

PHYLOGENETICS IN SPACE  
AND  
THE EVOLUTION OF THE HAUSTORIIDAE

A Dissertation

By

ZACHARY BROCK HANCOCK

Submitted to the Office of Graduate and Professional Studies of Texas A&M University in  
partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Mary K. Wicksten
Committee Members,	Jessica E. Light
	Anja Schulze
	Gregory A. Sword
Head of Department,	Thomas McKnight

December 2020

Major Subject: Ecology & Evolutionary Biology

Copyright 2020 Zachary Brock Hancock

## ABSTRACT

The tendency to discretize biology permeates taxonomy and systematics, leading to models that simplify the often continuous nature of populations. Even when the assumption of panmixia is relaxed, most models still assume some degree of discrete structure. In this dissertation, I review the many lines of evidence for how continuous spatial structure can impact phylogenetic inference. I illustrate and expand on these by using a combination of stepping-stone models and complex continuous-space demographic models that include distinct modes of speciation, as well as empirical datasets. I find that the impact of spatial structure permeates all aspects of phylogenetic inference, including gene tree stoichiometry, topological and branch-length variance, network estimation, divergence-time estimation, and species delimitation. Furthermore, I uncover dramatic genome size variation and possible historic drivers of divergence in a family of sand-burrowing crustaceans (Haustoriidae: Amphipoda). I conclude by utilizing my results to suggest how researchers can identify spatial structure in phylogenetic datasets.

## DEDICATION

*This is for the great friends I've made along the way. Especially the following people, whose persistent faith in me I find unnerving:*

*Faith Olivia*

*Emma Lehmberg*

*Thomas Strawn*

*and Dan Powell.*

*Thanks for always being down for a drink.*



“This Way to the Playa”. Photo by Zach Hancock.

## ACKNOWLEDGMENTS

This work is the result of many great collaborations and conversations across the diverse program that is EEB. Firstly, I'd like to thank Dr. Mary Wicksten for accepting me into her lab and affording me the freedom to explore my interests. One of these days, we'll take that boat to Matagorda Island. In addition, I'd like to thank my labmate Maureen Hayden for being around to bounce ideas off of and her willingness to make periodic treks down to the beach (odious task, I know!).

A special thanks to my committee member and collaborator Dr. Jessica Light, who from the beginning welcomed me into her lab, invited me to her lab meetings and gatherings, and provided a mountain of helpful comments on our manuscript (Chapter I of this dissertation). Without her help, this dissertation would not exist in its present form. Furthermore, I would like to thank Dr. Adrian Castellanos and Bridgett Benedict, who always tolerated my presence in the lab and were good sports about my many questions.

A second special thanks to Dr. Heath Blackmon, who, you'll notice, is the last author on the manuscripts of the final two chapters of this dissertation. You can blame him almost exclusively for Chapter III, which was the result of a question he asked during a Biology seminar. Thanks for seeing potential in me.

I'd like to thank Justin Hilliard and Drs. Anja Schulze and María Nuria Méndez Ubach for letting me crash their sampling trip along the Mexico coastline, and Thomas and Shelby Strawn for help sampling across the Gulf states. The samples I was able to collect during these

trips made Chapter I possible. Thanks also to Dr. José Salgado-Barragán and Sergio Rendón Rodríguez for driving the Mexican coastline – I enjoyed all the Coronas and blue crab.

The entire lab of Dr. Gil Rosenthal has a special place in the formation of this dissertation. In no particular order, thanks to Emma Lehmberg, Dr. Dan Powell, Dr. Gaston Jofré, Stephen Bovio, Matteo García, and Megan Exnicios for all the great conversations about science, life, and the state of the world. I'd also like to thank Gil for letting me tag along on the many trips to Calnali, Hidalgo – CICHAZ will always have a special place in my heart.

Finally, I'd like to thank my great friend and partner Faith Olivia, who helped out on most of the projects presented here and tolerated my many quirks during the seemingly endless months of quarantine leading up to the completion of this dissertation. You're the best.

## CONTRIBUTORS & FUNDING SOURCES

### Contributors

This work was supervised by a dissertation committee consisting of Professor Mary K. Wicksten [chair] of the Department of Biology and Professors Jessica E. Light of the Department of Ecology and Conservation Biology, Gregory A. Sword of the Department of Entomology, and Anja Schulze of the Department of Marine Biology at Texas A&M University-Galveston.

Chapter I was performed in the lab of Jessica E. Light, and analyses were conducted with the aid of Faith O. Hardin in the Department of Wildlife and Fisheries Sciences. The results of this chapter were published in 2019 in the *Journal of Biogeography* (Hancock et al. 2019).

Chapter II was performed in the lab of Jessica E. Light and J. Spencer Johnston of the Department of Entomology, collections were performed with the aid of Faith O. Hardin, and Archana Murthy, an undergraduate in the Department of Biology, performed DNA extractions and amplification. Results of this chapter were submitted to *Genome Biology and Evolution*.

Chapter III was performed in collaboration with Heath Blackmon in the Department of Biology. The results of this chapter are currently in press at the *Journal of Heredity*.

Chapter IV was performed in collaboration with Emma S. Lehmberg in the Department of Biology and Ecology & Evolutionary Biology IDP, alongside Heath Blackmon.

### Funding

The projects included in this dissertation were funded by Grants-In-Aid of Research from Texas Sea Grant, Texas Ecolabs, and the Ecology & Evolutionary Biology New Research Grant.

## TABLE OF CONTENTS

	Page
ABSTRACT.....	iv
DEDICATION.....	viii
ACKNOWLEDGEMENTS.....	ix
CONTRIBUTORS AND SOURCES OF FUNDING.....	xi
TABLE OF CONTENTS.....	xii
LIST OF FIGURES.....	xiv
LIST OF TABLES.....	xvi
PRELUDE.....	1
<b><u>VOLUME I EVOLUTION OF THE HAUSTORIIDAE</u></b> .....	<b>12</b>
CHAPTER I PHYLOGEOGRAPHY OF THE HAUSTORIIDAE IN THE GULF OF MEXICO.....	13
Introduction.....	13
Methods.....	19
Results.....	25
Discussion.....	34
CHAPTER II RAPID GENOMIC EXPANSION AND PURGING	

ASSOCIATED WITH HABITAT TRANSITIONS.....	43
Introduction.....	43
Methods.....	47
Results.....	53
Discussion.....	57
 INTERLUDE.....	64
 <u>VOLUME II PHYLOGENETICS IN SPACE.....</u>	72
 CHAPTER III IMPACTS OF ANCESTRAL ISOLATION-BY-DISTANCE ON DIVERGENCE-TIME ESTIMATION.....	73
Introduction.....	73
Methods.....	79
Results.....	85
Discussion.....	90
 CHAPTER IV PHYLOGENETICS IN SPACE .....	96
Trees in Space: Models.....	100
Slatkin's Skew and Gene Tree Asymmetry.....	102
Wicked Forests: Species-Tree Inference and Gene Tree Variance.....	107
Species Delimitation Delimits Structure.....	113
Out of Space: Continuity for Discontinuous Data.....	119
 BENEDICTION.....	130
 REFERENCES.....	138

APPENDIX A..... 163

APPENDIX B..... 166

## LIST OF FIGURES

FIGURE		Page
1.1	Biogeographic hypotheses in the Gulf of Mexico.....	14
1.2	Distribution of sampling sites in the Gulf of Mexico.....	17
1.3	Total evidence tree produced in MrBayes.....	26
1.4	Divergence-time tree produced in BEAST2.....	28
1.5	Dated species-tree inferred using *BEAST.....	31
1.6	Pairwise $F_{ST}$ matrix.....	34
1.7	Effective population sizes ( $N_e$ ) over time.....	37
2.1	Map of sample sites.....	45
2.2	Calibrated phylogeny of the Haustoriidae.....	54
2.3	Repetitive content.....	55
2.4	Ancestral state reconstructions.....	60
2.5	Transitions to brackish water.....	63
3.1	Population model for SLiM simulations.....	78
3.2	Box plots of the estimated $T_{MRCA}$ in SNAPP.....	88

3.3	Density plot of scaled ancestral $N_e$ .....	90
3.4	Estimated divergence-time for the Domingos et al. (2017) dataset.....	93
4.1	Phylogenetics in space.....	100
4.2	Demonstrating Slatkin's Skew.....	105
4.3	Gene tree variance.....	112
4.4	Substitution-rate estimates from *BEAST2.....	114
4.5	Species delimitation in continuous populations.....	119
4.6	How space impacts phylogenetic datasets.....	123

## LIST OF TABLES

TABLE	Page	
1.1	Summary statistics.....	33
2.1	Summary of samples and their genome sizes.....	49
2.2	Summary of genomic correlation models.....	57

## PRELUDE

“The man who believes that the secrets of the world are forever hidden lives in mystery and fear. Superstition will drag him down. The rain will erode the deeds of his life. But that man who sets himself the task of singling out the thread of order from the tapestry will by the decision alone have taken charge of the world and it is only by such taking charge that he will effect a way to dictate the terms of his own fate.”

