat stress treatment to gain insight on the mechanistic basis of the evolution of thermal adaptations.

Materials and Methods

Copepod rearing and hybridization

Copepods were collected from high intertidal rock pools from SD (32° 45' N, 117° 15' W) and SC (36° 57' N 122° 03' W) locations. Stock cultures of the populations were maintained in 400 ml beakers filled with 250 ml filtered (0.45 μm) seawater under constant 20 °C with 12-hour photoperiod. The animals were fed ground *Spirulina* wafer *ad libitum*. F1 hybrids were produced and maintained following a previous study (Barreto, Pereira, et al. 2015). Two replicates of F1 hybrids were made using independent sets of crosses between SD and SC populations.

Heat stress experiment and RNA sequencing

For each sample in a treatment, 50 adult copepods of equal sex ratio (25 males and 25 non-gravid females) were collected from the two parental populations and the two reciprocal F1 hybrids. The animals were put in 1.5 ml Eppendorf tubes filled with 0.5 ml filtered sea water. For the heat stress treatment, the tubes were first submerged in a water bath at 20 °C. Temperature was increased by 5 °C per hour until it reached 35 °C and then held at 35 °C for 1 hour. For control treatment, the copepods were similarly treated by transfer into 1.5 ml tubes, but they remained in 20 °C incubator for the whole duration of the heat stress treatment. Copepods were immediately sacrificed by homogenization in Tri Reagent (Sigma) for RNA extraction. RNA was extracted following the manufacturer's protocol. All four genotypes and two

treatments were performed at the same time as one batch. Each treatment and genotype had two biological replicates. cDNA libraries were prepared using oligo-dT priming and Illumina's Truseq standard mRNA protocol. All 16 libraries were sequenced (100-bp single-read) on a single lane on the Illumina HiSeq4000 platform. The 100-bp sequencing length was chosen to be able to differentiate population specific alleles based on 2.04% median coding sequence divergence between SD and SC population (Barreto, Pereira, et al. 2015).

RNA mapping and filtering

Reads were aligned to a transcriptome reference (Barreto, Pereira, et al. 2015) that included paired orthologs (15747 nuclear contigs and 13 mitochondrial contigs) from both populations using CLC Genomics Workbench 11. Parameters for mapping included a cutoff at 0.8 length fraction and 0.98 sequence similarity based on approximately 2% median divergence in coding sequences (Barreto 2015). Two types of read counts were obtained from mapping including unique reads and total reads. The number of sequences that only mapped to one of the two alleles were counted as "unique reads". Unique reads from both alleles were combined with the number of reads that mapped equally to either allele to make "total reads".

For unique reads,13 mitochondrial genes and contigs with low raw reads count (average < 20 reads) were filtered out from future analysis. We examined parental libraries and filtered out genes that were not assigned to the correct populations. For example, for an SC library, contigs that had more reads mapped to the SD allele than the SC allele were filtered out. Contigs which had reads mapped to both SD and SC but binomial test results were not significant in any parental library were also removed from further analyses. The final 9679 nuclear contigs were used for further allele specific expression analyses.

For total reads, contigs with low read counts (counts per million, CPM < 2 in 12 out of 16 libraries) were removed from further analyses. 10454 nuclear contigs and 13 mitochondrial contigs were kept for further analyses. The list of 9679 nuclear contigs of unique reads completely overlaps with the list of 10454 nuclear contigs of total reads.

Differential expression and modes of inheritance analyses

Total reads were used to identify differential expressed genes between control and heat stress treatments. Total reads were analyzed using package edgeR (Robinson, et al. 2010; Lund, et al. 2012) in R statistical environment (R core team). Reads were TMM normalized before fitting negative binomial generalized linear models. Significant values were obtained by quasi-likelihood F test followed by 5% FDR. Analyses were performed separately for each parental line and F1; both reciprocal hybrid crosses were combined as F1 (4 replicates) while each parental line has 2 replicates.

To infer modes of inheritance, total reads were compared between F1 and each parental line at each temperature treatment separately. Significant values were obtained the same way as tests for differential expressed genes between temperature treatments described above. Modes of inheritance were assigned using three criteria (Table 2-1) following previously suggested classifications (McManus, et al. 2010; Schaefke, et al. 2013; Chen, et al. 2015). Although previous studies used both significant test and 1.25-fold cut-off threshold, only significant values were used as a threshold to classify modes of inheritance in this study.

Allele specific expression and cis-trans regulatory assignment

In order to classify patterns of regulatory divergence, three statistical testes were used to identify divergence in *cis* and *trans* regulatory elements (McManus, et al. 2010). First was the

test for allele specific expression in F1 hybrids using unique reads (Test H). Binomial tests (H_0: F1 allelic expression SD = SC) with 5% FDR were performed on each gene in each replicate separately. A gene would be considered having significant biased expression only when both replicates of a reciprocal cross in a treatment tested significant for the same direction of the allele expression. Significant results for Test H infer divergence in *cis* regulatory elements. The next test was the test for differential expression between two parental populations (Test P: H_0: SD(total) = SC(total)). Significant values were obtained using total reads comparing differential expression between SD and SC parents for each temperature treatment the same way as tests for differential expressed genes described above. The last test (Test T) compared the ratio between allelic expression in F1 to the ratio of gene expression between two parents (H_0: F1(SD/SC) = SD(total)/SC(total)). This test identifies divergence in *trans* regulatory elements. Significant values were obtained with Fisher's Exact tests following by 5% FDR. Since Fisher's test only takes integers, rounded average normalized counts were used for this analysis. CPM of unique reads were used for F1 hybrids while CPM of total reads were used for F0 parents.

Divergence in *cis* and *trans* regulatory elements between the two populations was identified following three significant tests and classification criteria (Table 2-2) previously suggested by McManus et al. (2010). Conserved genes did not show any evidence for divergence between populations or for divergence in *cis* and *trans* regulatory elements: all three tests are not significant. Ambiguous genes show some evidences of divergence through significant tests in one of the three tests, but they could not be confidentially classified into any of *cis* and *trans* combinations since they do not fit the classification criteria. Compensatory genes are the genes that have *cis* and *trans* regulatory elements favoring expression in the opposite directions so that the total expression levels remain conserved between populations. Thus, they test significant for

Test H (*cis* effect) and Test T (*trans* effect) but not Test P. Regulatory divergence of *cis x trans* genes are similar to compensatory genes but *cis* and *trans* effects do no completely offset each other so that the expression levels are not conserved between populations. On the other hand, *cis+trans* genes have *cis* and *trans* regulatory elements favoring the same allele. Genes that diverge only in the *cis* regulatory elements would have biased expression in the F1 hybrid with the same direction in parental differential expression (Test P) and similar allele specific expression ratio between F1 and parents (Test T not significant). Lastly, *trans* only genes would have significant differential expression between populations (Test P) and significant difference in ratio between F1 allelic expression and parental expression (Test T) but do not have significant biased expression in F1 hybrids (Test H).

Tests for maternal effects

Both unique and total reads were compared between reciprocal crosses for each treatment to see whether the direction of reciprocal crosses affects total gene expression and allele specific expression. To test for maternal effects on the total gene expression, total reads were compared between two reciprocal crosses for each treatment the same way as the analyses for differential and modes of inheritances described above. To test for maternal effects on allele specific expression, allelic expression ratios of two reciprocal crosses were compared (H_0: SCf(SD/SC) = SDf(SD/SC)). Significant values were obtained with Fisher's Exact tests on rounded average CPM between two reciprocal crosses similar to Test T described above.

Functional group enrichment

Lists of statistically significant gene sets from the analyses above were further tested for functional group enrichment using BLAST2GO software. Statistical values for overrepresented Gene Ontology (GO) terms were obtained using Fisher's exact tests followed by 5% FDR.

Results

Illumina sequencing and read mapping

The numbers of raw reads and mapped reads (both unique and total reads) from each library are listed on Table 2-3. Each library's raw reads range from 19.5 to 28.9 million reads. Unique read mapping ranges from fewer than 2 million reads to more than 10 million reads (less than 9% to more than 37%). Both biological replicates of each population and treatment showed high degree of positive correlation to each other (Table 2-4).

Differential gene expression in response to heat stress

Differential gene expression analyses identified significant up and down regulated genes in F1 and parental lines (Figure 2-1). F1 hybrids have more significantly up-regulated genes (212) than both parental lines (SD:133, SC:29). Although up-regulated genes in F1 did not test significant for functional group enrichment, overlapped up-regulated genes between SD and F1 and up-regulated genes in SD tested significant for functional group enrichment. Both lists tested significant for enrichment in two categories: response to stress (GO:0006950) (FDR p: 0.001 in overlapping SD and F1 and FDR p: 0.02 in SD) and unfolded protein binding (GO:0051982) (FDR p: 0.002 in overlapping SD and F1 and FDR p: 0.02 in SD). HSP genes are listed within both GO categories. Although there is not any significant test for any functional group enrichment, there are 9 HSPs among 29 up-regulated genes in SC. HSPs are among the highest

expressed gene in response to heat stress in all F1 hybrid and parents (Figure 2-2). There are fewer down-regulated genes than up-regulated genes in all hybrids and parents. Most of them are population specific. Only 2 genes are overlapped between F1 and SD and 1 gene overlapped between SD and SC. All the lists of down-regulated genes in all populations did not test significant for functional group enrichment.

Modes of Inheritance

Comparisons of gene expression between F1 hybrids and parents (based on total read counts) reveals how expression phenotypes are inherited (Figure 2-3). The numbers of contigs assigned to modes of inheritance are listed on Table 2-5. The majority of the contigs are classified as conserved which indicates that there is not a significant difference between hybrid and parental gene expression. Also, most of the conserved contigs overlap between the two treatments. However, under heat stress treatment, more genes showed different modes of inheritance. Although the numbers of additive genes are low in both treatments, 624 contigs under the control treatment and 974 contigs under heat stress treatment are potentially additive genes. They show F1 expression levels intermediate between two parental expression levels but statistical tests were not significant. The most striking difference between control and heat stress treatments is the number of SD dominant genes where F1 hybrid expression is more similar to the SD parent than the SC parent. Only 100 contigs in the control treatment tested significant for SD dominant while 907 genes tested for SD dominant under heat stress. Lists of contigs in each mode of inheritance and treatment were tested for functional group enrichment. However, only SD dominant genes under heat stress condition show significant results (Table 2-6).

Regulatory divergence

In order to classify patterns of regulatory divergence, three statistical tests were employed. First, significant allelic imbalance in F1 hybrids (Test H) indicates divergence in *cis* regulatory elements between two populations. The numbers of biased genes in each allelic direction are similar in both treatments. Under control treatment, there are 367 contigs that showed elevated expression of the SD allele and 328 contigs that showed elevated expression of the SC allele in both reciprocal crosses. Heat stress treatment increased the numbers of biased genes in both directions and crosses with 549 SD biased genes and 551 SC biased genes. Fisher's Exact tests for functional group enrichment only revealed that SD biased genes under control treatment are enriched in organic cyclic compound binding (GO: 0097159) (FDR p = 0.049) and heterocyclic compound binding (GO: 1901363) (FDR p = 0.049) while no significant enrichment was observed for SC biased genes under control treatment. For heat stress treatment, SC biased genes are enriched in many metabolic pathway functions (Table 2-6) while SD biased genes did not test significant for enriched functional groups. However, SD biased genes under heat stress include the highest expression HSPs, including small HSPs (HSPB1) and HSPs 70 (Table 2-7, Figure 2-2).

Next, significant Test P results indicate the divergence in expression phenotypes between the two parental populations. Under control treatment, 343 genes were more highly expressed in the SC population while 414 genes were more highly expressed in the SD population. Neither gene group was significantly enriched in any functional groups. Heat stress treatment increased the number of genes that were differentially expressed between the two populations. The SC population showed elevated expression of 672 genes compared to SD, while SD showed elevated expression of 923 genes compared to SC. Of the total of 44 remaining HSPs genes in final list of 9679 contigs analyzed, 3 HSPs were more highly expressed in SC and 9 HSPs were more highly

expressed in SD. Fisher's Exact tests revealed electron carrier activity (GO: 0009055) (FDR p: 0.01) as the only significant enriched functional group among genes with elevated expression in SD under heat stress. There was no significant test for genes that expressed more in SC under heat stress.

Lastly, differences between F1 allele specific expression ratio and parental expression ratio (Test T: F1(SD/SC) = SD (total)/SC (total)) imply divergence in *trans* regulatory elements. Under the control treatment, 152 genes tested significant for divergence in *trans* regulatory elements. Similar to Test H and Test P, the number of genes showing divergence in *trans* regulatory element in both reciprocal crosses also increased under heat stress treatment (385 genes). Only genes with divergence in *trans* regulatory elements under heat stress are enriched in 3 GO categories, including extracellular matrix structural constituent (GO: 0005201) (FDR p: 0.004), collagen trimer (GO: 0005581) (FDR p: 0.005), and aminoglycan metabolic process (GO: 0006022) (FDR p: 0.04).

With all three test results combined, each contig could be classified into a pattern of regulatory divergence (Figure 2-4, Table 2-8). The majority of genes in both treatments are conserved (no significant differential expression between parental populations and no biased allelic expression in hybrid). However, under heat stress treatment in both reciprocal hybrid crosses, more genes show evidence of divergence in both *cis* and *trans* regulatory elements in every category. Among these genes, many were classified as ambiguous. They showed some evidence of regulatory divergence, but they did not fit all the criteria for classification. Among the genes that could be classified, more genes show divergence in *cis* regulatory elements than *trans* regulatory elements in both treatments. The number of genes showing divergence in *cis* regulatory elements nearly doubles under heat stress treatment comparing to control treatment.

Under control treatment, Fisher's Exact tests did not identify significance of functional group enrichment in any *cis* and *trans* combinations. On the other hand, genes that diverge only in *cis* regulatory elements in both crosses under heat stress treatment tested significant in a few GO categories (Table 2-6). Within this group, genes where *cis* regulatory elements favor the SD allele are enriched in electron carrier activity (GO: 0009055) (FDR p: 0.01) and cofactor binding (GO: 0048037) (FDR p: 0.01). There was no significant enrichment of functional groups in other *cis* and *trans* patterns under the heat stress treatment.

Maternal effects

The direction of reciprocal hybrid crosses did not seem to have a strong effect on both total RNA expression or allele specific expression in F1 hybrids in any of the treatments. Only two unannotated contigs tested significant for differential expression between reciprocal crosses under control treatment and one unannotated contig under heat stress conditions. For allele specific expression, only 3 contigs from the control treatment and 4 contigs from the heat stress treatment (2 contigs overlapped) of the 9679 contigs tested significant in different allelic ratios between reciprocal crosses.

Discussion

Analysis of allele specific expression allowed us to explore genome-wide patterns of regulatory divergence and to identify candidate genes contributing to the evolution of thermal tolerance in *T. californicus*. Overall, our results revealed more genes showing evidence of regulatory divergence under the heat stress treatment compared to the control treatment. This suggests that differences in thermal environment are a major driving force for divergence

between these two *T. californicus* populations. Regulatory divergence between populations can be identified by examining differentially expressed genes between genotypes and allelic imbalance in F1 hybrid expression.