—Cormac McCarthy, *Blood Meridian*

“Of the many marine environments populated by macroscopic animal life, perhaps none is more difficult of access or more challenging to body form and function than the unstable, highly oxygenated, silt-free sands of wave-exposed intertidal beaches.”

—E.L. Bousfield, “Adaptive radiation in sand-burrowing amphipod crustaceans”

My initial inclination was to title this section *A Call to Worship*. The southern Baptist church I attended throughout my young life distributed a folded pamphlet as you entered the sanctuary. The pamphlet, printed on course yellow paper with pressed green writing as if someone had individually stamped each letter into place like a medieval monk, detailed the order of events for the service. First tended to be announcements – the business of when, where, and what the congregation had planned over the course of the foreseeable future. These usually bland announcements that meant next to nothing to a child was followed by the “call to worship”. This was initiated by gentle notes from the piano and the organ and was meant as a period of preparation to accept the message to be delivered. In a way, it’s an introduction.

But where to begin the task of introducing the science that has colored the past 5 years of one’s life? “In the beginning” may have been an ideal way to start the Pentateuch, but the origin of life seems unnecessarily far removed from marine crustaceans or the field of phylogenetics. Traditionally, one might begin by reviewing the literature – so-and-so did this (citation, citation), which was expanded by so-and-so (citation). A string of white men and their accomplishments in the admittedly obscure field of amphipod systematics may be an expected introduction.

I've never cared much for tradition. It's painful enough to follow the rigid formatting guidelines imposed by the esteemed Office of Graduate and Professional Studies (OGAPS). Regardless if the rest of the world is a dumpster fire, thou shalt *not* use bold font! I get it: people like uniformity. Especially paper-pushers and administrative bureaucrats. As my old history professor, Tom Wagy, would say, "Wagy's First Law of Bureaucracy is: *never argue with the person at the front desk*". They don't have the power to help you and they're not really sure who does.

The format I've adopted here should not deter you. In fact, I hope it saves you from slogging through endless jargon and details that you don't really need to understand the point of each chapter. Furthermore, each chapter has been written in the form of the published or submitted manuscript, complete with a section titled "Introduction" all its own. There you can find all the historic literature you want. Instead, I use the three sections "Prelude", "Interlude", and "Benediction" as the *me* in all of this – the human behind the science. I adopt this approach because this dissertation means more to me than the sum of each chapters' results. Because some of the most interesting parts of a seminar talk is when the speaker divulges a hint of what it was like in the field – some especially interesting locals or equipment malfunctions, etc. It's a nice break, a time to allow you to both ruminate on the information you've received thus far while connecting with the speaker on a human level. This dimension – our shared humanity – is too often neglected in academia.

This dissertation can be read in two ways: 1) as four chapters on phylogenetics and systematics with some personal anecdotes in between; or 2) as my story through graduate school interwoven with some real science. Those who favor the former option may choose to skip the

rest of the Prelude, the Interlude, and the Benediction, and just focus on the data chapters. Those with interest in the latter, I hope you enjoy the stories that are to follow.

Listen now as the organ begins to play.

\*\*\*

I was the oldest person in my cohort at 27. I was also the only one that was married with a stepdaughter, and the only one that did not go straight to college after finishing high school. I had virtually no college aspirations while in grade school, and in fact barely graduated. My fellow classmates likely remember me as a trouble-maker – those who remember me at all, given that two of the four years of high school I spent in a form of juvenile detention. Given this background, I was well set-up for a life in the service industry. I worked a series of odd jobs, mostly fast-food but also retail. I hated these. If anything can be blamed for my socialist politics, it was the decade I spent working for incompetent managers for minimum wage in conditions most Americans would be appalled at keeping a dog in.

Despite this, I did not take my education seriously until I got married at 24. By this time, I had largely converted to a law-abiding citizen who managed a Domino's Pizza 50 hours a week. I married a woman I'd known since I was a child – she often joked about tossing me around on the playground at our elementary school – and she came with a 2-year-old who quickly took to calling me *Daddy*. Her real father had died in a drunk driving accident when she was 6 months old.

We lived in an efficiency apartment in Texarkana for most of this time. The place was infested with fleas and we'd suffer nightly raps at the door by homeless people who could be

found sleeping in the hallways querying about money for a bus ticket or diapers for their non-existent children. I'd get off work every night after midnight and cruise 7<sup>th</sup> Street, passing dilapidated buildings with boarded windows and the streetwalkers flagging you down at the intersections, towards MLK where only the flickering lights from the dozens of liquor stores illuminated the cracked, potholed street. Maybe it was on one of these drives that I decided to go back to school.

My evolution as an academic is meandering. I started as an English major because my life-long dream was to become a novelist; my mother fostered this (she is a high school English teacher) by feeding me classical literature and engaging in hours of conversation about the exact meaning of every word. I must have submitted over a hundred short stories and three novels to different publishing agencies and literary magazines from the time I was 16 until I was 25 or so. In all this time, I successfully published only two short stories. I kept every rejection slip in a drawer on my desk because I had read that is what all famous writers did.

Literature naturally led me into history, which I would eventually minor in. If there was a common thread between these it was that both provide insights into the human condition – why are we here? What is our purpose? Why is the world the way that it is? History seemed to offer a window into these questions, but I consistently found that this window was opaque. There were no answers here; no unifying principles that could explain our existence. The childhood version of myself was comfortable with the southern Baptist explanations; but anyone who reads enough literature, studies enough history, will eventually find these unsatisfactory.

My academic path turned toward science after I read Richard Dawkins and Yan Wong's 2004 book, *The Ancestor's Tale*. The Chaucerian narrative immediately appealed to my English major's literary sense, and I was enamored at their ability to present, with reasoned argument and

evidence, the entire history of life one ancestor at a time. I quickly burned through Dawkins' entire catalogue, eventually making my way to the writings of Stephen Jay Gould. His 1989 book *Wonderful Life* about the Burgess Shale and the evolutionary trajectory of body forms (what he called *bauplans*) I credit for sparking my passion for invertebrate zoology and evolution. His collections of essays, *Ever Since Darwin* (1977) and *Bully for Brontosaurus* (1991), provided me with a political perspective into the history of science and evolutionary biology, and challenged the simplistic, adaptationist view of evolution often espoused by Dawkins and his ilk. These readings coincided with an undergraduate course in invertebrate zoology, taught by my undergrad mentor David Allard (whom I would later name a species for: *Haustorius allardi*; Hancock & Wicksten 2018).

Perhaps important to mention is that from the time I enrolled at Texas A&M University-Texarkana (TAMUT) I intended to get my PhD. This has not wavered; from the time I made this decision until now, I have never doubted that academia is where I want to be. Throughout my undergrad, I worked full-time managing a Domino's Pizza at night and attending classes full-time during the day. This undoubtedly put a strain on my marriage, but it would at least survive until my first year at graduate school.

I have Dr. Allard to thank for being in the lab of Dr. Mary Wicksten. We often chatted over lunch in the cafeteria about future goals, and I told him that I would probably apply to Texas A&M since my wife wanted to stay within driving distance of her hometown. He knew I had a passion for crustaceans – I completed an undergraduate project with him as the advisor sampling and identifying microcrustaceans across the various cow ponds that the university had acquired – and so he preemptively reached out to Dr. Wicksten on my behalf. I did not apply to any other graduate program.