Comparisons of mRNA expression between hybrids and parents reveal how divergent expression phenotypes are inherited. Genes that diverge in expression phenotype can be either additive or dominant. The most pronounced result with regard to modes of inheritance is that the number of SD dominant genes increased in response to heat stress. This list of genes is enriched in GO categories involved in muscle assembly, cuticle structure, and chitin metabolism (Table 2-6). Since elevated temperature leads to protein unfolding (Somero 1995; Gidalevitz, et al. 2011), heat stress could disrupt muscle structure and function. For example, heat has been shown to cause muscle atrophy in intertidal shrimp (Madeira, et al. 2015), so it is possible that *T.californicus* experienced similar symptoms during heat stress. The mechanisms involved in heat tolerance in the SD population might also require evolution in genes involved in muscle and skeletal structure (chitin exoskeleton in arthropods) formation or remodeling. Another study found downregulation of several chitinase genes in response to heat stress in mussels (Negri, et al. 2013), and Jovic et al. (2017) found that heat stress affected differential expression of cuticle related genes in *Caenorhabditis elegans* (despite different cuticle structure). In the *C. elegans* model, cuticle genes are also regulated by Heat Shock Transcription Factor (HSF) (Brunquell, et al. 2016). These observations in other systems suggest the important role of the structural remodeling process in response to heat stress.

In a previous *T.californicus* transcriptomic study, Schoville et al. (2012) found that the heat tolerant SD population upregulated several cuticle related genes while the SC population downregulated the same genes and suggested they might play a role in local adaptation. In this

study, modes of inheritance analysis confirmed differential expression of these genes between populations. However, these genes were not differentially expressed between temperature treatments. Also, there was not any consistent pattern of *cis* and *trans* regulatory divergence, but dominant inheritance suggest contributions of *trans* regulatory divergence since *cis* regulatory divergence would result in additive inheritance (Lemos, et al. 2008).

Previous studies in *T. californicus* also suggested a crucial role of HSPs in mediating responses to heat stress (Schoville, et al. 2012; Barreto, Schoville, et al. 2015). Among 212 significantly upregulated genes in F1 hybrids, 19 are HSPs genes (out of 44 HSPs in the 9769 unique gene list). Within the 19 HSPs genes, 5 expressed more SC allele and 7 express more SD allele (Table 2-7). All HSPs with biased expression toward SC allele do not differ in expression between SD and SC populations. The allelic imbalance indicates the divergence in *cis* regulatory elements but similar expression between populations suggest divergence in *trans* regulatory elements that compensate changes in *cis* elements to conserve expression phenotypes across populations (although not all 5 genes tested significant for *trans* effect). Among all upregulated HSPs, none of them significantly express more in the SC population than SD. On the other hand, 4 out of 7 HSPs with allelic imbalance toward SD allele (Figure 2-2B) also showed elevated expression in the SD population relative to the SC population. Overall, few HSPs are differentially expressed between populations. Conservation of HSP expression could result from the energetic cost of excessive HSP expression that can lead to negative consequences (Krebs and Feder 1997, 1998; Hoekstra and Montooth 2013). Therefore, HSPs that have higher expression in more heat tolerant populations imply benefits of expressing more transcripts and could be potential candidate genes for further investigations to see whether these HSPs contribute to increased heat tolerance. In addition, 5 out of 7 SD biased HSPs have biased

expression only under heat stress treatment. Among all the HSPs, HSPB1 is the only HSP gene which upregulates in F1 and both populations, has biased toward heat tolerant SD allele only under heat stress treatment, and has divergence in expression phenotype (increased upregulation in SD population than SC in response to heat stress). This HSPB1 had the highest magnitude of upregulation and was one of the most highly expressed small HSPs. Even though there are many small HSPs in the genome, knocking down the mRNA of this HSPB1 led to decrease in survivorship (Barreto, Schoville, et al. 2015). This confirms the necessity of this HSPB1 expression in heat stress response and suggests its crucial role in the evolution of thermal tolerance in *T. californicus*. Allelic imbalance under heat stress treatment suggests that HSPB1 has divergence in the *cis* regulatory elements that might be in the binding site of HSF. HSF is known to get activated by heat stress and induce expression of HSP genes in many eukaryote systems (Wu 1995; Pirkkala, et al. 2001). A subsequent study has confirmed divergence in *cis* regulatory elements as latitudinal cline of SNPs in Heat Shock Elements (HSEs), HSF binding sites, across *T. californicus* populations (Tangwancharoen, et al. 2018).

Besides HSPB1, E3 ubiquitin-protein ligase (contig4499) is the only gene in the genome that matches the same criteria. This gene is in ubiquitin proteolytic pathway. Some of these genes work directly with chaperones to degrade unfolded proteins from heat stress damage (Ciechanover 1998; Labbadia and Morimoto 2015). There are 71 E3 ubiquitin-protein ligase genes in *T. californicus* genome, but only 3 homologs are upregulated in hybrids. Similar to some HSPs, E3 ubiquitin-protein ligase could be transcriptionally regulated by HSF (Pirkkala, et al. 2000). Although there are many E3 ubiquitin-protein ligase copies, the few heat stress inducible homologs could be good candidates for further investigations on the mechanisms underlying thermal tolerance.

In addition to HSPs and the ubiquitin pathway genes discussed above, 923 genes were significantly expressed higher in heat tolerant SD than in SC populations during heat stress. However, gene ontology analysis only found significant enrichment in the electron carrier functional group (GO: 0009055) with 19 genes. Nine of these are Cytochrome P450 (CYP) genes. The role of CYP genes in physiological response to heat stress is unclear especially in marine organisms. Some studies found heat stress suppressed CYP expression (Rosic, et al. 2010; Wang, et al. 2016). Others found that high temperature can be associated with increased expression of CYP (Andersson and Forlin 1992; Vergara-Amado, et al. 2017). In *T. californicus*, none of the CYP were upregulated in response to heat stress and almost all CYP genes that SD expressed more than SC in under heat stress also expressed more in SD under control condition. Another possible explanation is that this regulatory divergence is not related to heat tolerance. One of CYP functions is detoxification of xenobiotic compounds (Snyder 2000; Anzenbacher and Anzenbacherova 2001) and F1 hybrid expressed more SD allele for organic compound binding proteins (GO: 0097159 and GO: 1901363) under control condition, potentially suggesting evolution in response to pollutants in the habitats.

Moreover, under heat stress treatment, F1 SC biased genes are significantly enriched in energy metabolism (Table 2-6) including glycolysis and Krebs cycle genes such as 6-phosphofructo-2-kinase (PFK), glyceraldehyde-3-phosphate dehydrogenase (GAPDH), pyruvate kinase (PK) and isocitrate dehydrogenase (IDH). Some of these genes including PFK, GAPDH and IDH also have multiple copies in the genome. This is an interesting observation even though the ecological relevance is unclear. Most of these genes were not significantly differential expressed between treatments or between parental populations suggesting that the expression phenotypes are conserved across treatments and populations. Allelic imbalance in F1 despite

conserved expression across populations suggests that interacting *cis* and *trans* regulatory elements diverge in the opposing directions resulting in compensatory expression. Also, allele specific expression pattern did not differ between reciprocal crosses. This suggests that these patterns did not result from mismatch in mitochondrial genotype.

Conclusion

The goal of this study was to identify the genetic basis underlying local adaptation and evolution of thermal tolerance. Schoville et al. (2012) identified HSPs, proteolysis pathways and cuticle related genes as groups of genes with the most pronounced difference in expression between SD and SC populations. Our results also confirmed the important role of these genes in heat stress response and local adaptation in *T. californicus* and revealed how the expression of these genes diverged. Divergence in transcriptional regulation was more pronounced under heat stress treatment. The number of SD dominant expression phenotypes increased with heat treatment. Many of these genes are related to building body structure including muscle organization and chitin metabolic processes. Divergence in *cis* regulation was more common than divergence in *trans* regulation. Among *cis* divergence genes are key genes involving in heat shock response, including HSPs and genes in ubiquitin proteolytic pathway. Our study provides candidate genes and pathways for further investigations on the genetic basis of local adaptation and provides some insights of the molecular mechanisms underlying thermal tolerance.

Acknowledgements

Chapter 2, in full, is being prepared for submission: Tangwancharoen S, Burton RS. F1 RNA-seq reveals regulatory divergence underlying thermal adaptation in the copepod *Tigriopus californicus*. The dissertation author was the primary investigator and author of this paper.

Table 2-1: List of all three criteria for assigning modes of inheritance including tests for differential expressed genes between F1 and each parental line ("Yes" means test significant and "No" means not significant) and a criterion for pattern of expression.

Modes of Inheritance	F1 vs SD	F1vs SC	Pattern of expression
Conserved	No	No	
Additive	Yes	Yes	SD>F1>SC or SC>F1>SD
Overdominance	Yes	Yes	F1 > both SD and SC
Underdominance	Yes	Yes	F1 < both SD and SC
SD Dominance	No	Yes	
SC Dominance	Yes	No	

Table 2-2: List of criteria for regulatory divergence pattern classification. “Yes” means statistical test significant on rejecting null hypothesis and “No” means fail to reject the null hypothesis. Contigs that do not any of these categories were assigned as “ambiguous”

Categories	Test H	Test P	Test T	Other criteria
Conserved	No	No	No	
cis only	Yes	Yes	No	
trans only	No	Yes	Yes	
cis+*trans*	Yes	Yes	Yes	*cis* and *trans* favor the same allele
*cis*x*trans*	Yes	Yes	Yes	*cis* and *trans* favor opposite alleles
Compensatory	Yes	No	Yes	

Table 2-3: The numbers of both unique and total mapped reads in each replicate of both treatments and populations

Population	Replicate	Treatment °C	#Raw (M)	#Unique (M)	%Unique	#Total (M)	%Total
SCfxSDm	1	20	22.8	1.99	8.72	2.46	10.75
		35	26.6	6.07	22.82	7.36	27.68
	2	20	27.9	10.11	36.19	12.47	44.67
		35	28.9	10.75	37.16	13.11	45.34
SDfxSCm	1	20	23.1	3.50	15.15	4.32	18.70
		35	23.6	6.29	26.62	7.69	32.54
	2	20	27.7	9.98	36.06	12.26	44.24
		35	21.0	7.24	34.50	8.74	41.65
SD	1	20	26.3	3.56	13.53	4.35	16.53
		35	21.3	6.93	32.50	8.36	39.18
	2	20	22.9	7.86	34.33	9.67	42.20
		35	19.5	7.10	36.46	8.64	44.37
SC	1	20	25.9	9.04	34.85	11.03	42.54
		35	22.6	8.00	35.41	9.75	43.11
	2	20	21.1	6.87	32.56	8.42	39.91
		35	23.2	8.24	35.48	9.96	42.86

Table 2-4: Correlations between replicates for both unique and total reads in each population and treatment.

	Population	Treatment	Pearson's r	95% confident interval		p value
				Lower bound	Upper bound	
Unique	SCfxSDm	control	0.8112	0.8063	0.8160	<0.0001
	SCfxSDm	heat stress	0.9567	0.9555	0.9579	<0.0001
	SDfxSCm	control	0.8239	0.8193	0.8284	<0.0001
	SDfxSCm	heat stress	0.9623	0.9613	0.9634	<0.0001
Total	SCfxSDm	control	0.9307	0.9280	0.9332	<0.0001
	SCfxSDm	heat stress	0.9209	0.9179	0.9237	<0.0001
	SDfxSCm	control	0.9312	0.9286	0.9337	<0.0001
	SDfxSCm	heat stress	0.9725	0.9715	0.9735	<0.0001
	SD	control	0.9713	0.9702	0.9723	<0.0001
	SD	heat stress	0.9677	0.9665	0.9689	<0.0001
	SC	control	0.9667	0.9654	0.9679	<0.0001
	SC	heat stress	0.9899	0.9895	0.9902	<0.0001

Table 2-5: The number of genes in each mode of inheritance under each treatment and overlapped genes between treatments.

Modes of Inheritance	Control	Heat Stress	Overlapped
Conserved	10218	9404	9334
Additive	16	28	9
SD dominance	100	907	71
SC dominance	132	127	63
Overdominance	1	1	1
Underdominance	0	0	0

Table 2-6: Significant over-represent enrichment gene ontology (GO) categories.

Conditions	GO ID	GO Category:Term	FDR	P-Value
Heat stress	GO:0006022	B: aminoglycan metabolic process	0.0037	4.00E-07
SD dominant	GO:0006040	B: amino sugar metabolic process	0.0361	1.04E-05
	GO:0006030	B: chitin metabolic process	0.0361	1.18E-05
	GO:0030241	B: skeletal muscle myosin thick filament assembly	0.0395	5.60E-05
	GO:1901071	B: glucosamine-containing compound metabolic process	0.0395	2.09E-05
	GO:0070986	B: left/right axis specification	0.0395	5.60E-05
	GO:0014866	B: skeletal myofibril assembly	0.0395	5.60E-05
	GO:0060972	B: left/right pattern formation	0.0395	5.60E-05
	GO:0007527	B: adult somatic muscle development	0.0395	5.60E-05
	GO:0007496	B: anterior midgut development	0.0395	5.60E-05
	GO:0000146	M: microfilament motor activity	0.0395	5.60E-05
	GO:0031036	B: myosin II filament assembly	0.0395	5.60E-05
	GO:0031038	B: myosin II filament organization	0.0395	5.60E-05
	GO:0060361	B: flight	0.0422	6.45E-05
	GO:0006023	B: aminoglycan biosynthetic process	0.0436	7.14E-05
	GO:0035017	B: cuticle pattern formation	0.0472	8.75E-05
	GO:0005703	C: polytene chromosome puff	0.0472	8.75E-05
Heat stress	GO:0044710	B: single-organism metabolic process	0.0001	1.51E-08
F1 SC biased	GO:0003824	M: catalytic activity	0.0006	1.32E-07
Both reciprocal crosses	GO:0008152	B: metabolic process	0.0033	1.10E-06
	GO:0005975	B: carbohydrate metabolic process	0.0080	3.56E-06
	GO:0044281	B: small molecule metabolic process	0.0243	1.35E-05
	GO:0016491	M: oxidoreductase activity	0.0340	2.64E-05
	GO:0006022	B: aminoglycan metabolic process	0.0340	2.49E-05
	GO:0006629	B: lipid metabolic process	0.0383	3.40E-05
	GO:0006766	B: vitamin metabolic process	0.0497	6.62E-05
	GO:0006767	B: water-soluble vitamin metabolic process	0.0497	5.07E-05
	GO:0032787	B: monocarboxylic acid metabolic process	0.0497	7.16E-05
	GO:0071704	B: organic substance metabolic process	0.0497	6.06E-05
	GO:0044283	B: small molecule biosynthetic process	0.0497	7.73E-05
	GO:0006040	B: amino sugar metabolic process	0.0497	7.41E-05

GO category: B = biological process, C = cellular component, M = molecular function

Table 2-6: Significant over-represent enrichment gene ontology (GO) categories, *continued*

Conditions	GO ID	GO Category:Term	FDR	P-Value
Heat stress	GO:0009055	M: electron carrier activity	0.0025	3.05E-07
cis only divergence	GO:0003824	M: catalytic activity	0.0025	5.52E-07
Both reciprocal crosses	GO:0006022	B: aminoglycan metabolic process	0.0074	2.46E-06
	GO:0006118	B: obsolete electron transport	0.0099	4.39E-06
	GO:0006040	B: amino sugar metabolic process	0.0184	1.02E-05
	GO:0009055	M: electron carrier activity	0.0025	3.05E-07

GO category: B = biological process, C = cellular component, M = molecular function

Table 2-7: All HSPs that are significantly up-regulated under heat stress treatment in F1 hybrids with differential expression and regulatory divergence test results. In the significant tests for regulatory divergence, SD and SC represent significant biased expression toward the respective allele (Test H) in F1 hybrids or more total expression in the respective parental population (Test P). In differential expression, up and down mean significant up or down-regulation under heat stress treatment. NS means test results not significant.