I had originally intended on continuing research on cladocerans. I had some half-baked ideas about neoteny and larval development in daphniids. Of course, I ended up with several boxes of thousands of haustoriid amphipods in vials of ethanol that had been collected by previous students, and my mission was to go forth and see if I could identify any new species unique to the Texas coast. A student from the late 90s had written a dissertation on this very topic – the systematics of haustoriid amphipods – but he never formally described any species and gave them names like “genus 1 species a”. I was able to use some of his keys to help out in identifying the Texas coast ‘pods, but I ultimately found that most of his descriptions were either lacking in detail, confused the between location variation with interspecific differences, or I was completely unable to find the “species” supposedly at a given location. My first year was dedicated to this kind of traditional taxonomy – count spines, setae, dissect legs and mount them on slides of glycerin. I poured through old books from the 60s and 70s, mostly by James Lauren Barnard and E.L. Bousfield, and yellow-paged editions of the *Journal of Crustacean Biology*. Eventually I convinced myself that I understood the bizarre morphology of these obscure crustaceans well enough to actually describe them.

I should give an anecdote at this point about the first time I saw a haustoriid under a microscope. Instead of excitement or awe, I was actually filled with panic. If it hadn’t been for its beady white eyes, I wouldn’t have been able to tell its head from its ass. I remember poking and prodding it with forceps and probes, trying to find the different appendages that had been etched so neatly in McLaughlin’s *Biology of the Recent Crustacea*. I mistook legs for antennae, pleopods for uropods – it must’ve taken me three weeks to actually dissect a mouthpart that wasn’t completely obliterated beyond recognition. All the while, I was attending the first year EEB modules with my cohort that I knew at any moment would see me for the imposter I was.

What was this man from rural East Texas doing at a Tier 1 research institution? My first year was the only time during my academic career that I thought I might not be smart enough to do a PhD. I never thought about quitting – way too much pride for that – but I definitely questioned if I belonged. It wasn’t until much later that I found out that *everyone* thinks this way. I needed to hear it from the other people in my cohort to actually believe it, regardless of the seminars for first year grad students that assured me I wasn’t alone. Bullshit. Everyone looks like they have it together. Turns out we’re all just pretending. Fake it ‘till you make it, you know. And I did. I faked it until I figured it out.

By the end of my first year it had become clear that my marriage wasn’t going to survive, and we split that May. It was the one thing that faking it just never seemed to work. I’m sure I did not make it easy on her. I was obsessive with my work; I’d stay at the lab until late at night, and when I was home, I was distant, often lost in schoolwork or some new grant I was writing. She had never wanted to make the move in the first place and had begged me to take a job in Texarkana – just be a high school biology teacher or something. She followed me to College Station, but she was a shadow of her former self. I’m not sure why she came at all. Maybe it was for the financial security or because her daughter had taken to calling me *daddy*. Whatever her reasoning, she seemed to drift through each day, an empty hull, indifferent and stranded in a place she never wanted to be.

In the spring of my second year, I went on a sampling trip with a grad student from TAMU-Galveston and a team of Mexican scientists to the states of Veracruz, Tabasco, and Campeche. The entire cab ride from the airport in Mexico City to UNAM I was glued to the window. The city goes on forever, both outward and upward. Houses and buildings rise up the mountains that surround the city, painted bright pastels of pink and green and blue. Murals on

buildings depicted Aztec warriors and gods, cornfields with a lone farmer, beautiful women in flowing dresses. Despite the heat, doors were propped open and large box speakers blared mariachi. At every intersection, children waited to leap onto the hood of the van and squirt slimy green liquid from water bottles onto the windshield. The driver would drop a handful of pesos into their small hands. Without speaking a word of Spanish, I could only listen to the cacophony of sounds around me, the thick smell of the street – motor exhaust intertwined with sizzling *carne* from the open grills along every corner – and the soft chattering from the radio that the cabby occasionally honked over.



“No Pictures!”. Justin Hilliard and I sampling *Capitella*. Photo by Jose Salgado.

Nuri was the head scientist from Mazatlán, and she had a rambunctious, passionate personality. She hugged me the first time we met as if we were old friends (we'd only exchanged emails a few times), and she excitedly introduced us to her colleagues at the university. Each of these welcomed us warmly, and I must admit to never having felt such a welcome in the states. We toured their labs, most of which looked more like workshops where students tinkered with various limnological devices, and we listened as Nuri translated for us what each of the graduate

students were working on. After these pleasantries, we met Tama and Pepe – the two field hands that had agreed to drive the van and help us sample – in the parking lot.

One last thing about UNAM. Unlike the concrete monstrosities that are many American universities, UNAM somehow managed to be both sprawling but also a part of the jungle that grew up around it. The university feels as if it were built inside a nature park – everything is green, old growth forest that actually obscures the size of the university. While the facilities were old and the labs in dire need of upgrades, the campus itself was an escape from the bustling city that it was nestled within.

The sampling expedition was funded as a joint venture between TAMUG and UNAM with the purpose of collecting polychaete worms of the genus *Capitella*, indicators of pollution and environmental degradation. These worms live in sewage outflows and polluted bays, burrowing into the anoxic sediment where they thrive where nothing else lives. The grad student who led the TAMUG side was interested in collecting enough worms for genome and transcriptome sequencing, and on the plane ride to Mexico City we found that our two projects had a lot in common. By this time, I had decided that I would use some small-scale sequencing to help delineate cryptic haustoriid species and also to build a phylogeny of the haustoriid family tree to go along with my taxonomic work. He was doing a similar thing with the worms – biogeography, alongside searching for loci with signatures of selection that may contribute to their persistence in filth. We had many great conversations and to some degree his persistence that I could afford next-generation sequencing if I just planned well led to Chapter II.

I don't think it's a violation of their trust to say that we worked hard and played hard. Some days, we did one harder than the other. We drove the entire length of the eastern coastline, from Tamiahua in Veracruz down to Ciudad del Carmen in Campeche. We ate breakfast at noon,

lunch at four, and dinner at eight, and beer was a staple of each. We would stay up late into the evening after a long day of digging through mucky bays in search of polychaetes drinking and chatting about science, politics, the politics of science. Trump, a fascist, had been in office for a year at this point, and his harsh rhetoric about Mexico was still fresh on everyone's minds.



"Revelation 16:3 or March of the Cows". Dock leading from a slaughterhouse in Laguna de Términos. Photo by Zach Hancock.

At a field station on Laguna de Alvarado in southern Veracruz, we met dozens of undergrads from UNAM who spent three months at the end of their senior year performing water quality measurements in a makeshift wet lab overlooking the bay. On our last night, the students put on a mock gameshow for us in which volunteers from the audience participated in a series of competitions, from eating ghost peppers to who had the most impressive *grito*. The show ended with a panel of scientists, including Nuri, Tama, and Pepe, sharing their experiences in academia and a bit about their work – this, as Nuri would later explain to us, was a kind of elevator pitch to

the undergrads to go to grad school in their labs. Following the panel, we danced a variety of local routines, which I was terrible at, and then several of us sat around and talked politics.

At Ciudad del Carmen we pulled off onto a dirt road that ended at the beach. I would look for amphipods, and everyone else would enjoy the cool, crystal blue waters that characterize the Yucatan peninsula. A crowd had gathered on the shelly beach, pointing and murmuring, some snapping photos, all in awe of the large, algae-coated and barnacle-ridden yacht that had run aground some ten meters from the shore. We approached the crowd and asked if they knew why the boat was there. Nuri, grinning, explained to us that the boat had belonged to a politician in Veracruz that had been arrested on corruption charges. He could no longer make payments to the marina and so they cut the boat loose. A few weeks later the currents dumped it onto the barrier island. Several men had already managed to scale its towering hull and loot whatever might remain in the quarters of any value. Now it loomed as nothing more than an empty metal beast casting its shadow across the beach where I would dig for amphipods.

Nuri and Tama stayed to talk with the people. I think they talked politics, but I don't know. Pepe and the other grad student collected shells, filling their pockets and their shirts. The currents flowed such that the recently deceased bivalves were deposited onto the white sands of the barrier island; the shells all in pristine condition. I just sat on the sand, arms hugging my knees, and watched the waves pound the hull of the yacht – a hollow sound like ringing a bell underwater.

I didn't find a single amphipod on Ciudad del Carmen.

VOLUME I

THE EVOLUTION OF THE HAUSTORIIDAE

## CHAPTER I

# PHYLOGEOGRAPHY OF THE HAUSTORIIDAE IN THE GULF OF MEXICO<sup>1</sup>

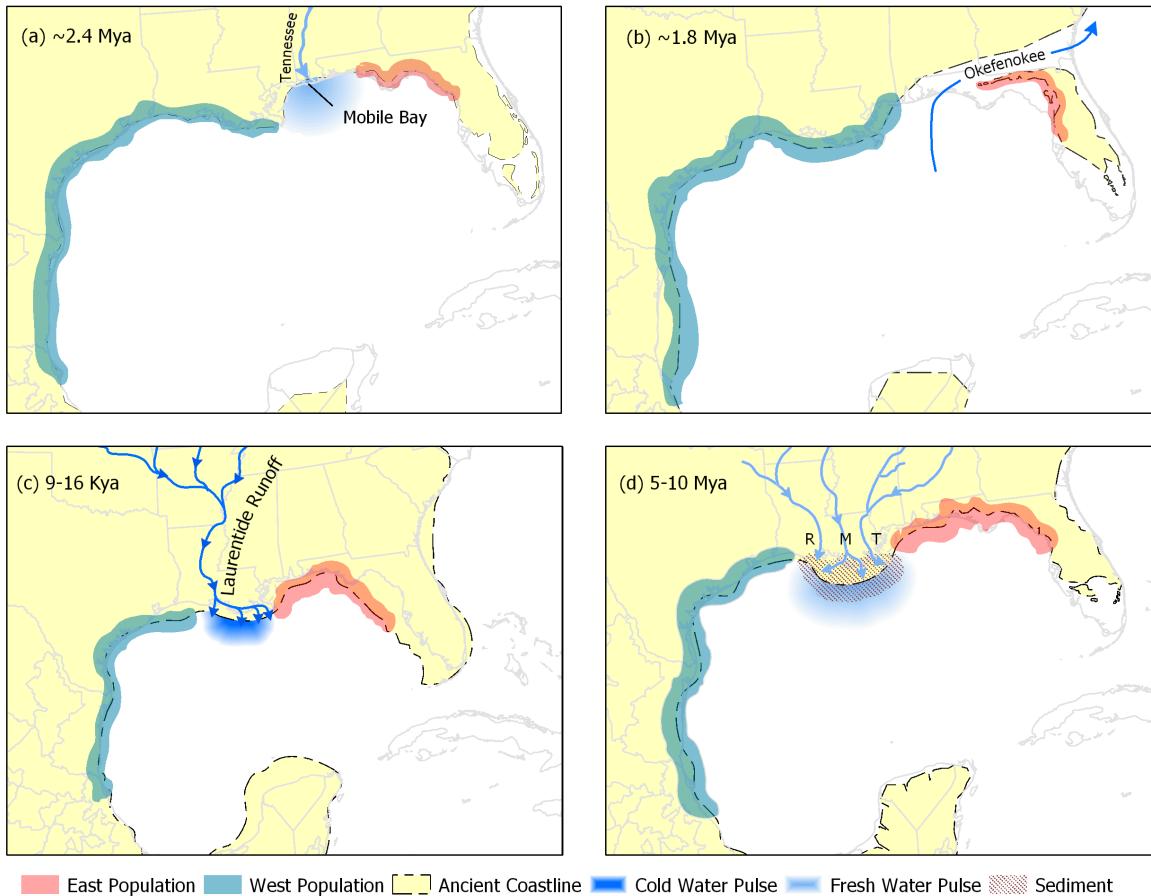
### Introduction

Suture zones in the Gulf of Mexico have been studied since at least the early 1900s (e.g., Simpson, 1900; Bert, 1986; Avise, 1992; Portnoy & Gold, 2012). These suture zones (or vicariant zones) are regions in the Gulf basin that severed population connectivity in the past, leading to divergence in genetic and/or morphological characters between the divided populations. Recent divergences along an east-west gradient within the Gulf of Mexico have been noted for a diverse range of taxa (reviewed by Portnoy & Gold, 2012), including stone crabs (*Menippe* spp.; Bert 1986, but see Schneider-Broussard et al., 1998), sea robins (*Prionatus* spp.; McClure & McEachran, 1992), snapping shrimp (*Alpheus* spp.; McClure & Greenbaum, 1999; Matthews et al., 2002), and sheepshead (*Archosargus* spp.; Anderson et al., 2007). The exact location of species divergence appears to shift depending on the life-history of the species involved. For example, a hybrid zone suspected to reflect secondary contact between species of *Menippe* crabs lies along northern Florida and into Mobile Bay, Alabama (Bert, 1986). Other sister species or populations appear divided on either side of the Mississippi River, such as the lane snapper (*Lutjanus synagris*; Portnoy & Gold, 2012), Gulf squid (*Loligo plei*; Herke & Foltz, 2002), and the dusky pipefish (*Syngnathus scovelli*; Flanagan et al., 2016). A second issue arises in that the timing of divergence is not the same across these sister groups, indicating there were

---

<sup>1</sup> Hancock ZB, Hardin FO, Light JE (2019). Phylogeography of sand-burrowing amphipods (Haustoriidae) supports an ancient suture zone in the Gulf of Mexico. *Journal of Biogeography*, 46(11): 2532–2547.

likely multiple vicariant events in the Gulf of Mexico periodically throughout the history of the basin (Avise, 1992). However, generally low genetic distances between sister taxa indicate that these splits are all less than 5 million years old (Bert, 1986; Avise, 1992; Portnoy & Gold, 2012).



**Figure 1.1.** Biogeographic hypotheses in the Gulf of Mexico: a) Tennessee River drainage system which split sister taxa on either side of Mobile Bay, Alabama, approximately 2-4 Mya; b) Okefenokee Trough, which separated taxa along the Florida panhandle 1.8 Mya; c) recent cold-water pulses down the Mississippi River 9-16 Kya, which isolated taxa on either side of the Louisiana coastline; d) ancient Miocene sedimentation 5-10 Mya, which isolated taxa on either side of the Louisiana coastline; M (Mississippi River), R (Red River), T (Tennessee River). Map created with ESRI 2011. ArcGIS Desktop: Release 10. Redlands, CA: Environmental Systems Research Institute.

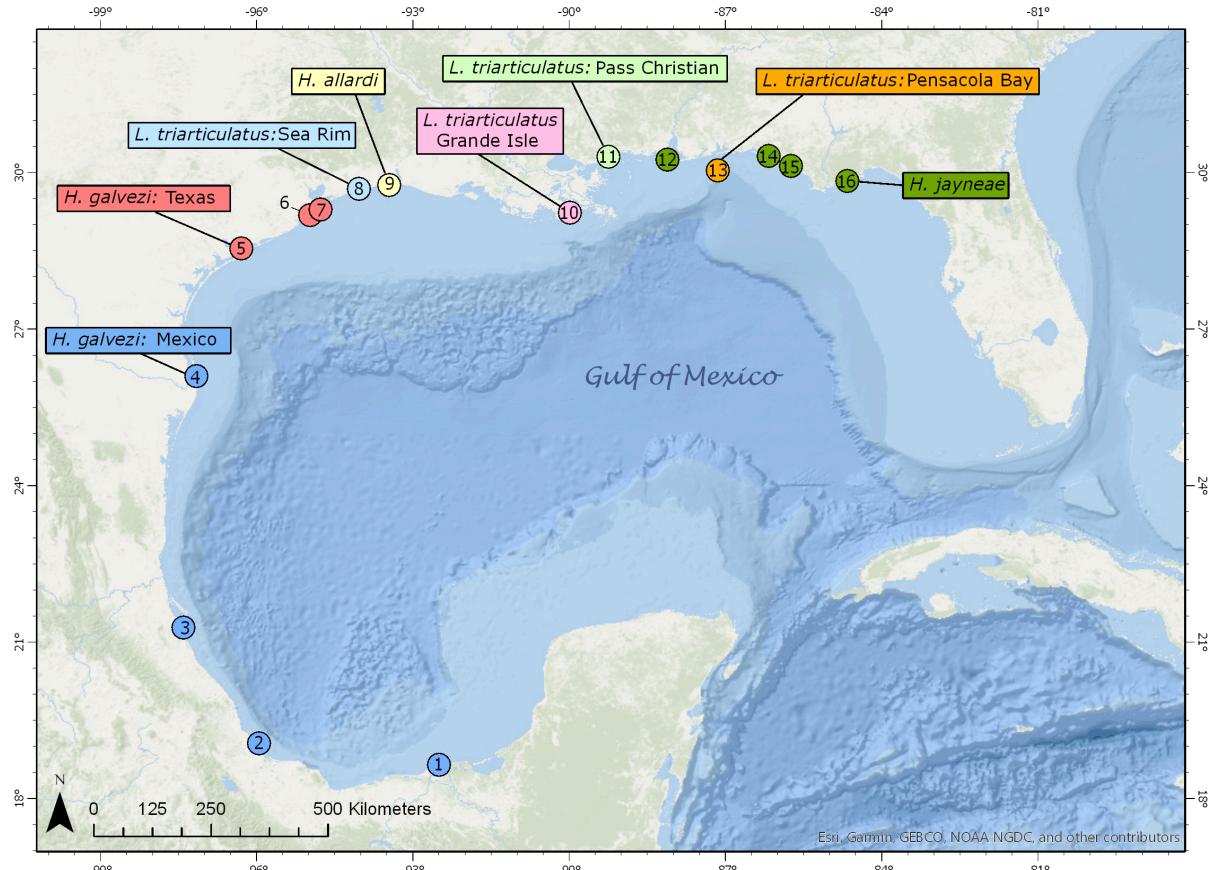
Apart from the Mississippi River, there appears to be no obvious geological or hydrological barriers today within the Gulf basin. However, major suture zones have been identified in the Gulf by past research. These are not visible today, but include: 1) the