Contig	Annotation	HSP	Log_2	Log_2	Test P		Allelic imbalance (Test H)				Differential	
number		gene	fold	CPM	Control	Heat	Control		Heat stress		expression	
		family	change	heat stress		stress	SCfxSDm	SCfxSDm	SCfxSDm	SCfxSDm	SD	SC
1765cap	heat shock protein beta-1	small HSP	6.85	10.54	NS	SD	NS	NS	SD	SD	up	up
6736	kda small heat shock protein	small HSP	6.62	11.94	NS	NS	SC	SC	SC	SC	up	up
4361_2	heat shock protein 70	HSP70	6.41	11.97	NS	NS	NS	NS	SD	SD	up	up
4097_2	heat shock protein beta-1	small HSP	5.62	11.74	NS	NS	SC	SC	SC	SC	up	up
70407_3	heat shock protein	small HSP	5.34	11.83	NS	NS	NS	SC	SD	SD	up	up
2724	heat shock 70 kda protein cognate 4-like isoform 1	HSP70	4.97	12.74	NS	NS	SD	SD	SD	SD	up	up
7380	heat shock protein 40	HSP40	3.80	9.86	NS	NS	NS	NS	NS	NS	up	up
162_2	heat shock protein 90	HSP90	3.29	13.61	NS	NS	NS	NS	NS	NS	up	up
3121	heat shock protein 67b2	small HSP	2.57	8.38	NS	NS	NS	NS	NS	NS	up	NS
4411	lethal2essential for life	small HSP	2.49	7.77	NS	SD	NS	NS	SD	SD	up	NS
753cap	heat shock protein beta-1	small HSP	2.36	8.37	NS	NS	NS	NS	SC	SC	NS	NS
24253	dnaj homolog subfamily a member 1	HSP40	2.20	6.05	NS	NS	NS	NS	NS	NS	up	NS
69610_3	heat shock protein 70	HSP70	2.07	11.84	NS	NS	NS	SC	NS	NS	up	up
728_2	heat shock protein 4	HSP70	2.04	10.73	NS	SD	SD	SD	SD	SD	up	NS
1508	heat shock protein 70	HSP70	1.98	7.23	NS	NS	NS	SC	SC	SC	NS	NS
1748cap	heat shock protein 40	HSP40	1.80	10.14	NS	SD	NS	NS	SD	SD	up	NS
11608	dnaj homolog subfamily b member 9 isoform x2	HSP40	1.67	6.78	NS	NS	NS	NS	NS	NS	up	NS
7972	dnaj homolog subfamily b member 5 isoform x2	HSP40	1.61	7.18	NS	NS	NS	NS	NS	NS	up	NS
6659	chaperone protein dnaj 16	HSP40	0.78	7.42	NS	NS	NS	NS	SC	NS	up	NS

Table 2-8: The number of genes in each category of regulatory divergence under both treatments and both reciprocal hybrid crosses.

	Control		Heat stress	
	SCfxSDm	SDfxSCm	SCfxSDm	SDfxSCm
Conserved	8084	8048	6977	6866
Compensatory	65	93	157	180
cis only	222	229	417	427
trans only	24	23	58	65
cis + *trans*	8	10	34	30
cis x *trans*	3	3	18	22
ambiguous	1273	1273	2018	2089

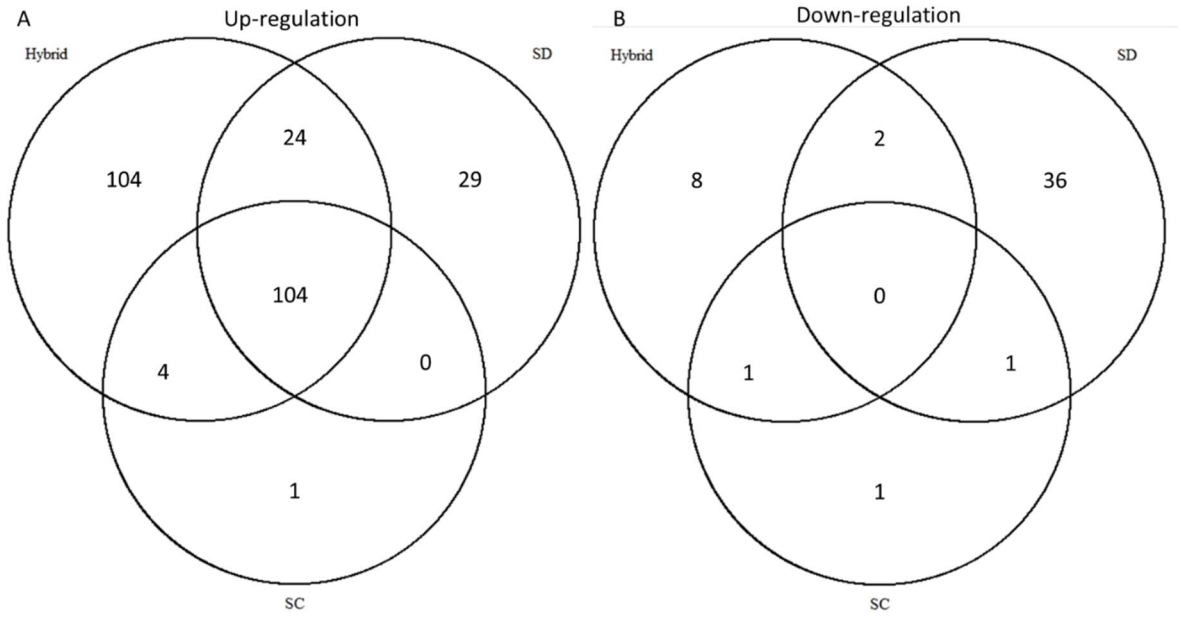


Figure 2-1: Venn diagrams showing the number of significantly differential expressed genes both up (A) and down-regulation (B) in response to heat stress treatments in F1 hybrids and parental populations.

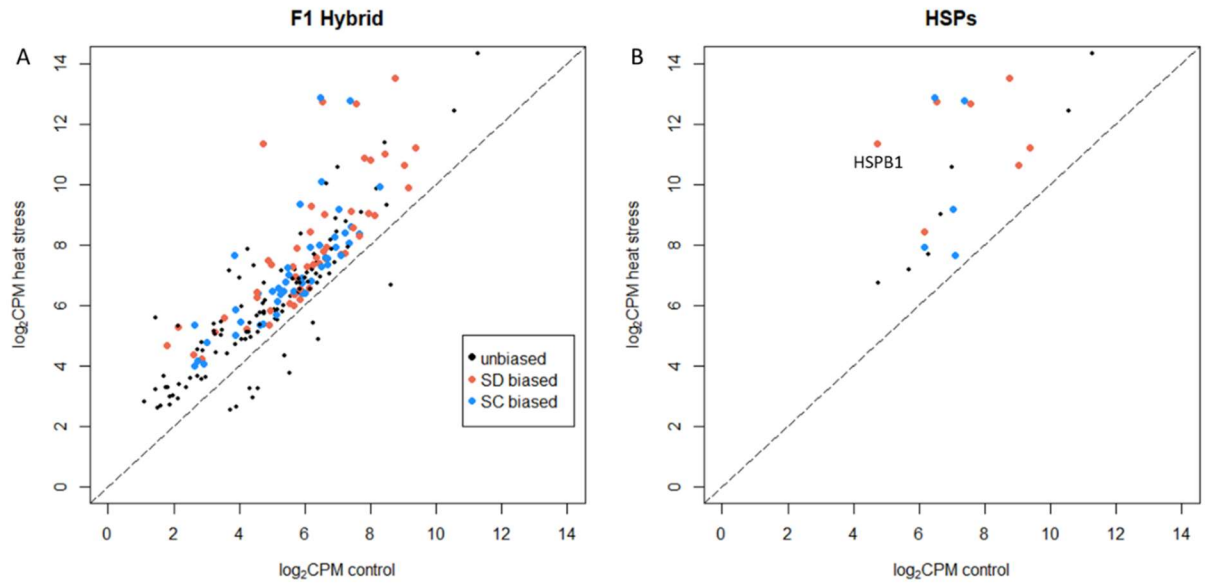


Figure 2-2: Significant differential expressed genes in F1 hybrids (A). Dots above the dash line represent significantly up-regulated genes in response to heat stress and dots below the dash line represent down-regulated genes. Colors designate direction of significant allele specific expression (Blue = SC and Coral = SD). Plot B only shows significantly differential expressed HSP genes in F1 hybrids.

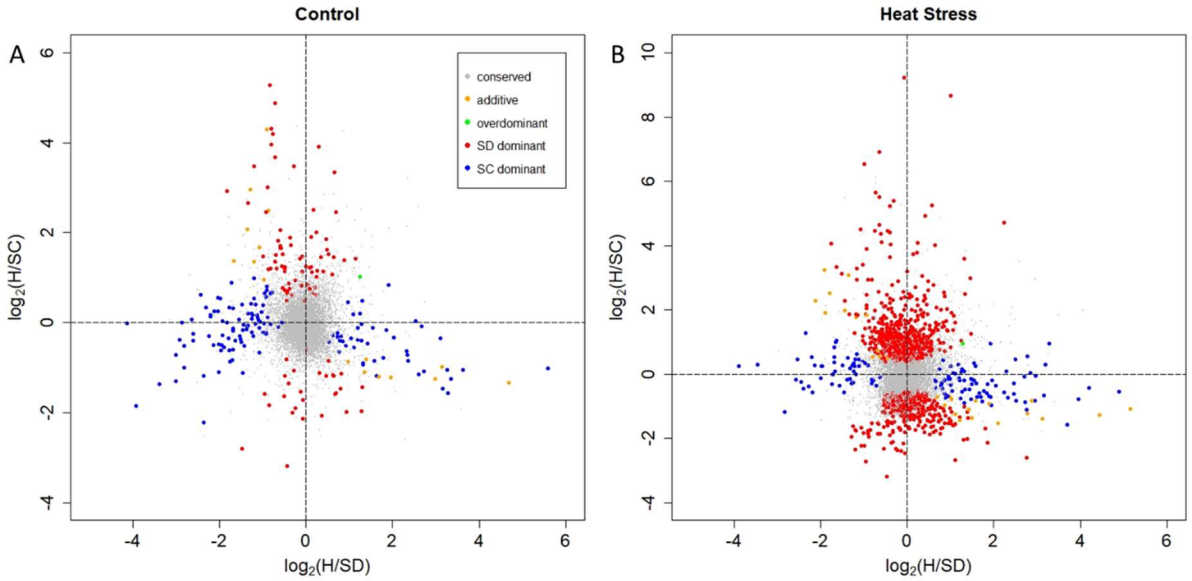


Figure 2-3: Scatter plots showing relative expression between F1 hybrid and each parent and inheritances of gene expression phenotypes for control (A) and heat stress (B) treatments. Each dot represents a single gene and each color represents a mode of inheritance.

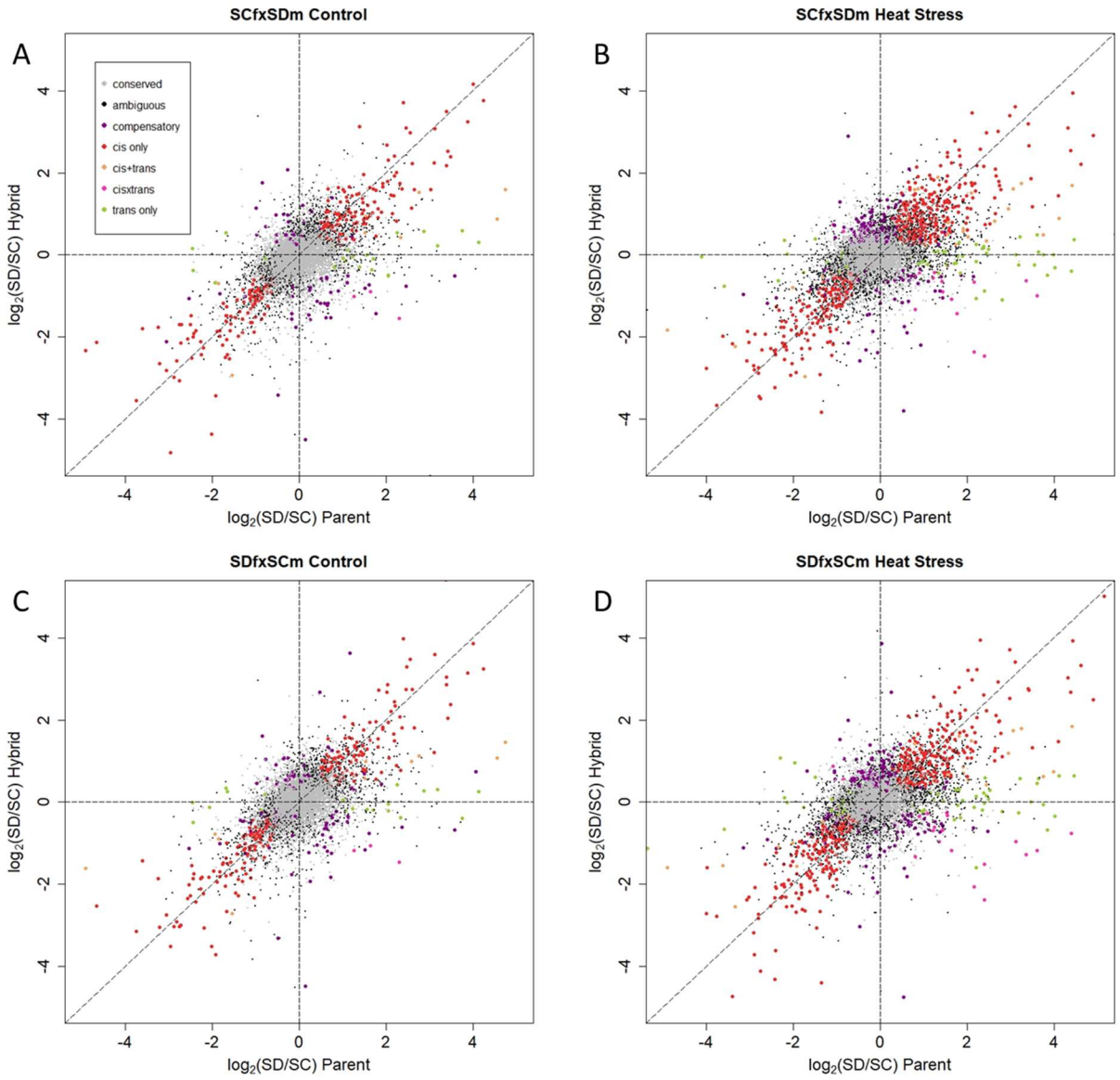


Figure 2-4: Scatter plots showing relative allele specific expression in F1 hybrids and relative total expression in parents. Each plot represents a specific reciprocal hybrid cross in a treatment. The colors represent different combinations of *cis* and *trans* regulatory divergence underlying gene expression patterns.

References

Andersson T, Forlin L. 1992. Regulation of cytochrome-P450 enzyme-system in fish. Aquatic Toxicology 24:1-19.

Anzenbacher P, Anzenbacherova E. 2001. Cytochromes P450 and metabolism of xenobiotics. Cellular and Molecular Life Sciences 58:737-747.

Barreto FS, Pereira RJ, Burton RS. 2015. Hybrid Dysfunction and Physiological Compensation in Gene Expression. Molecular Biology and Evolution 32:613-622.

Barreto FS, Schoville SD, Burton RS. 2015. Reverse genetics in the tide pool: knock-down of target gene expression via RNA interference in the copepod *Tigriopus californicus*. Molecular Ecology Resources 15:868-879.

Barshis DJ, Ladner JT, Oliver TA, Seneca FO, Traylor-Knowles N, Palumbi SR. 2013. Genomic basis for coral resilience to climate change. Proceedings of the National Academy of Sciences of the United States of America 110:1387-1392.

Brunquell J, Morris S, Lu Y, Cheng F, Westerheide SD. 2016. The genome-wide role of HSF-1 in the regulation of gene expression in Caenorhabditis elegans. Bmc Genomics 17.

Chen B, Feder ME, Kang L. 2018. Evolution of heat-shock protein expression underlying adaptive responses to environmental stress. Molecular Ecology 27:3040-3054.

Chen J, Nolte V, Schlotterer C. 2015. Temperature Stress Mediates Decanalization and Dominance of Gene Expression in Drosophila melanogaster. Plos Genetics 11.

Ciechanover A. 1998. The ubiquitin-proteasome pathway: on protein death and cell life. Embo Journal 17:7151-7160.

Coolon JD, Wittkopp PJ. 2013. *cis*- and *trans*-Regulation in *Drosophila*Interspecific Hybrids. In: Chen ZJ, Birchler JA, editors. Polyploid and Hybrid Genomics: Wiley-Blackwell. p. 37-57.

Cowles CR, Hirschhorn JN, Altshuler D, Lander ES. 2002. Detection of regulatory variation in mouse genes. Nature Genetics 32:432-437.

Feder ME, Hofmann GE. 1999. Heat-shock proteins, molecular chaperones, and the stress response: Evolutionary and ecological physiology. Annual Review of Physiology 61:243-282.

Gidalevitz T, Prahlad V, Morimoto RI. 2011. The Stress of Protein Misfolding: From Single Cells to Multicellular Organisms. Cold Spring Harbor Perspectives in Biology 3.

Gleason LU, Burton RS. 2015. RNA-seq reveals regional differences in transcriptome response to heat stress in the marine snail *Chlorostoma funebralis*. Molecular Ecology 24:610-627.

Gould BA, Chen Y, Lowry DB. 2018. Gene Regulatory Divergence Between Locally Adapted Ecotypes in Their Native Habitats. Molecular ecology.

Hochachka PW, Somero GN. 2002. Biochemical adaptation: mechanism and process in physiological evolution. Biochemical adaptation: mechanism and process in physiological evolution.:i-xi, 1-466.

Hoekstra LA, Montooth KL. 2013. Inducing extra copies of the Hsp70 gene in Drosophila melanogaster increases energetic demand. Bmc Evolutionary Biology 13.

Kelly MW, Sanford E, Grosberg RK. 2012. Limited potential for adaptation to climate change in a broadly distributed marine crustacean. Proceedings of the Royal Society B-Biological Sciences 279:349-356.

Kingsolver JG. 2009. The Well-Temperatured Biologist. American Naturalist 174:755-768.

Krebs RA, Feder ME. 1997. Deleterious consequences of Hsp7O overexpression in Drosophila melanogaster larvae. Cell Stress & Chaperones 2:60-71.

Krebs RA, Feder ME. 1998. Hsp70 and larval thermotolerance in Drosophila melanogaster: how much is enough and when is more too much? Journal of Insect Physiology 44:1091-1101.

Labbadia J, Morimoto RI. 2015. The Biology of Proteostasis in Aging and Disease. Annual Review of Biochemistry, Vol 84 84:435-464.

Lang RP, Bayne CJ, Camara MD, Cunningham C, Jenny MJ, Langdon CJ. 2009. Transcriptome Profiling of Selectively Bred Pacific Oyster Crassostrea gigas Families that Differ in Tolerance of Heat Shock. Marine Biotechnology 11:650-668.

Lemos B, Araripe LO, Fontanillas P, Hartl DL. 2008. Dominance and the evolutionary accumulation of cis- and trans-effects on gene expression. Proceedings of the National Academy of Sciences of the United States of America 105:14471-14476.

Leong W, Sun PY, Edmands S. 2018. Latitudinal Clines in Temperature and Salinity Tolerance in Tidepool Copepods. Journal of Heredity 109:71-77.

Lerman DN, Feder ME. 2005. Naturally occurring transposable elements disrupt hsp70 promoter function in *Drosophila melanogaster*. Molecular Biology and Evolution 22:776-783.

Lerman DN, Michalak P, Helin AB, Bettencourt BR, Feder ME. 2003. Modification of heat-shock gene expression in *Drosophila melanogaster* populations via transposable elements. Molecular Biology and Evolution 20:135-144.

Lindquist S. 1986. The Heat-Shock Response. Annual Review of Biochemistry 55:1151-1191.

Lockwood BL, Sanders JG, Somero GN. 2010. Transcriptomic responses to heat stress in invasive and native blue mussels (genus Mytilus): molecular correlates of invasive success. Journal of Experimental Biology 213:3548-3558.

Lovell JT, Schwartz S, Lowry DB, Shakirov EV, Bonnette JE, Weng XY, Wang M, Johnson J, Sreedasyam A, Plott C, et al. 2016. Drought responsive gene expression regulatory divergence between upland and lowland ecotypes of a perennial C-4 grass. Genome Research 26:510-518.

Lund SP, Nettleton D, McCarthy DJ, Smyth GK. 2012. Detecting Differential Expression in RNA-sequence Data Using Quasi-likelihood with Shrunken Dispersion Estimates. Statistical Applications in Genetics and Molecular Biology 11.

Mack KL, Campbell P, Nachman MW. 2016. Gene regulation and speciation in house mice. Genome Research 26:451-461.

Madeira D, Mendonca V, Dias M, Roma J, Costa PM, Larguinho M, Vinagre C, Diniz MS. 2015. Physiological, cellular and biochemical thermal stress response of intertidal shrimps with different vertical distributions: Palaemon elegans and Palaemon serratus. Comparative Biochemistry and Physiology a-Molecular & Integrative Physiology 183:107-115.

McManus CJ, Coolon JD, Duff MO, Eipper-Mains J, Graveley BR, Wittkopp PJ. 2010. Regulatory divergence in *Drosophila* revealed by mRNA-seq. Genome Research 20:816-825.

Naranjo S, Smith JD, Artieri CG, Zhang M, Zhou YQ, Palmer ME, Fraser HB. 2015. Dissecting the Genetic Basis of a Complex cis-Regulatory Adaptation. Plos Genetics 11.

Negri A, Oliveri C, Sforzini S, Mignione F, Viarengo A, Banni M. 2013. Transcriptional Response of the Mussel Mytilus galloprovincialis (Lam.) following Exposure to Heat Stress and Copper. Plos One 8.

Parsell DA, Lindquist S. 1993. The function of Heat-Shock Proteins in stress tolerance - degredation and reactivation of damaged proteins. Annual Review of Genetics 27:437-496.

Peterson DL, Kubow KB, Connolly MJ, Kaplan LR, Wetkowski MM, Leong W, Phillips BC, Edmands S. 2013. Reproductive and phylogenetic divergence of tidepool copepod populations across a narrow geographical boundary in Baja California. Journal of Biogeography 40:1664-1675.

Pirkkala L, Alastalo TP, Zuo XX, Benjamin IJ, Sistonen L. 2000. Disruption of heat shock factor 1 reveals an essential role in the ubiquitin proteolytic pathway. Molecular and Cellular Biology 20:2670-2675.

Pirkkala L, Nykanen P, Sistonen L. 2001. Roles of the heat shock transcription factors in regulation of the heat shock response and beyond. Faseb Journal 15:1118-1131.

Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26:139-140.

Rosic NN, Pernice M, Dunn S, Dove S, Hoegh-Guldberg O. 2010. Differential Regulation by Heat Stress of Novel Cytochrome P450 Genes from the Dinoflagellate Symbionts of Reef-Building Corals. Applied and Environmental Microbiology 76:2823-2829.

Schaefke B, Emerson JJ, Wang TY, Lu MYJ, Hsieh LC, Li WH. 2013. Inheritance of Gene Expression Level and Selective Constraints on Trans- and Cis-Regulatory Changes in Yeast. Molecular Biology and Evolution 30:2121-2133.

Schoville SD, Barreto FS, Moy GW, Wolff A, Burton RS. 2012. Investigating the molecular basis of local adaptation to thermal stress: population differences in gene expression across the transcriptome of the copepod *Tigriopus californicus*. BMC Evolutionary Biology 12.

Snyder MJ. 2000. Cytochrome P450 enzymes in aquatic invertebrates: recent advances and future directions. Aquatic Toxicology 48:529-547.

Somero GN. 1995. Proteins and temperature. Annual Review of Physiology 57:43-68.

Tangwancharoen S, Burton RS. 2014. Early life stages are not always the most sensitive: heat stress responses in the copepod Tigriopus californicus. Marine Ecology Progress Series 517:75-83.

Tangwancharoen S, Moy GW, Burton RS. 2018. Multiple modes of adaptation: regulatory and structural evolution in a small heat shock protein gene. Molecular biology and evolution 35:2110-2119.

Tautz D. 2000. Evolution of transcriptional regulation. Current Opinion in Genetics & Development 10:575-579.

Teranishi KS, Stillman JH. 2007. A cDNA microarray analysis of the response to heat stress in hepatopancreas tissue of the porcelain crab Petrolisthes cinctipes. Comparative Biochemistry and Physiology D-Genomics & Proteomics 2:53-62.

Tomanek L, Somero GN. 1999. Evolutionary and acclimation-induced variation in the heat-shock responses of congeneric marine snails (genus *Tegula*) from different thermal habitats: Implications for limits of thermotolerance and biogeography. Journal of Experimental Biology 202:2925-2936.

Vergara-Amado J, Silva AX, Manzi C, Nespolo RF, Cardenas L. 2017. Differential expression of stress candidate genes for thermal tolerance in the sea urchin Loxechinus albus. Journal of Thermal Biology 68:104-109.

Wang W, Hui JHL, Williams GA, Cartwright SR, Tsang LM, Chu KH. 2016. Comparative transcriptomics across populations offers new insights into the evolution of thermal resistance in marine snails. Marine Biology 163.

Willett CS. 2010. Potential fitness trade-offs for thermal tolerance in the intertidal copepod *Tigriopus californicus*. Evolution 64:2521-2534.

Willett CS, Ladner JT. 2009. Investigations of fine-scale phylogeography in Tigriopus californicus reveal historical patterns of population divergence. Bmc Evolutionary Biology 9.

Wittkopp PJ. 2010. Variable Transcription Factor Binding: A Mechanism of Evolutionary Change. Plos Biology 8.

Wittkopp PJ, Kalay G. 2012. Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. Nature Reviews Genetics 13:59-69.

Wray GA. 2007. The evolutionary significance of *cis*-regulatory mutations. Nature Reviews Genetics 8:206-216.

Wray GA, Hahn MW, Abouheif E, Balhoff JP, Pizer M, Rockman MV, Romano LA. 2003. The evolution of transcriptional regulation in eukaryotes. Molecular Biology and Evolution 20:1377-1419.

Wu C. 1995. Heat shock transcription factors: Structure and regulation. Annual Review of Cell and Developmental Biology 11:441-469.

Chapter 3

Multiple modes of adaptation*:* regulatory and structural evolution in a small heat shock protein gene

Sumaetee Tangwancharoen, Gary W Moy, Ronald S Burton

Abstract

Thermal tolerance is a key determinant of species distribution. Despite much study, the genetic basis of adaptive evolution of thermal tolerance, including the relative contributions of transcriptional regulation versus protein evolution, remains unclear. Populations of the intertidal copepod *Tigriopus californicus* are adapted to local thermal regimes across their broad geographic range. Upon thermal stress, adults from a heat tolerant southern population (San Diego) upregulate several heat shock proteins (HSPs) to higher levels than those from a less tolerant northern population (Santa Cruz). Suppression of a specific HSP, HSPB1, significantly reduces *T. californicus* survival following acute heat stress. Sequencing of HSPB1 revealed population specific nucleotide substitutions in both promoter and coding regions of the gene. HSPB1 promoters from heat tolerant populations contain two canonical heat shock elements (HSEs), the binding sites for heat shock transcription factor (HSF), while less tolerant populations have mutations in these conserved motifs. Allele specific expression of HSPB1 in F1 hybrids between tolerant and less tolerant populations showed significantly biased expression favoring alleles from tolerant populations and supporting the adaptive divergence in these *cis*-regulatory variants. The functional impact of population-specific non-synonymous substitutions in HSPB1 coding sequences was tested by assessing the thermal stabilization properties of SD versus SC HSPB1 protein variants. Recombinant HSPB1 from the southern SD population showed greater capacity for protecting protein structure under elevated temperature. Our results indicate that both regulatory and protein coding sequence evolution within a single gene appear to contribute to thermal tolerance phenotypes and local adaptation among conspecific populations.

Introduction

Populations of species that range across heterogeneous habitats frequently show evolutionary adaptation to their local environments. Adaptive phenotypes can stem from genetic variation in coding sequences, gene regulatory sequences or both. The relative contributions to adaptation from structural and gene regulatory variation are the subject of considerable debate (Carroll 2005; Hoekstra and Coyne 2007; Wray 2007). Clear evidence exists for both structural (Place and Powers 1979; Alahiotis 1982; Perutz 1983; Wirgin, et al. 2011) and regulatory (Schulte, et al. 2000; Juneja, et al. 2016) variations that lead to adaptive physiological traits that correspond to organisms' local habitats.

In organisms ranging from bacteria to vertebrates, thermal stress induces the expression of Heat Shock Proteins (HSPs) that help mitigate cellular damage from misfolded proteins (Lindquist 1986; Parsell and Lindquist 1993; Feder and Hofmann 1999). Heat shock response is among the best-established models for studying gene regulatory mechanisms {e.g., *Drosophila* HSP70 gene expression (Perisic, et al. 1989; Fernandes, et al. 1995)}. Differential expression of HSPs has been linked to differences in thermal tolerance within (Schoville, et al. 2012; Gleason and Burton 2015) and among (Tomanek and Somero 1999) species and available evidence suggests that the evolution of thermal tolerance may be at least partially driven by changes in HSP gene expression (Sørensen , et al. 2003). The eukaryotic HSP gene regulatory system is well-studied. Heat shock response is largely transcriptionally regulated by heat shock transcription factor (HSF) (Wu 1995). HSF is known to bind to a specific DNA sequence motif called the heat shock element (HSE) (Amin, et al. 1988) upon thermal stress and mediate transcriptional response of HSPs (Pelham 1982). An HSE unit consists of three inverted tandem repeats of a 5 base pair motif with 3 conserved base pairs in the middle as 'nGAAn'. The 15 bp

units with all consensus sequence among eukaryotic lineages is called the 'perfect' or 'canonical' HSE. Each 5bp motif binds a subunit of HSF which is trimeric when active (Fernandes, et al. 1994). Though the mechanism of HSF-HSP transcription regulation has been well characterized, to date few examples (Lerman and Feder 2001; Lerman, et al. 2003) exist that demonstrate a causal connection between point mutations in heat shock gene regulatory sequences and adaptation to different temperature regimes across a species range.