“Tennessee River drainage system” that increased the freshwater outflow into Mobile Bay, Alabama, around 2.4 million years ago (Mya) (**Fig. 1.1a**; Simpson, 1900); 2) the Okefenokee Trough, a deep channel that separated a partially submerged Florida peninsula from the mainland during the Miocene (>5.33 Mya), and was closed by 1.75 Mya (**Fig. 1.1b**; Bert, 1986; McClure & Greenbaum 1999); and 3) cold-water pulses down the Mississippi River following the retreat of the Laurentide ice sheet, approximately 9–16 Kya (**Fig. 1.1c**; Aharon, 2003; Portnoy & Gold, 2012). Furthermore, prior to the formation of these suture zones the Gulf of Mexico had an ancient and complex geological history that may have affected species distributions in the distant past. For example, Galloway et al. (2011) examined a series of fluvial systems with major depositional impact on the Gulf across the Cenozoic (66 Mya–Present). During the Middle-Late Miocene, levels of sedimentation into the Gulf increased twofold due in part to the combined depositional contributions of an ancient Red River, the Mississippi River, and the Tennessee River, which may have resulted in a fourth ancient suture zone (**Fig. 1.1d**). This major sedimentation event would have created a coastal environment unsuitable for many nearshore marine organisms due both to lower salinities and heavier silt loads and would also create a pattern of east-west divergence comparable to suture zones described above. Each of these suture zones creates predictable patterns of species distributions and degrees of genetic divergence (**Fig. 1**). For example, sister taxa divided by the Tennessee drainage system are expected to be distributed on either side of Mobile Bay, Alabama, and divergence-time estimation should converge on ages of ~2 Mya (**Fig. 1.1a**). Similarly, for an ancient sedimentation event concentrated at the Mississippi River, sister taxa should be distributed on opposite sides of the river and divergence dating would indicate an older split during the Miocene (5–10 Mya; **Fig. 1.1d**). However, highly mobile organisms or those with planktonic larvae may obscure patterns

of historical vicariance by high rates of gene flow, modern hybridization, and introgression (e.g., Avise, 1992; Drovetski et al., 2015; Cahill et al., 2017). Examination of less mobile organisms may provide better insights into a potential ancient suture zone due to Miocene sedimentation down the Mississippi (**Fig. 1.1d**).

Amphipods (Crustacea: Amphipoda) are brooding crustaceans, a life-history strategy that has been proposed to limit their dispersal ability (Kinlan et al., 2005). Due to their limited mobility, amphipods have been used extensively to study biogeographic hypotheses, including inferring ancient glacial refugia (Witt & Herbert, 1999), land-mass uplift in the Carpathian paleo-archipelago (Copilaş-Ciocianu & Petrusek, 2017), and the impact of the Tsushima Straits in the Sea of Japan on population continuity (Takada et al., 2018). Haustoriidae, a family of burrowing amphipods that are morphologically specialized to a fossorial, filter-feeding lifestyle, live on fine sand beaches across the Gulf of Mexico. Haustoriids lack eye pigmentation and are poor swimmers that are believed to mate in the sediment, which may explain their rarity in the water column (Sameoto, 1969; Conlan, 1991; Hancock & Wicksten, 2018; but see Croker, 1967). Bousfield (1970) hypothesized that this family represented a recent rapid radiation into sandy-beach habitats, and that the group was likely evolutionarily young. Within the haustoriid genus *Haustorius*, there appears to be considerable species divergence across the Gulf of Mexico. Using morphological traits, Hancock & Wicksten (2018) reported an east-west divergence pattern of *Haustorius* species in the Gulf of Mexico. In the western Gulf, *Haustorius galvezi* Hancock & Wicksten, 2018 is the dominant intertidal amphipod and appears to range from southern Mexico to approximately Galveston Island, Texas, and has been reported as far north as Sea Rim State Park, Texas (Witmer, 2014; **Fig. 1.2**). In the eastern Gulf, *H. jayneae* Foster & LeCroy, 1991 ranges from Mississippi Sound to Carrabelle Beach, Florida (**Fig. 1.2**),

and is hypothesized to be the sister species to *H. galvezi* (Hancock & Wicksten, 2018). These species differ in overall body length, with *H. jayneae* being generally larger than *H. galvezi*. A larger body size may explain differences related to an inflated number of antenna articles, mandibular palp combs, and mandible accessory blades in *H. jayneae* compared to *H. galvezi*.



**Figure 1.2.** Distribution of amphipod species in the Gulf of Mexico including sampling sites for this study. 1) Tabasco-Campeche border (Mexico); 2) Antón Lizardo, Veracruz (Mexico); 3) Tamiahua, Veracruz (Mexico); 4) South Padre Island, Texas (U.S.A.); 5) Matagorda Island, Texas (U.S.A.); 6) Jamaica Beach, Texas (U.S.A.); 7) Galveston Island, Texas (U.S.A.); 8) Sea Rim State Park, Texas (U.S.A.); 9) Holly Beach, Louisiana (U.S.A.); 10) Grand Isle, Louisiana (U.S.A.); 11) Pass Christian, Mississippi (U.S.A.); 12) Dauphin Island, Alabama (U.S.A.); 13) Pensacola Bay, Florida (U.S.A.); 14) Grayton Beach State Park, Florida (U.S.A.); 15) St. Andrew's State Park, Florida (U.S.A.); 16) Carrabelle Beach, Florida (U.S.A.). Map created with ESRI 2011. ArcGIS Desktop: Release 10. Redlands, CA: Environmental Systems Research Institute.

Morphological differences not related to size include the telson cleft and the shape and spination of pereopod 7 article 4 (or merus; convention in Haustoriidae is to refer to article segments by number instead of name; see Bousfield, 1962; LeCroy, 2002; Hancock & Wicksten, 2018). A third species of *Haustorius*, *H. allardi* Hancock & Wicksten, 2018, is endemic to the Louisiana coastline where it occurs intertidally on muddy beaches.

A second genus of haustoriid amphipod, *Lepidactylus*, also occurs in the Gulf of Mexico. Despite co-occurring with *Haustorius* species on many beaches, the two genera are not known to hybridize and tend to vertically partition the beach (LeCroy, 2002; Shelton & Roberston, 1981) with *Haustorius* found in the lower intertidal and *Lepidactylus* in the high intertidal. *Lepidactylus* can be distinguished from *Haustorius* by epimeron 3 that is continuous with the urosome (it forms a shelf in *Haustorius*), as well as by the distinct pereopod 7 article 4 (merus) setation and shape (Roberston & Shelton, 1980). In the Gulf of Mexico, this genus is represented by a single species, *L. triarticulatus* Robertson & Shelton, 1980, which is known to range throughout the entire northern Gulf (Hancock & Wicksten, 2018; LeCroy, 2002). The only morphological differences in *L. triarticulatus* across the Gulf appear to be related to increased body size (number of spines and setae) in individuals sampled in the eastern Gulf, resulting in the recognition of a single, widely distributed species (Hancock & Wicksten, 2018; LeCroy, 2002). These morphological data indicate that despite occupying the same niche and having similar life-histories, the two amphipod genera may not share the same biogeographic history in the Gulf of Mexico. Alternatively, *L. triarticulatus* may represent a cryptic species complex, a common finding in amphipods (e.g., Vergilino et al., 2012; Fišer et al., 2015; Fišer et al., 2017).