Although intraspecific HSP regulatory variation appears to contribute to the evolution of thermal tolerance (Sørensen , et al. 2001; Fangue, et al. 2006), additional modes of adaptation, such as structural variation in the HSPs themselves, have not been widely considered. There is some evidence for correlations between small HSP genotypes and thermal environment (Frydenberg, et al. 2003; Healy, et al. 2010; Graham, et al. 2012). Although these correlations suggest a contribution of HSP structural variation to the evolution of thermal tolerance, meaningful comparison of the functional difference of HSP alleles, especially for the small HSP family, is lacking.

The copepod *Tigriopus californicus* inhabits high intertidal pools along the west coast of North America, spanning a broad latitudinal gradient from Alaska, USA (Dethier 1980) to Baja California, Mexico (Ganz and Burton 1995). Previous studies have shown that populations along the coast exhibit different tolerances to acute thermal stress, with southern populations being significantly more tolerant of high temperatures than northern populations (Willett 2010; Kelly, et al. 2013). Populations differ in HSP gene expression following exposure to heat stress; specifically, a heat tolerant San Diego (SD) population (32^{o}45 N 117^{o}15W) showed a greater degree of upregulation of HSPs than a less tolerant Santa Cruz (SC) population (36^{o}57 N 122^{o}03W) (Schoville, et al. 2012). Among upregulated genes following 1-hour acute heat stress

in *T. californicus*, HSPB1 (Accession number: JW506233) showed > 100X increase in transcript abundance in the SD population while only 5X upregulation was observed in the SC population. In addition to differential expression, there is also significant non-synonymous variation in the protein coding region of HSPB1 gene, making this an attractive system to evaluate the functional consequences of both structural and regulatory variation in adaptive phenotypes.

One difficulty encountered in assessing the role of single HSP genes in adaptation is that in many organisms (including *T. californicus*), there are multiple families of HSP genes and even within families, each HSP gene can have multiple copies suggesting functional redundancy (de Jong, et al. 1998). However, a few studies had shown that small HSPS can have an essential and non-redundant functional role. A study in *Drosophila* showed that different levels of one small HSP transcripts can lead to differences in thermal tolerance in *Drosophila* larvae (Lockwood, et al. 2017). HSPB1 knocked out mice did not show compensation by other HSPs both at mRNA and protein levels and HSPB1 knocked out cell-line showed less viability after heat treatment (Huang, et al. 2007). Similarly, in *T. californicus*, when RNAi was used to knock down HSPB1 expression, mortality dramatically increased after heat stress (average 5-day survivorship following stress was reduced by approximately 80%), indicating there was no direct back-up capacity for its function (Barreto, Schoville, et al. 2015). These results suggest a critical role for HSPB1 and its transcriptional regulation in survivorship following heat stress exposure and possibly in driving local adaptation among *T. californicus* populations.

The current study examines the potential role of DNA sequence variation in both the proximal promoter and protein coding regions of the small HSP gene, HSPB1, in generating population differences in thermal tolerance in *T. californicus*. We hypothesize that the observed divergence in HSPB1 promoter sequences account for differences in transcript abundances

across populations while coding sequence variation results in allelic differences in the thermal protectant properties of HSPB1; combined, the experimental results present a compelling case for the roles of both regulatory and structural gene evolution in molding adaptation to local thermal regimes across a species range.

Results and discussion

HSPB1 promotor sequencing

Given the evidence suggesting HSF regulation of HSP genes from other model systems and taking advantage of the existing draft genome sequence for the SD *T. californicus* population (see https://i5k.nal.usda.gov/Tigriopus_californicus), we examined the 5' flanking region of HSPB1 for population differences in potential gene regulatory sequences. The gene does not have introns; therefore, only flanking promoter sequences were searched for HSEs using JASPAR (Mathelier, et al. 2014). From the SD sequence, we identified 2 HSEs within the promoter region 584bp upstream of the transcription start site of HSPB1. The upstream sequence beyond this region is a repeat sequence. Both of these HSEs have the three inverted repeats of the 'canonical HSE'.

We then sequenced the proximal promoter region of the HSPB1 gene from a set of 11 *T. californicus* populations from distinct geographic regions spanning from Baja California to Vancouver Island, Canada (Figure 3-1A). By sequencing multiple individuals from a population, we identified two geographic regions with different promoter genotypes. Southern populations from Southern California and Baja California all have two 'canonical HSEs' while northern populations have nucleotide substitutions within the conserved regions of both HSEs (Figure 3-

1B, full promoter alignment Figure A3-1). The AB population from Los Angeles County, California, appears to be a transition between the two regions, with one intact 'canonical HSE' and one HSE with polymorphic site in the conserved GAA motif.

Though there are additional sequence polymorphisms in the HSPB1 promoter among populations, research in other systems suggests that the observed variation in the HSEs alone could result in different gene expression phenotypes, with promoter strength declining when nucleotide substitutions result in deviations from the canonical HSEs with conserved GAA sequences (Fernandes, et al. 1994, 1995; Dierick, et al. 2007). Unlike *Drosophila* (Lerman, et al. 2003), we found no evidence of transposon insertion in the promoter region between the transcription start sites and the HSEs in the populations sequenced.

RNA sequencing and allele specific expression

In order to determine if the observed mutations in the HSEs within the HSPB1 promoter actually result in differential transcription, we tested for differences in allele specific expression (ASE) in F1 hybrids between the SD and the SC populations. In hybrids, the two HSPB1 alleles are present in the same cellular environment including all *trans*-acting factors, so differences in allelic expression are isolated to the effects of different *cis*-regulatory elements (Tirosh, et al. 2009; McManus, et al. 2010). Given its canonical HSEs, we hypothesized that expression of the SD HSPB1 allele would be favored in SD/SC F1 hybrids.

Using RNA-seq analysis, we found only low levels of HSPB1 transcripts in both parental and F1 hybrid animals under control temperature (20^{o}C) conditions. In agreement with previous findings in parental populations SD and SC (Schoville, et al. 2012), HSPB1 was strongly upregulated after heat stress (Table 3-1). HSPB1 expression was significantly biased in hybrids

under heat stress conditions, with the SD allele elevated in both biological replicates in both reciprocal crosses (i.e., all four tests). Notably, this strong expression bias was only observed under the heat stress conditions when the HSF/HSE mediated upregulation of expression is expected to occur (Table 3-2). The strong bias in allelic expression in F1 hybrids suggests that the causal mutation is in the *cis*-regulatory elements of the HSF gene regulatory network, most likely the substitutions in the conserved motif of HSEs in the promoter region. The hybrids also showed slightly biased HSPB1 expression (toward the SD allele) under control condition, which may be due to either a low level of mapping bias between two alleles (Table A3-2) due to higher polymorphism in SC population (Pereira, et al. 2016), or it may reflect low levels of HSF/HSE mediated expression favoring the SD allele under the control conditions. However, any bias due to variation in mapping efficiency is relatively minor (>90% of hybrid reads under heat stress treatment mapped correctly) and could not account for the large expression bias between the alleles observed under heat stress condition.

Allele specific expression in additional interpopulation crosses

To further confirm the functional consequences of the substitutions in the HSEs, we examined levels of ASE in HSPB1 in hybrids between different population pairs, including pairs with the same HSE structure: BR and SD each have two canonical HSEs while PES and SC have substitutions in the conserved motif of both HSEs. We predicted that only the F1 hybrid between populations with different promoter structure (eg. BRxPES and SDxSC) would show ASE in HSPB1 after heat stress treatment; F1 hybrids between populations with similar promoters would not show allelic imbalance. To identify ASE, fragments of HSPB1 coding sequence were PCR amplified from both genomic DNA and cDNA followed by restriction digests to discriminate between the alleles. Following agarose gel electrophoresis, ASE was evaluated by comparing

intensity of allele-specific fragments between F1 genomic DNA and cDNA templates (Figure 3-2A-D, Table 3-3). Band intensity of digested amplicons from genomic DNA template from F1 hybrids should reflect equal proportions of template expected in HSPB1 heterozygotes; thus, comparing cDNA template band intensity to genomic DNA template reveals any biased expression. To validate this approach, SDxSC genomic DNA and mRNA from the RNAseq experiment were used also analyzed by restriction digest.

No evidence for ASE was observed in F1 hybrids between SDxBR and SCxPES, confirming that no ASE occurs in hybrids between populations with the same structure of HSEs in the promoter of HSPB1 (for both heat tolerant and heat sensitive population pairs). In contrast, BRxPES showed significant ASE; under heat stress treatment, there was substantial bias toward the BR allele (containing two canonical HSEs in the promoter region) over the PES allele similar to SDxSC (Figure 3-2A). The average band intensity of BR allele in F1 cDNA was approximately three times higher than the intensity of the genomic DNA template band (Figure 3-2E). ASE results from BRxPES hybrids further reinforces the SDxSC RNA-seq evidence for functional divergence in *cis*-regulatory elements, with enhanced expression of HSPB1 alleles from more heat tolerant populations.

Taken together, our data suggest that divergence in *cis*-elements of HSF-HSP gene regulatory network may underlie differential HSPB1 gene expression and ultimately contribute to differences in thermal tolerance among *T. californicus* populations. Biased expression of HSPB1 alleles in F1 hybrids suggests that observed single nucleotide polymorphisms (SNPs) in the *cis*-regulatory HSEs have significant functional consequences on regulation of gene expression. Furthermore, expression bias toward southern alleles (SD allele in F1 SDxSC hybrids and BR allele in BRxPES hybrids) suggests that alleles with canonical HSEs act as

stronger promoters than northern alleles that have substitutions in the conserved HSE motifs. The geographic pattern of stronger promoters for HSPB1 in southern *T. californicus* populations suggests that natural selection is favoring either regulatory variation that enhances the heat shock response in the warmer portion of the species range or animals that do not express HSP excessively in cooler climate. However, we cannot conclude which genotype is the ancestral trait with our current knowledge in phylogeographic history of the species. Edmands (2001) found reduced population differentiation among populations north of San Francisco Bay which might be partly attributed to recent post-glacial recolonization. This could explain HSPB1 promoter similarity between populations from the northern range (PAC and FHL) and Central California populations (SC, PES and PL) which covers more distance than the distance between SD and SC populations (Figure 3-1). However, the phylogeographic relationship among populations between Central California (including SC) and Southern California (including SD) remains unresolved (Edmands 2001; Peterson, et al. 2013).

Structural variation

Although our results strongly implicate adaptive variation in *cis*-regulation, amino acid sequence divergence in both HSF (*trans*-regulatory element) and HSPB1 itself may also contribute to differential thermal tolerance and local adaptation across populations. From existing transcriptome data (Schoville, et al. 2012), we identified a single HSF gene in *T. californicus* with 529 amino acids. HSF comparison between SD and SC revealed 10 amino acid substitutions between populations (Figure A3-3). Three additional populations' HSF sequences were obtained from unpublished RNA-seq data including BR, AB, and PES. The functional significance (if any) of these amino acid substitutions in HSF is unknown; however dN/dS analyses (estimated using PAML 4.7) (Yang 2007) found no significant evidence of positive

selection at HSF (ω, dN, dS: 0.1251 , 0.0066, 0.0525) (Table A3-3A). Any potential contribution of population differences in HSF *trans*-acting regulatory elements was factored out of the ASE studies by the experimental design.

In addition to HSF amino acid sequence divergence, we examined structural variation of HSPB1 across populations of *T. californicus*. Small HSPs including HSPB1 are characterized by an α-crystallin domain towards the C-terminal end of the protein (de Jong, et al. 1998). There are two α-crystallin domains (pfam00011) in *T. californicus* HSPB1. There is substantial structural variation of HSPB1 between SD and SC populations: 17 amino acid substitutions and one indel occur within the 277 amino acids (Figure A3-2). The amino acid substitutions between SD and SC populations were found throughout the gene including inside the α-crystallin domains suggesting the potential for functional difference between two HSPB1 alleles. We aligned HSPB1 coding sequences of five populations and found a relatively elevated dN/dS (ω = 0.440, dN=0.0393, dS=0.0886) (Table A3-3B) compared to transcriptome-wide mean (ω = 0.120 between SD and SC populations)(Barreto, et al. 2011). We further identified polymorphisms (> 1% of mapped reads) within SD and SC populations from our transcriptomes for calculating Neutrality Index (NI) (McDonald and Kreitman 1991) and Direction of Selection (DoS) (Stoletzki and Eyre-Walker 2011) to look for signal of positive selection (NI < 1, and DoS > 0). Both indices (NI = 0.642 (Fisher exact test p = 0.47) and DoS = -0.04) do not suggest any significant evidence of positive or purifying selection in the HSPB1 gene between SD and SC populations.

However, given the high number of fixed amino acid substitutions between SD and SC HSPB1alleles, we used *in vitro* functional assays to directly test the potential adaptive significance of the extensive population differentiation in HSPB1 coding sequence. A previous

study found that thermal tolerance of *E. coli* was increased when expressing a truncated *Tigriopus japonicus* HSPB1 homolog with only one α-crystallin domain (Seo, et al. 2006). Based on this evidence and HSPB1's putative function as a member of the small heat shock protein family, we hypothesized that adaptive evolution would lead to enhanced thermal protectant properties in southern alleles where populations are more frequently exposed to high temperatures. Using an *in vitro* thermal protection assay (Gong, et al. 2009), we tested the function of SD and SC variants of HSPB1 proteins expressed in *E. coli*. Purified HSPB1 protein (Figure A3-3) from each population was found to reduce *in vitro* aggregation of a test protein (porcine citrate synthase, CS) held at a high temperature. Furthermore, the SD allele consistently out-performed the SC allele in reducing the measured protein aggregation observed in each temperature treatment and in all four HSPB1 concentrations tested (Figure 3-3); although each of these tests was not replicated, we view the consistent differences across all temperatures and all HSPB1 concentrations as appropriate validation of the functional differences between the SD and SC variants, especially because the relevant *in vivo* concentrations are unknown. We further analyzed the functional differences between SD and SC HSPB1 alleles using an enzyme activity assay (Hristozova, et al. 2016). Adding recombinant HSPB1 protected citrate synthase from enzymatic activity loss by high temperature, $F_{2,44} = 21.72688$, $p = <0.0001$ (Table 3-4). Again, the SD allele outperformed SC allele in retaining CS enzyme activity (Figure 3-4) with Tukey pairwise comparison test showed a significant difference between SD and SC allele ($p < 0.0001$) (Table 3-5). Results from both experiments confirmed our hypothesis that HSPB1 from more heat tolerant population has enhanced thermoprotectant properties.

Conclusions

In this study, we demonstrate that variation in HSPB1 expression and function among populations of *T. californicus* can, in part, be attributed to both *cis*-regulatory variation and coding sequence variation in the HSPB1 gene. ASE assays in F1 interpopulation hybrids confirmed the functional significance of SNPs in *cis*-regulatory elements between populations that differ in thermal tolerance phenotypes. Additionally, *in vitro* assays showed that HSPB1 produced by the heat tolerant SD was more potent at preventing protein aggregation and preserving enzyme activity at high temperature than HSPB1 from the less heat tolerant SC population. Both findings, in *cis*-regulatory sequences and protein function, are consistent with geographic differences in the thermal regimes experienced by different copepod populations. Together with previous studies verifying the key role of HSPB1 in thermal response in *T. californicus* (Schoville, et al. 2012; Barreto, Schoville, et al. 2015), the present study demonstrates that selection can act on both protein structure and regulation of expression within a single gene, and that each mode of selection may contribute to local adaptation among populations.

Materials and methods

Copepod culturing and hybridization

Copepods were collected from high intertidal rock pools along the Pacific coast of North America (Figure 3-1, Table A3-1). Stock populations were maintained in 400 ml beakers filled with 250 ml of filtered (0.45 μm) seawater under constant 20ºC and 12 hour light:dark photoperiod. Copepods were fed ground *Spirulina* wafer fish food *ad libitum*. *T. californicus* F1 hybrids were produced following Barreto et. al. (Barreto, Pereira, et al. 2015). For the RNA

sequencing experiment, each replicate of each reciprocal cross between SD and SC consisted of F1 hybrids that came from more than 30 successfully mated females.

Heat stress experiment and RNA sequencing

Both parental populations and reciprocal F1 hybrids between SD and SC were exposed to constant 20 ºC as a control treatment; experimental heat stress presumed to activate HSF involved increasing temperatures by 5 ºC per hour up to 35 ºC. After 60 minutes at 35 ºC, animals were sacrificed and RNA was immediately extracted with Tri Reagent (Sigma) following the manufacturer's protocol. Each sample contained 50 adult copepods of equal sex ratio and each treatment had two biological replicates. cDNA libraries were prepared using oligo-dT priming and Illumina's Truseq standard mRNA protocol. Libraries were sequenced (100-bp single-read) on the Illumina 4000 platform. Reads were aligned to a transcriptome reference (Barreto, Pereira, et al. 2015) that included paired orthologs from both populations. Parameters for mapping included a cutoff at 0.8 length fraction and 0.98 sequence similarity. Full analysis of ASE across the transcriptome will be presented elsewhere; here we focused only on the expression of HSPB1 alleles. Only reads that uniquely mapped to one of the two alleles were counted. We performed binomial tests with 5% FDR to identify significantly biased HSPB1 allelic expression in hybrids.

HSPB1 promoter sequencing and HSE identification

To obtain genomic DNA for PCR amplification, individual copepods were put in 15 µl lysis buffer with Proteinase K (Willett and Burton 2001) then heated to 55 ºC for 90 minutes followed by 95 ºC for 15 minutes. For pooled extraction, 15-50 individuals from each population were used following DNeasy blood and tissue kit protocol (QIAGEN). We use the SD

population draft genome to design primers to amplify a 361-bp product including 2 HSEs in the promoter region (forward primer: 5'-ACTAGTTGTCCGATACACAAACAAACTAT-3', reverse primer: 5'-GAAACAAAAGAGCCATGGTTTA-3'). We sequenced the promoter region from at least 10 individuals from each population. Sequences were aligned using Sequencher and Geneious.

Restriction digest

In addition to the RNA sequencing experiment involving the SD x SC hybrids, three further crosses were made to assess the role of promoter sequence in ASE. BRxPES cross is an independent test with different SNPs in the HSEs corresponding to SDxSC cross in the RNA-seq experiment. SDxBR and SCxPES (two southern and two northern populations, respectively) are crosses between populations with similar SNPs in the HSEs. Hybrids were raised in the same conditions as described above. For SCxPES and BRxPES crosses, F1 hybrids from a single female were used as a biological replicate. For SDxSC cross, we used pooled F1 hybrids from multiple females from independent crosses as a biological replicate. For each replicate, genomic DNA was obtained from F1 male using the same methods described above. cDNA was made from RNA extracted from F1 animals subjected to the same heat stress treatment as in the RNA sequencing experiment. Both genomic DNA and cDNA were used as template for PCR amplification of HSPB1 coding sequences. We used HSPB1 coding sequences of the four populations (unpublished data) to design PCR primers and identify population specific restriction cut sites using Webcutter 2.0 (Maarek, et al. 1997). The primers and the restriction enzymes used in the experiment are listed in Table A3-4. Restriction digest reactions were performed on 200 ng of PCR products following the manufacturer protocols. Each digested sample was run on gel electrophoresis up to 3 times to estimate the variability of the band intensity. Gel images were

analyzed for band intensity using Image Lab software (Bio-Rad). We performed nested ANOVA using the package nlme (Pinheiro, et al.) in R 3.4.2 (R core team, 2017) on percent band intensity of the largest digested band between genomic DNA and cDNA templates. We used the band intensity of the sample with genomic DNA template as a baseline for unbiased expression. Significant difference in band intensity in samples with cDNA template indicates biased ASE.

Expression and purification of HSPB1

Full length SD and SC variants of HSPB1 were amplified from their respective cDNA with N-terminal 6xHis-tag, cloned into the pProEx Htb expression vector (Invitrogen) and transformed into *E. coli* BL21 (DE3) pLysE cells. Ten ml of an overnight culture in LB medium were added to 200ml and grown for 2 h at 37 ºC. Expression of the recombinant HSPB1 was induced by the addition of isopropyl-ß-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. After an additional 2.5 h incubation at 37 ºC, cells were harvested by centrifugation at 3,000g for 20 minutes at 4 ºC and then lysed in lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 20 mM imidazole, 8 M urea, pH 7.4) The cell lysate was centrifuged 15,000g for 15 minutes to pellet the insoluble material and to collect the supernatant fraction of the cell lysate. The supernatant was then loaded onto a His60 Ni Superflow Resin column (Clontech). The column was washed in wash buffer (50 mM NaH_2PO_4, 300 mM NaCl, 40 mM imidazole, 8 M urea, pH 7.4) and then the His-tagged recombinant HSPB1 was eluted in wash buffer containing 300 mM imidazole. We did not remove His-Tag from the recombinant protein. The eluted HSPB1 fractions were then dialyzed against phosphate buffered saline (PBS) and protein concentration determined using the BCA Protein Assay Kit (Pierce).

Citrate Synthase (CS) aggregation and activity assays

Thermal aggregation experiments were performed as described in (Gong, et al. 2009). For each test, 10 µg CS from porcine heart (C3260, Sigma, 9.4 µg/µl) was incubated with either SD or SC HSPB1 in 1 ml PBS at 45 ºC and aggregation monitored by measuring turbidity at 320 nm in a spectrophotometer. Thermal inactivation of CS activity was done as described in (Hristozova, et al. 2016). The reaction was performed with 0.5 µg/ml CS (0.329 units/µg), 0.45 mM Acetyl-coA, 0.5mM oxaloacetate, and 0.1 mM Ellman's reagent (DTNB) in PBS and followed for 3 min at room temperature. CS activity was fitted into a linear mixed model with time and allele fixed effects and replicates (SD = 5, SC = 4, and no HSP = 6 with 56 total observations) as a random effect using R package nlme (Pinheiro, et al.). We performed Tukey's pairwise comparison among two alleles and no HSP control using R package multcomp (Hothorn, et al. 2008).

Author contributions

S.T. and R.S.B. conceived and planned the project. S.T. conducted the RNA-seq experiment, HSPB1 promoter sequencing, restriction digest experiment, and all data analyses. G.W.M. conducted the cloning of HSPB1 and aggregation assay. S.T., G.W.M and R.S.B. wrote the manuscript.

Acknowledgment

This work was supported by US National Science Foundation grants (DEB1551466 and IOS1155030) to R.S.B. The authors thank Dr. Ricardo Pereira, Brian Hong, and Dr. Jon Shurin for providing some of the copepod populations. We also thank undergraduates Wenpei Li and Kiana Woodward for their help in this work. Dr. Tim Healy, Dr. Satomi Tsuboko-Ishii and Alice Harada provided helpful comments on the manuscript.

Chapter 3, in full, was published in Molecular Biology and Evolution: Tangwancharoen S, Moy GW, Burton RS (2018) Multiple modes of adaptation: regulatory and structural evolution in a small Heat Shock Protein gene. Molecular Biology and Evolution. 35: 2110-2119. The dissertation author was the primary investigator and author of this paper.

Table 3-1: Differential expression pattern of HSPB1 across genotypes. Significant values were obtained from Likelihood Ratio Test in edgeR package (Robinson, et al. 2010).

Crosses	Heat stress expression (average log_2 CPM±SE)	log_2 Fold Change (heat stress / control)	p value
SDfxSCm	11.05±0.83	6.81	4.70E-45
SCfxSDm	11.68±0.30	6.77	7.96E-37
SD	12.14±0.41	7.57	5.24E-65
SC	10.49±0.30	5.88	1.54E-40

Table 3-2: F1 allele specific expression of HSPB1(raw mapped reads to each parental reference) of two independent reciprocal crosses and significant values (5% FDR p values) from binomial tests of equal expression between the two alleles in F1 hybrids

Cross	Control			Heat Stress		
	SD counts	SC counts	5% FDR p value	SD counts	SC counts	5% FDR p value
SCfxSDm #1	25	13	0.28	16365	6252	1.87E-321
SCfxSDm #2	152	128	0.27	29701	12960	1.61E-321
SDfxSCm #1	119	60	1.37E-3	13452	4976	1.80E-321
SDfxSCm #2	141	123	0.43	23157	7906	1.31E-321

Table 3-3: Significant values from nested ANOVA model with template as a fixed effect and a random batch effect for the test for ASE by restriction digest. Family numbers represent the number of independent F1 hybrid crossed used in the experiment. Observations are the total number of digested bands measured for intensity (%) and compared for ASE analysis.

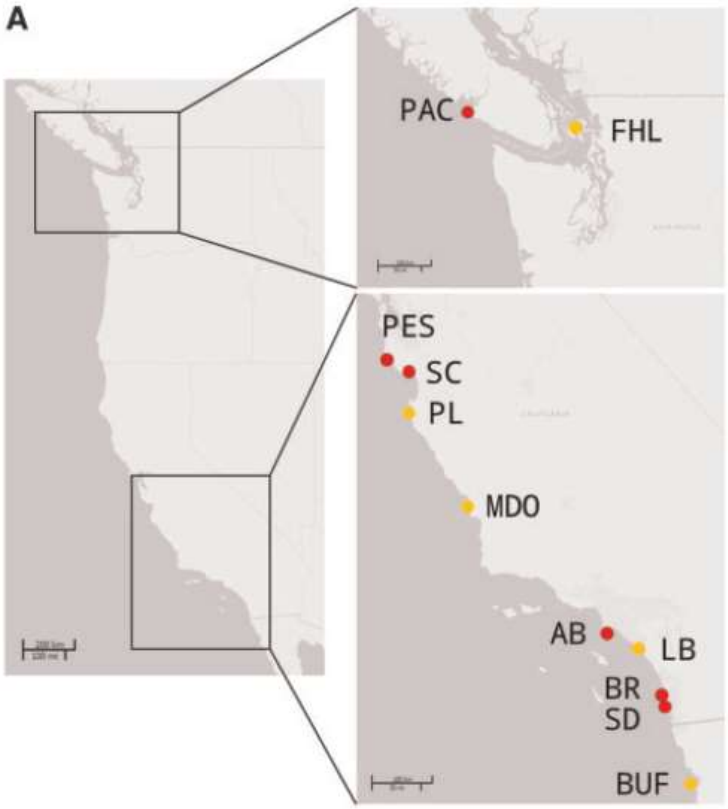
Cross	#family	#observations	F value	p value
SDxSC	4	6	78.96	<0.001
BRxPES	6	31	145.79364	<0.001
SDxBR	4	16	0.06362	0.8055
SCxPES	7	34	0.00703	0.9338

Table 3-4: Significant values from linear mixed model for the CS activity assay.

Source of variation	df	F value	p value
Intercept	1	33.21778	<0.0001
Time	1	243.71921	<0.0001
Allele	2	21.72688	<0.0001
Time x Allele	2	2.61886	0.0842

Table 3-5: Significant values from Tukey pairwise comparison among SD and SC allele treatments and no HSP control on the CS activity assay.

Hypothesis	p value
SD = no HSP	< 0.0001
SC = no HSP	0.775
SD = SC	< 0.0001



B

	-227	-50
PAC	CCAACGTTCTTGAAC	TTTCGAGAGATTTCA
FHL	CCAACGTTCTTGAAC	TTTCTAGACATTTCA
PES	CCAACGTTCTTGAAC	TTTCGAGASATTTCA
SC	CCAACGTTCTTGAAC	TTTCGAGASATTYCA
PL	CCAACGTTCTTGAAC	TTTCGAGACATTCCA
MDO	CCAACGTTCTTGAAC	TTTCGAGACATTTCA
AB	CGAACGTTCTTGAAC	ATTCGAGAAATTTYA
LB	CGAACGTTCTTGAAC	TTTCGAGAAATTTCA
BR	CGAACGTTCTTGAAC	TTTCGAGAAATTTCA
SD	CGAACGTTCTTGAAC	TTTCGAGAAATTTCA
BUF	CGAACGTTCTTGAAC	TTTCGAGAAATTTCA
	-*----------	*---*---*---**-

Figure 3-1: HSEs in the promoter region of HSPB1 gene across populations of *T. californicus*. A map showing populations of *T. californicus* along the Pacific coast of North America (A) with corresponding HSEs sequences in the HSPB1 promoter. Red dots indicate the populations that we sequenced individually. Orange dots indicates the populations that we performed pooled DNA extraction. (B). The numbers mark the positions upstream from transcription start site for the SD population. Full promotor alignment is shown in the supplementary figure S1. Red characters indicate the consensus GAA motif of HSEs. Blue characters indicate SNPs or polymorphism in the conserved part of HSEs that deviate from the consensus sequence. Green characters indicate SNPs outside of the conserved motif.