This study expands on previous morphological work by analyzing four genes (two mitochondrial and two nuclear) to examine population divergence of *Haustorius* and

*Lepidactylus* across the Gulf of Mexico, and to determine if relatively contemporaneous or more ancient suture zones best explain modern species distributions. Since sister species of *Haustorius* are distributed on opposite sides of the Mississippi River and are suspected poor dispersers, I hypothesize that their genetic divergence would correspond with an ancient sedimentation event during the Miocene (5–10 Mya; **Fig. 1.1d**; Galloway et al., 2011) and thus an ancient suture zone in the Gulf of Mexico. Alternatively, these species may have been separated much more recently by cold-water pulses at the end of the last ice age (**Fig. 1.1c**; Portnoy & Gold, 2012), which would be supported by low levels of genetic divergence. Additionally, I hypothesize that *Lepidactylus triarticulatus* is composed of at least two cryptic lineages, with sister lineages minimally separated by the Mississippi River. Using genetic data to examine divergence patterns and timing of these two amphipod genera will allow us to determine if an ancient suture zone affects the distribution of dispersal-limited species in the Gulf of Mexico.

## Methods

### *Field collections*

Sixteen sites across the Gulf of Mexico were sampled, ranging from the border of Tabasco-Campeche in southern Mexico to Carrabelle Beach, Florida (**Fig. 1.2, Table S1.1**). Individual amphipods were collected with a shovel and 435 $\mu$ m sieve plate, removed with forceps, and preserved in 95% EtOH. Individuals were morphologically identified in the lab under a dissecting microscope and stored at -20°C prior to DNA extraction. Using the same methodologies, *Haustorius canadensis* Bousfield, 1962 was also collected from Cape May, New Jersey, to serve as an outgroup taxon for phylogenetic analyses of *Haustorius*. Representatives of

each identified lineage have been deposited at the Biodiversity Research and Teaching Collections (BRTC) at Texas A&M University, with additional specimens deposited at the Smithsonian Museum of Natural History (USNM; **Table S1.1**). Additionally, *H. arenarius* Slabber, 1769 specimens were loaned from the Royal Belgian Institute of Natural Sciences to serve as an additional outgroup (catalog #INV.138073).

#### *DNA extraction and alignment*

Whole genomic DNA was extracted from 1–2 pereopods from larger individuals or the entire organism for smaller individuals using an EZNA Tissue DNA Kit (Omega Bio-tek Inc., Norcross, Georgia) following the manufacturer’s protocol. Two mitochondrial genes, cytochrome oxidase I (*COI*) and 16S ribosomal RNA (16S), and two nuclear genes, 18S ribosomal RNA (18S) and 28S ribosomal RNA (28S), were amplified via polymerase chain reaction (PCR) using primer sets and conditions in **Table S1.2**. Amplicons were verified using gel electrophoresis and ExoSAP-IT (Affymetrix, Inc., Santa Clara, CA, U.S.A.) was used to purify positive PCR products. Forward and reverse strand sequencing was performed at DNA Analysis Facility on Science Hill at Yale University. Sequences were manually edited in Sequencher v.4.10.1 (Gene Codes Corp., Ann Arbor, MI, U.S.A.) and alignments were performed in MAFFT v.7 (Katoh & Standley, 2013). For *COI* (the only protein-coding gene used in this study), alignments were translated into amino acids in Mesquite v.3.5 (Maddison & Maddison, 2018) to check for premature stop codons.

#### *Phylogenetic methods*

Phylogenetic inference was first conducted for each gene separately, each of which was collapsed into unique haplotypes using FaBox v.1.41 (Villesen, 2007). Additional outgroup taxa were identified using a BLAST search (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>; **Table S1.2**). PartitionFinder2 (Lanfear et al., 2016) was used to identify the best partitioning scheme and substitution model by codon position for the protein-coding gene (*COI*) using Bayesian Information Criterion (BIC). For RNA genes, jModelTest2 (Darriba, 2012) was used to find the best model of DNA evolution using a sample-size corrected Akaike Information Criterion (AICc). Models not implemented in MrBayes were replaced with the next most overparameterized model (Lecocq et al., 2013; Huelsenbeck & Rannala, 2004). Maximum-likelihood (ML) inference was conducted in RAxML v.8.2.10 (Stamatakis, 2014) with 1,000 multiparametric bootstrap replicates in the CIPRES portal (Miller et al., 2010). Bayesian inference was performed using MrBayes v.3.2.6 (Huelsenbeck & Ronquist, 2001) with four independent Markov chain Monte Carlo (MCMC) chains of 100 million generations sampling every 5,000. Convergence was checked in TRACER v.1.7 (Rambaut et al., 2018) by visually examining trace graphs and estimated sample sizes (ESS), which was assumed if all parameter estimates were >200. All trees produced were summarized in Dendropy (Sukumaran & Holder, 2010) using the command *sumtrees* to produce a 50% majority-rule consensus tree. Consensus trees were visualized in FigTree v.1.4.3 (Rambaut, 2012). If no significant conflict was observed among the individual gene trees, a total evidence tree was inferred by concatenation with each gene as an independent partition in MrBayes as described above.

#### *Species delimitation methods*

To identify operational taxonomic units (OTUs), interpreted as putative species, four separate delimitation methods were used, and all were applied solely to the *COI* gene except for the Bayesian implementation of the Poisson tree processes method (bPTP), which also used the total evidence tree. The first method was the general mixed Yule coalescent (GMYC) model, which identifies the threshold between lineages following a Yule process with log-normal distributed lineages through time and those lineages that follow an intrapopulation coalescent model (Fujisawa et al., 2016). GMYC was implemented in the R platform using the package SPLITS from the CRAN repository and was performed on the ultrametric tree produced from BEAST2 as described above. A second delimitation model, bPTP, was implemented on the web-server [species.h-its.org](http://species.h-its.org) (Zhang et al., 2013). Since this method does not require an ultrametric tree, it was performed on both *COI* and the total evidence tree produced in MrBayes. Lastly, two distance-based approaches were used: 1) Lefébure et al. (2006) *a priori* designations of 0.16 substitutions per site (subs./site) as a universal threshold for delimiting crustacean species for *COI*, and 2) the Automatic Barcode Gap Discovery (ABGD; Puillandre et al., 2011) implemented on the web server <http://wwwabi.snv.jussieu.fr/public/abgd/abgdweb.html>. In the ABGD, Pmin was set to 0.001, Pmax to 0.1, X = 1.0, and the distance matrix was estimated using the Kimura-2 parameter (K2P) model (Kimura 1980; following Hurtado et al., 2010 and Copilaş-Ciocianu & Petrusek, 2017).

*Divergence-time and species-tree estimation*

Divergence-time estimation was performed on *COI* using the widely-used molecular clock of 1.4–2.6% divergence per My, or 0.007–0.013 subs./site per My (Knowlton & Weigt, 1998; Takada et al., 2018). This molecular clock was used because there are no known noncontentious amphipod fossils by which to calibrate the tree. Divergence-time estimation was conducted in BEAST2 using a Birth-Death prior and a relaxed log-normal clock rate with a mean rate of 0.01 and variance of 0.006 sites/My following Takada et al. (2018), and an MCMC chain of 10 million sampling every 1,000. A shared GTR + G model was used instead of a partitioned approach because the latter analysis struggled to reach convergence, and the topology was not different between runs or from the partitioned MrBayes *COI* tree (**Fig. S1.1**) Convergence was verified in Tracer by visually inspecting trace graphs and ESS values. Sampled trees were summarized in TreeAnnotator and visualized in FigTree.