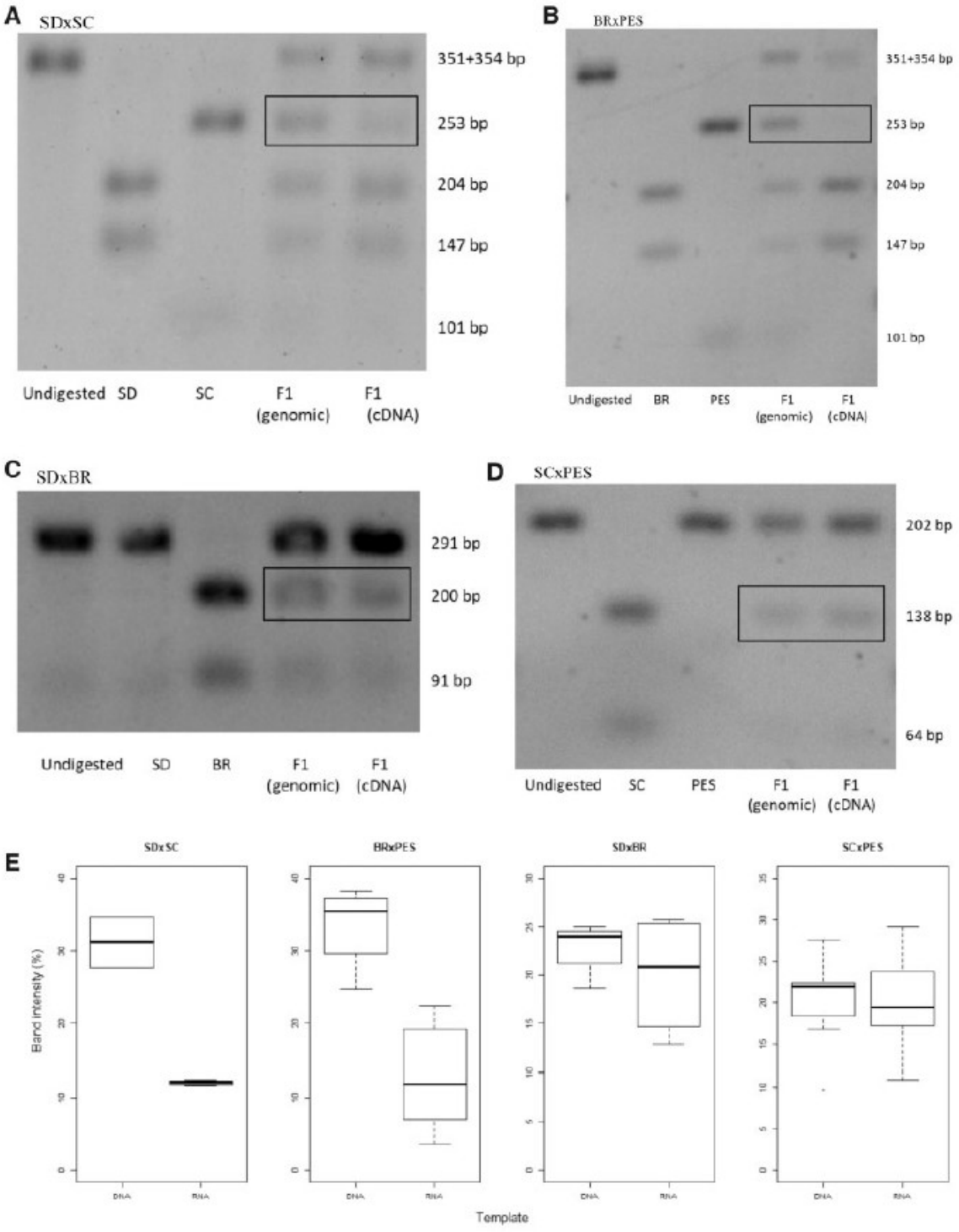


Figure 3-2: HSPB1 Allele Specific Expression in restriction digest experiment. Pictures form gel-electrophoresis showing restriction cut bands in each F1 hybrid and parent pairs (A-D). The boxes encircle the largest restriction cut bands in both genomic DNA and cDNA template used for evaluating ASE. Box plots showing percent band intensity between cut bands (encircled bands in A-D) between genomic DNA and cDNA templates (E).

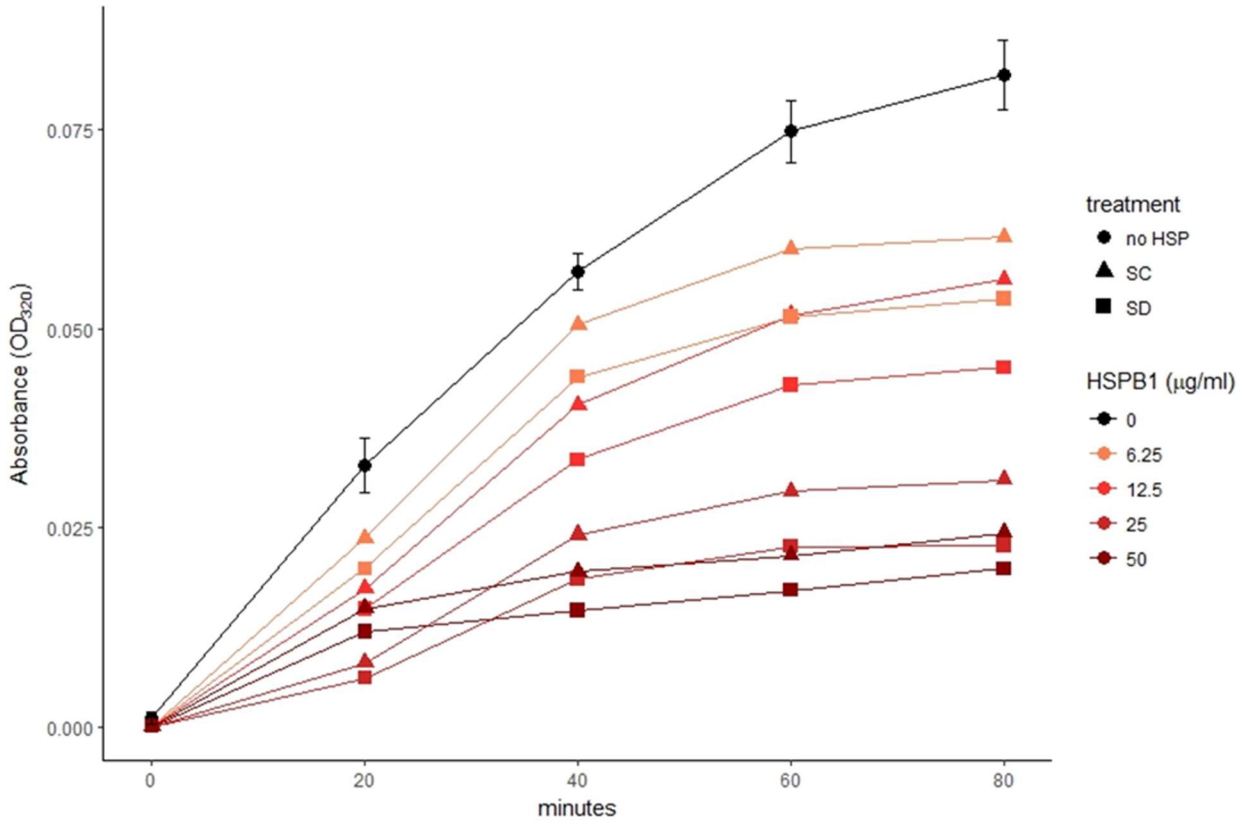


Figure 3-3: Effects of HSPB1 on thermal aggregation of CS. SD or SC HSPB1 at 6.25, 12.5, 25 and 50 µg/ml were incubated at 45ºC with CS (10 µg/ml). Insoluble CS aggregates formed over time were detected by light scattering at OD320. The OD of no HSP control is shown with standard errors (n= 4).

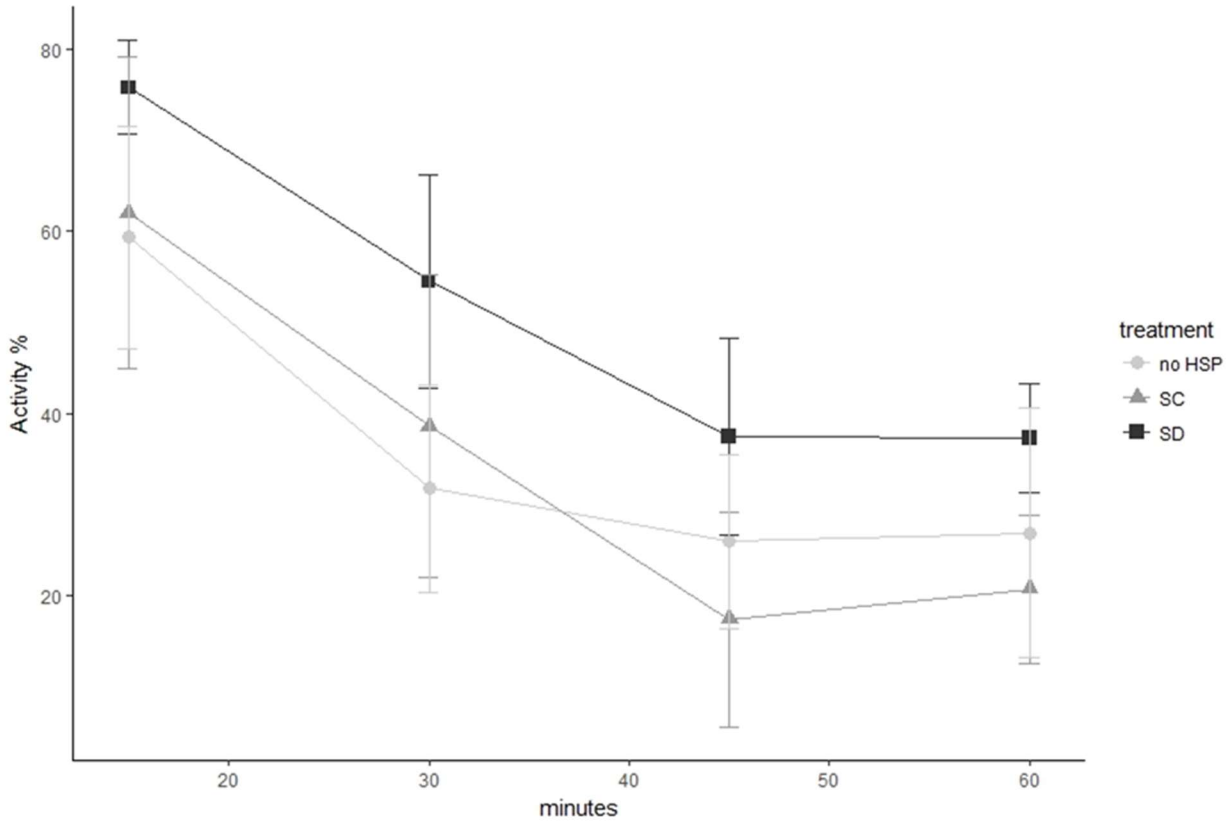


Figure 3-4: Effects of HSPB1 on thermal inactivation of CS activity. SD or SC HSPB1 (172 µg/ml for each) were incubated with 0.5 µg/ml CS (0.329 units/µg) CS at 44°C. The deactivation of CS is shown as the % relative remaining activity (relative to activity before 44 °C exposure) with standard errors.

Appendix

Table A3-1: Locations of *T. californicus* used in this study

Populations	Geographical Coordinates
Pacific Rim Reserve (PAC)	48°49N 125°09W
Friday Harbor Laboratory (FHL)	48°30N 123°09W
Pescadero (PES)	37°15 N 122°24W
Santa Cruz (SC)	36°57 N 122°03W
Point Lobos (PL)	36°30 N 121°57W
Montaña del Oro (MDO)	35°15 N 120°53W
Abalone Cove (AB)	33°44 N 118°22W
Laguna Beach (LB)	33°32 N 117°46W
Bird Rock (BR)	32°48 N 117°16W
San Diego (SD)	32°45 N 117°15W
Bufadora (BUF)	31°43 N 116°43W

Table A3-2: Mapping of parental population reads on hybrid references.

Population	Treatment	Reads mapped on SD	Reads mapped on SC
SD	control	86	6
SD	control	252	11
SD	heat stress	42373	20
SD	heat stress	27513	19
SC	control	61	177
SC	control	64	187
SC	heat stress	1910	9614
SC	heat stress	2850	13738

Table A3-3: Pairwise population dN/dS ratios (ω: Top; dN,dS: Bottom)

A. HSF

	SD	BR	AB	SC	PES
SD		0	0.0560	0.1251	0.1327
BR	0,0.0085		0.0624	0.1411	0.1509
AB	0.0033,0.0586	0.0033,0.0523		0.0810	0.0874
SC	0.0066,0.0525	0.0065,0.0464	0.0033,0.0403		0
PES	0.0066,0.0495	0.0065,0.0434	0.0033,0.0523	0,0.0028	

B. HSPB1

	SD	BR	AB	SC	PES
SD		0.3590	0.2576	0.4440	0.2565
BR	0.0079,0.0219		0.2428	0.4918	0.2694
AB	0.0286,0.1111	0.0270,0.1112		0.3634	0.2114
SC	0.0393,0.0886	0.0377,0.0766	0.0392,0.1079		0.3879
PES	0.0368,0.1443	0.0351,0.1304	0.0344,0.1581	0.0228,0.0588	

Table A3-4: Primers and restrictions endonucleases used in ASE experiments

Cross	Forward primers	Reverse primers	Product size	Restriction endonuclease	Allele cut
SDxBR	GATTGCGGTGATCATTTTGA	GGGAGTCTCATCGACATCCT	291	MseI	BR
SDxSC BRxPES	CAACCTCGATTGGCATTACA	TCCATTCGACAATTCTTTGG	351	MseI	SD, BR SC, PES
SCxPES	GTGGTTAGCATTTCGGGACA	TGATGGCTAATTTCCGAGGT	202	HincII	SC

Figure A3-1: Full sequence alignment of HSPB1 promoters from 11 *T. californicus* populations

A HSF

B HSPB1

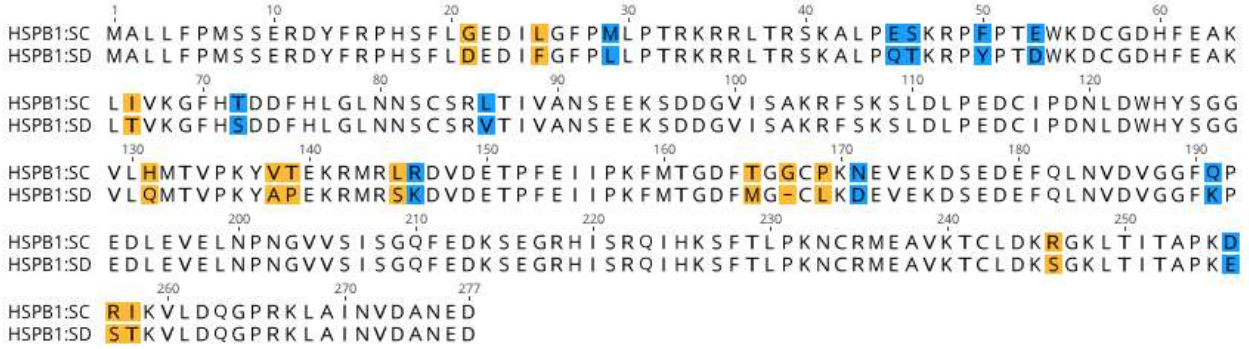


Figure A3-2: Protein alignment of HSF (A) and HSPB1 (B) between SC and SD populations. Blue highlights depict similar amino acid substitutions (similar charge or functional groups). Yellow highlights show dissimilar amino acid substitutions. HSPB1 sequences shown here were amplified from cDNA and cloned into *E. coli* for CS aggregation and CS activity assays. There are total of 23 amino acid substitutions between two HSPB1 alleles while the RNA sequencing has identified 17 of these are fixed substitutions.

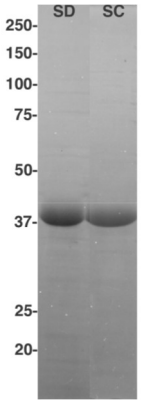


Figure A3-3: Coomassie Blue stained 10% acrylamide gel of purified SD and SC HSP beta 1. kDa MW standards are marked on the left side of the gel. The predicted molecular weight of both recombinant HSP beta 1 proteins is approximately 35 kDa.

References

Alahiotis SN. 1982. Adaptation of *Drosophila* enzymes to temperature IV. Natural selection at the alcohol-dehydrogenase locus. Genetica 59:81-87.

Amin J, Ananthan J, Voellmy R. 1988. Key features of Heat-Shock Regulatory Elements. Molecular and Cellular Biology 8:3761-3769.

Barreto FS, Moy GW, Burton RS. 2011. Interpopulation patterns of divergence and selection across the transcriptome of the copepod *Tigriopus californicus*. Molecular Ecology 20:560-572.

Barreto FS, Pereira RJ, Burton RS. 2015. Hybrid Dysfunction and Physiological Compensation in Gene Expression. Molecular Biology and Evolution 32:613-622.

Barreto FS, Schoville SD, Burton RS. 2015. Reverse genetics in the tide pool: knock-down of target gene expression via RNA interference in the copepod *Tigriopus californicus*. Molecular Ecology Resources 15:868-879.

Carroll SB. 2005. Evolution at two levels: On genes and form. Plos Biology 3:1159-1166.

de Jong WW, Caspers GJ, Leunissen JAM. 1998. Genealogy of the alpha-crystallin - small heat-shock protein superfamily. International Journal of Biological Macromolecules 22:151-162.

Dethier MN. 1980. Tidepools as refuges: predation and the limits of the Harpacticoid copepod *Tigriopus californicus* (Baker). Journal of Experimental Marine Biology and Ecology 42:99-111.

Dierick I, Irobi J, Janssens S, Theuns J, Lemmens R, Jacobs A, Corsmit E, Hersmus N, Van Den Bosch L, Robberecht W, et al. 2007. Genetic Variant in the HSPB1 Promoter Region Impairs the HSP27 Stress Response. Human Mutation 28:830-830.

Edmands S. 2001. Phylogeography of the intertidal copepod *Tigriopus californicus* reveals substantially reduced population differentiation at northern latitudes. Molecular Ecology 10:1743-1750.

Fangue NA, Hofmeister M, Schulte PM. 2006. Intraspecific variation in thermal tolerance and heat shock protein gene expression in common killifish, *Fundulus heteroclitus*. Journal of Experimental Biology 209:2859-2872.

Feder ME, Hofmann GE. 1999. Heat-shock proteins, molecular chaperones, and the stress response: Evolutionary and ecological physiology. Annual Review of Physiology 61:243-282.

Fernandes M, Xiao H, Lis JT. 1995. Binding of heat shock factor to and transcriptional activation of heat shock genes in Drosophila. Nucleic Acids Research 23:4799-4804.

Fernandes M, Xiao H, Lis JT. 1994. Fine-structure analyses of the *Drosophila* and *Saccharomyces* heat-shock factor - heat-shock element interactions. Nucleic Acids Research 22:167-173.

Frydenberg J, Hoffmann AA, Loeschcke V. 2003. DNA sequence variation and latitudinal associations in hsp23, hsp26 and hsp27 from natural populations of *Drosophila melanogaster*. Molecular Ecology 12:2025-2032.

Ganz HH, Burton RS. 1995. Genetic differentiation and reproductive incompatibility among Baja California populations of the copepod *Tigriopus californicus*. Marine Biology 123:821-827.

Gleason LU, Burton RS. 2015. RNA-seq reveals regional differences in transcriptome response to heat stress in the marine snail *Chlorostoma funebralis*. Molecular Ecology 24:610-627.

Gong WN, Yue M, Xie BY, Wan FH, Guo JY. 2009. Inhibition of Citrate Synthase Thermal Aggregation In Vitro by Recombinant Small Heat Shock Proteins. Journal of Microbiology and Biotechnology 19:1628-1634.

Graham AM, Merrill JD, McGaugh SE, Noor MAF. 2012. Geographic Selection in the Small Heat Shock Gene Complex Differentiating Populations of *Drosophila pseudoobscura*. Journal of Heredity 103:400-407.

Healy TM, Tymchuk WE, Osborne EJ, Schulte PM. 2010. Heat shock response of killifish (*Fundulus heteroclitus*): candidate gene and heterologous microarray approaches. Physiological Genomics 41:171-184.

Hoekstra HE, Coyne JA. 2007. The locus of evolution: Evo devo and the genetics of adaptation. Evolution 61:995-1016.

Hothorn T, Bretz F, Westfall P. 2008. Simultaneous inference in general parametric models. Biometrical Journal 50:346-363.

Hristozova N, Tompa P, Kovacs D. 2016. A Novel Method for Assessing the Chaperone Activity of Proteins. Plos One 11.

Huang L, Min JN, Masters S, Mivechi NF, Moskophidis D. 2007. Insights into function and regulation of small heat shock protein 25 (HSPB1) in a mouse model with targeted gene disruption. Genesis 45:487-501.

Juneja P, Quinn A, Jiggins FM. 2016. Latitudinal clines in gene expression and *cis*-regulatory element variation in *Drosophila melanogaster*. BMC Genomics 17.

Kelly MW, Grosberg RK, Sanford E. 2013. Trade-Offs, Geography, and Limits to Thermal Adaptation in a Tide Pool Copepod. American Naturalist 181:846-854.

Lerman DN, Feder ME. 2001. Laboratory selection at different temperatures modifies heat-shock transcription factor (HSF) activation in *Drosophila melanogaster*. Journal of Experimental Biology 204:315-323.

Lerman DN, Michalak P, Helin AB, Bettencourt BR, Feder ME. 2003. Modification of heat-shock gene expression in *Drosophila melanogaster* populations via transposable elements. Molecular Biology and Evolution 20:135-144.

Lindquist S. 1986. The Heat-Shock Response. Annual Review of Biochemistry 55:1151-1191.

Lockwood BL, Julick CR, Montooth KL. 2017. Maternal loading of a small heat shock protein increases embryo thermal tolerance in *Drosophila melanogaster*. Journal of Experimental Biology 220:4492-4501.

Maarek YS, Jacovi M, Shtalhaim M, Ur S, Zernik D, BenShaul IZ. 1997. WebCutter: A system for dynamic and tailorable site mapping. Computer Networks and Isdn Systems 29:1269-1279.

Mathelier A, Zhao X, Zhang AW, Parcy F, Worsley-Hunt R, Arenillas DJ, Buchman S, Chen C-y, Chou A, Ienasescu H, et al. 2014. JASPAR 2014: an extensively expanded and updated open-access database of transcription factor binding profiles. Nucleic Acids Research 42:D142-D147.

McDonald JH, Kreitman M. 1991. Adaptive protein evolution at the ADH locus in *Drosophila*. Nature 351:652-654.

McManus CJ, Coolon JD, Duff MO, Eipper-Mains J, Graveley BR, Wittkopp PJ. 2010. Regulatory divergence in *Drosophila* revealed by mRNA-seq. Genome Research 20:816-825.

Parsell DA, Lindquist S. 1993. The function of Heat-Shock Proteins in stress tolerance - degredation and reactivation of damaged proteisn. Annual Review of Genetics 27:437-496.

Pelham HRB. 1982. A regulatory upstream promoter element in the *Drosophila* HSP 70 Heat-Shock gene. Cell 30:517-528.

Pereira RJ, Barreto FS, Pierce NT, Carneiro M, Burton RS. 2016. Transcriptome-wide patterns of divergence during allopatric evolution. Molecular Ecology 25:1478-1493.

Perisic O, Xiao H, Lis JT. 1989. Stable binding of *Drosophila* heat-shock factor to head-to-head and tail-to-tail repeats of a conserved 5-bp recognition unit. Cell 59:797-806.

Perutz MF. 1983. Species adaptation in a protein molecule. Molecular Biology and Evolution 1:1-28.

Peterson DL, Kubow KB, Connolly MJ, Kaplan LR, Wetkowski MM, Leong W, Phillips BC, Edmands S. 2013. Reproductive and phylogenetic divergence of tidepool copepod populations across a narrow geographical boundary in Baja California. Journal of Biogeography 40:1664-1675.

Pinheiro J, Bates D, DebRoy S, Sarkar D. R Core Team.(2017). nlme: Linear and Nonlinear Mixed Effects Models [Software].

Place AR, Powers DA. 1979. Genetic variation and relative catalytic efficiencies: Lactate dehydrogenase B allozymes of *Fundulus heteroclitus*. Proceedings of the National Academy of Sciences of the United States of America 76:2354-2358.

Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26:139-140.

Schoville SD, Barreto FS, Moy GW, Wolff A, Burton RS. 2012. Investigating the molecular basis of local adaptation to thermal stress: population differences in gene expression across the transcriptome of the copepod *Tigriopus californicus*. BMC Evolutionary Biology 12.

Schulte PM, Glemet HC, Fiebig AA, Powers DA. 2000. Adaptive variation in lactate dehydrogenase-B gene expression: Role of a stress-responsive regulatory element. Proceedings of the National Academy of Sciences of the United States of America 97:6597-6602.

Seo JS, Lee YM, Park HG, Lee JS. 2006. The intertidal copepod *Tigriopus japonicus* small heat shock protein 20 gene (Hsp20) enhances thermotolerance of transformed *Escherichia coli*. Biochemical and Biophysical Research Communications 340:901-908.

Stoletzki N, Eyre-Walker A. 2011. Estimation of the Neutrality Index. Molecular Biology and Evolution 28:63-70.

Sørensen JG, Dahlgaard J, Loeschcke V. 2001. Genetic variation in thermal tolerance among natural populations of *Drosophila buzzatii*: down regulation of Hsp70 expression and variation in heat stress resistance traits. Functional Ecology 15:289-296.

Sørensen JG, Kristensen TN, Loeschcke V. 2003. The evolutionary and ecological role of heat shock proteins. Ecology Letters 6:1025-1037.

Tirosh I, Reikhav S, Levy AA, Barkai N. 2009. A Yeast Hybrid Provides Insight into the Evolution of Gene Expression Regulation. Science 324:659-662.

Tomanek L, Somero GN. 1999. Evolutionary and acclimation-induced variation in the heat-shock responses of congeneric marine snails (genus *Tegula*) from different thermal habitats: Implications for limits of thermotolerance and biogeography. Journal of Experimental Biology 202:2925-2936.

Willett CS. 2010. Potential fitness trade-offs for thermal tolerance in the intertidal copepod Tigriopus californicus. Evolution 64:2521-2534.

Willett CS, Burton RS. 2001. Viability of cytochrome C genotypes depends on cytoplasmic backgrounds in *Tigriopus californicus*. Evolution 55:1592-1599.

Wirgin I, Roy NK, Loftus M, Chambers RC, Franks DG, Hahn ME. 2011. Mechanistic Basis of Resistance to PCBs in Atlantic Tomcod from the Hudson River. Science 331:1322-1325.

Wray GA. 2007. The evolutionary significance of *cis*-regulatory mutations. Nature Reviews Genetics 8:206-216.

Wu C. 1995. Heat shock transcription factors: Structure and regulation. Annual Review of Cell and Developmental Biology 11:441-469.

Yang ZH. 2007. PAML 4: Phylogenetic analysis by maximum likelihood. Molecular Biology and Evolution 24:1586-1591.

Conclusion

Results from this dissertation expands knowledge on both ecological and evolutionary aspects of thermal tolerance. In first chapter, I found that early life stages of *T. californicus* could survive acute heat stress at higher temperatures than adults. This contradicts a popular belief that early life stages are more sensitive to environmental stresses. Since all life stages of *T. californicus* inhabit the same tidepool habitats and experience the same temperature changes, selection would favor more tolerant early life stages so that they can survive until reaching reproductive age. From this finding, we could use adult's upper thermal limit to predict populations' future survival and distribution. In addition, heat stress during larval stage comes with a cost; heat stressed copepods delayed metamorphosis compared to their siblings that did not experience heat stress. There are a couple research directions that can be expanded from our findings in chapter 1. First, I only used 1-hour acute heat stress treatment to look for survivorship and developmental effects. However, this did not reflect how the animals experience temperature fluctuation in their habitats. Since then, we now obtained temperature logger data from tidepools. We can use this data to design future heat stress experiments. Besides improving future experiments with more relevant temperature pattern, investigating the effects of chronic heat stress seems appropriate as mean temperatures are projected to increase with climate change.

In second and third chapters, the objective was to investigate the genetic basis underlying thermal tolerance. Chapter 2 used RNA sequencing of F1 hybrids and their parental populations to identify *cis* and *trans* regulatory divergence and how expression phenotypes are inherited. Results from chapter 2 revealed candidate genes underlying thermal tolerance that exhibit regulatory divergence between populations. These genes span over different functional groups including electron carrier, muscle and cuticle assembly, and heat shock response. These lists

provide candidates for future investigations on the physiological processes underlying heat stress response and genetic basis underlying thermal adaptation. In *T. californicus* system, we can use RNAi to knock down transcription one of these candidate genes to determine its functional significance. We can also investigate the roles of theses gene categories in response to heat stress beyond *T. californicus* system. For example, the role of chitin metabolism during heat stress response remains mostly as speculations despite many organisms having chitin structures.

In chapter 3, we further investigated on HSPB1 which was one of the highest induced HSPs in response to heat stress and Chapter 2 results indicate *cis* divergence with allelic imbalance toward heat tolerant SD population only under heat stress treatment. We tested for functional differences in both *cis* regulatory and coding sequence divergence in HSPB1 between populations. For *cis* regulatory divergence, we identified potential causal mutations as SNPs in the conserved motif of HSEs in the promoter region of HSPB1. Allele specific expression in more F1 hybrid crosses showed biased expression favoring alleles from more heat tolerant populations with 2 canonical HSEs in HSPB1 promoter. This allelic imbalance validated the SNPs in HSEs as the potential cause underlying differential expression of HSPB1 and heat tolerance across populations. For coding sequence divergence, we found that recombinant HSPB1 from heat tolerant SD population is more efficient at preventing protein aggregation and protecting enzyme function at elevated temperatures than HSPB1 from SC population. Several research aspects can be explored further from chapter 3 results. For example, though we used the best available tools to validate functional consequences of SNPs inside HSEs, the next step would be obtaining the definitive proof of HSF binding to HSPB1 promoter. HSPB1 gene expression results from chapter 2 and 3 were only a snap shot of a single gene response. Future projects could expand in different dimensions of both temperature gradient and time course of

HSPB1 gene expression both RNA and protein. In addition, future projects could also investigate the functional consequences of amino acid substitutions in HSF between populations even though we could not identify the functional difference of HSF in the allele specific expression experiment.

Overall, thermal tolerance is a complex trait that involves many processes across biological organizations from the genetic level to the organismal level. This dissertation has explored different aspects of thermal tolerance in *T. californicus*. In chapter 1, we demonstrated different survivorship phenotypes in different life stages. In chapter 2, even only looking at divergence in transcriptional regulation underlying thermal adaptation, we identified divergence underlying thermal tolerance in several gene categories. In chapter 3, even within a single gene, we found that selection could act on different mechanisms to achieve increase of heat tolerance.