To account for possible gene tree incongruence, ancestral polymorphism, and to jointly estimate ancestral population sizes using all four genes, I inferred a species-tree using \*BEAST (Heled & Drummond, 2010). I applied a strict molecular clock to *COI* with a mean rate of 0.01 (0.007–0.024); for the other three loci, relaxed log-normal clocks were applied. The population function parameter in the multispecies coalescent was set to “linear\_with\_constant\_root” to allow population sizes to fluctuate across branches. Since this analysis required an *a priori* designation of individuals to “species”, I grouped individuals into OTUs identified using the tree-based methods described above to avoid biases in population size inference due to significant population structure. I performed four independent MCMC runs of 100 million steps sampling every 5,000, which provided ESS values of >700 for all estimated parameters.

#### *Population genetic methods*

Population genetic analyses were estimated on all OTUs identified using the tree-based delimitation methods described above. Distance-based delimitation methods result in significant population structure within OTUs (see Results), which can bias methods for estimating effective population sizes ( $N_e$ ) (Heller et al., 2013); thus, OTUs determined using these methods were not assessed at the population level. All population genetic analyses were conducted on the *COI* gene as that marker had the greatest sampling (**Table S1.1**) and variation. Population parameters including number of segregating sites ( $S$ ), nucleotide diversity ( $\pi$ ), and haplotype diversity ( $h$ ) were estimated in DnaSP v.6 (Rosas et al., 2018).

Additional analyses were conducted only on *Haustorius* OTUs due to low numbers of sampled *Lepidactylus* OTUs (**Table S1.1**). The probability of including the most recent common ancestor (MRCA) in a sample of  $n$  individuals is  $(n - 1) / (n + 1)$ , which is derived from Kingman's coalescent (Kingman, 1982) by Saunders et al. (1984). The *Haustorius* OTUs all had at least 19 individuals, which provides a 90% probability of including the MRCA. Pairwise Slatkin's linearized  $F_{ST}$  (Slatkin, 1991; 1995) was estimated in Arlequin v.3.5 (Excoffier & Lischer, 2010), which calculates  $F_{ST}$  as  $(t_1 - t_0) / t_1$ , where  $t_1$  is the mean coalescence time between two genes from different populations, and  $t_0$  is the mean coalescence time between two genes drawn from the same population.  $F_{ST}$  significance was determined by a permutation test with 110 replicates. A pairwise distance matrix was generated in MEGA7 (Kumar et al., 2016) using a K2P model. TCS haplotype networks for *COI* were constructed in PopART (Leigh & Bryant, 2015). Isolation-by-distance (IBD) was tested in FSTAT v.1.2 (Goudet, 1995) using a Mantel test conducted by comparing a pairwise matrix of  $F_{ST} / (1 - F_{ST})$  (after Rousset, 1997) against a pairwise geographical distance matrix. Significance of the  $\beta$  coefficient was determined

by 2,000 randomizations. To visualize population structure, a TCS *COI* haplotype network was constructed using PopART (Leigh & Bryant, 2015).

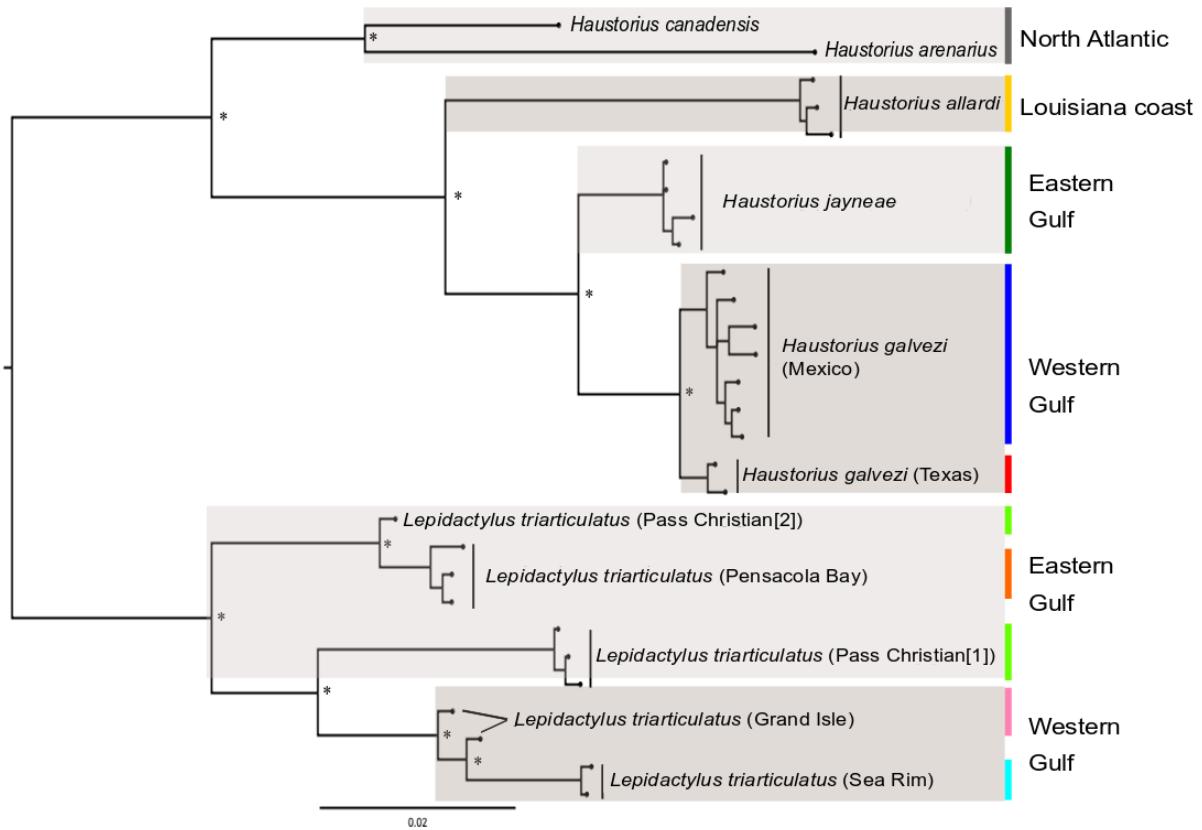
To estimate changes in effective population sizes ( $N_e$ ) over time, three separate methods were used: 1) Bayesian coalescent skylines were constructed for each identified OTU using BEAST2 (Bouckaert et al., 2014) with a Bayesian coalescent skyline tree prior and an MCMC chain of 10 million. This was performed under a strict molecular clock rate of 0.01 subs./My. Preliminary work indicated that these analyses struggled to reach convergence (even after increasing the MCMC chain to >500 million), so a simpler substitution model was specified (TN93; Tamura & Nei, 1993). ESS values >1,000 for all parameters were then obtained for all runs. 2) Mismatch distribution plots were generated in Arlequin, and Harpending's raggedness index (HRI; Harpending, 1994), which measures the "smoothness" of the distribution, was calculated with 100 bootstrap replicates under a sudden population expansion model. The fit of the sudden expansion model was tested by the sum of squared deviations (SSD). 3) Neutrality tests were performed in Arlequin, including Tajima's *D* (Tajima, 1989) and Fu's *F<sub>s</sub>* (Fu, 1997) with significance determined by 1,000 simulations.

## Results

### *Field work*

All sites from Galveston Island, Texas, U.S.A., to Tabasco, Mexico were dominated intertidally by *Haustorius galvezi*, often in numbers >100 per sample. Similar abundances were seen at Sea Rim State Park, Texas, and Holly Beach and Grand Isle in Louisiana, with *H. allardi* collected at Holly Beach. *Haustorius jayneae* was sampled from Dauphin Island, Alabama, and

Grayton Beach, St. Andrews State Park, and Carrabelle Beach, all in Florida. *Lepidactylus triarticulatus* was sampled at Sea Rim and Grand Isle and was the only intertidal haustoriid found at Pass Christian, Mississippi, and Pensacola Bay, Florida. Only Carrabelle Beach had population abundances comparable to the western Gulf populations for both species. The other locations presented difficulties collecting enough specimens; most of these locations were represented by five or fewer individuals (**Table S1.1**).

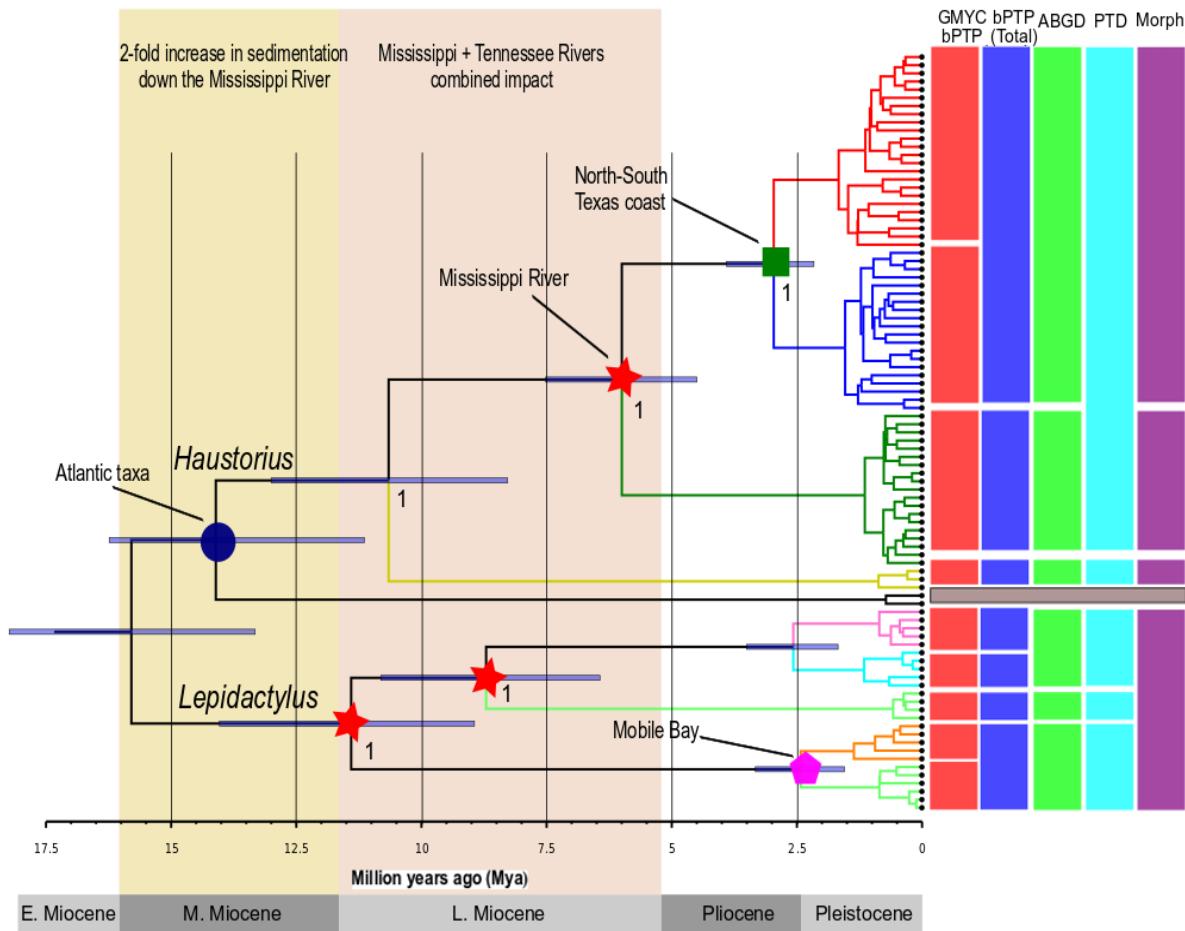


**Figure 1.3.** Total evidence tree produced in MrBayes with each gene independently partitioned. Asterisks on nodes represent >0.95 posterior support. Distributions are indicated to the right of each lineage. Outgroups only selected for individual genes were removed from the concatenated dataset and genera were rooted by one another. Bar color for each clade represent OTU location in **Fig 1.8**.

### Phylogenetic results

After trimming and alignment, the two mitochondrial genes, *16S* and *COI*, were 458 bp and 392 bp, respectively; the nuclear genes produced alignments of 2,050 bp for *18S* and 1,143 bp for *28S*. Thus, the concatenated dataset consisted of 4,043 bp. Details on the number of individuals sequenced for each gene can be found in **Table S1.1**.

All phylogenetic inference methods of all molecular markers revealed the nominal Gulf *Haustorius* species and *Lepidactylus triarticulatus* to be monophyletic (**Fig. 1.3–1.4; Fig. S1.1–S1.4**). Within *Haustorius*, *H. galvezi* and *H. jayneae* were supported as sibling species with *H. allardi* sister to them. The Atlantic haustoriids (*H. arenarius* and *H. canadensis*) grouped together outside the Gulf-specific taxa according to *18S* (**Fig. S1.4**) causing *Haustorius* to be paraphyletic, albeit with low support (bootstrap support = 0.60), but showed monophyly at all other markers and in the concatenated tree (**Fig. 1.3, Fig. S1.1–S1.3**). Relationships and support for within-species Gulf clades sometimes differed with genetic marker and inference methods. For *COI*, Bayesian and Maximum Likelihood methods found differing and generally weak levels of support for reciprocal monophyly between a northern Texas *H. galvezi* clade and southern Texas and Mexico *H. galvezi* clade (**Fig. S1.1**). In contrast, Bayesian and Maximum Likelihood inference found strong (MrBayes PP = 0.98; **Fig. S1.2**) and moderate (RAxML bootstrap = 0.64; **Fig. S1.2**) support for these clades using the *28S* marker. Analyses using *18S* failed to distinguish these subgroups (likely as a result of low variation: **Fig. S1.4**) and this subdivision was not tested with *16S* due to failure of *H. galvezi* (Texas) specimens to amplify (**Fig. S1.3; Table S1.1**). The concatenation approach found high support (PP = 0.99) for monophyly between *H. galvezi* (Texas) and *H. galvezi* (Mexico) (**Fig. 1.3**).



**Figure 1.4.** Divergence-time tree using the mitochondrial *COI* gene produced in BEAST2 with Birth-Death prior and relaxed log-normal molecular clock. Blue bars at nodes represented 95% HPD; numbers under nodes are posterior support (only values > 0.95 are shown). Clades are colored by OTUs from **Fig 1.2**. Colored boxes on the right are inferred species from each delimitation method, with a comparison to morphologically defined lineages (Morph) as determined by Hancock & Wicksten (2018). Tree-based: GMYC = General mixed Yule coalescent; bPTP = Bayesian Poisson tree processes. Distance-based: ABGD = Automatic Barcode Gap Discovery; PTD = Patristic distance (defined by Lefébure et al., 2006).

Solid gray line is *H. canadensis*, which was not considered in the delimitation analysis. Symbols at nodes indicate major splits and specific geographic breaks, and ancient suture zones are indicated in yellow and tan shading.

For *Lepidactylus triarticulatus* sample sites, high divergence and monophyly were recovered across most sites, genes, and methods (**Fig. 1.4**, **Fig. S1.1**). Two distinct clades were recovered that were diverged by >0.16 subs./site at *COI*, and these can superficially be separated into a western (Sea Rim State Park, Texas, and Grand Isle, Louisiana) and eastern (Pass Christian, Mississippi, and Pensacola Bay, Florida) clade (**Figs. 1.4–1.5**). Notably, two distantly

related lineages were recovered at Pass Christian, Mississippi (**Fig. 1.4; Table S1.1**). These lineages were separated by 0.1858 subs./site, and lineage 1 (*L. triarticulatus*, Pass Christian[1]) grouped with, yet was highly diverged from (>0.15 subs./site), the western clade of *L. triarticulatus*. Gene trees for *16S*, *28S*, and *18S* all found high support for an *L. triarticulatus* (Sea Rim) and *L. triarticulatus* (Grand Isle) clade, *L. triarticulatus* (Pensacola Bay) and *L. triarticulatus* (Pass Christian[2]) clade, and an *L. triarticulatus* (Pass Christian[1]) clade, but lacked variation to distinguish site-specific lineages as *COI* did (**Fig. S1.1–S1.3**). The concatenation approach largely agreed with the individual gene trees except for the placement of the *L. triarticulatus* (Sea Rim) and *L. triarticulatus* (Grand Isle) clade, which was not found to be monophyletic with high support (**Figs. 1.9, S1.1–S1.4**).

#### *Species delimitation results*

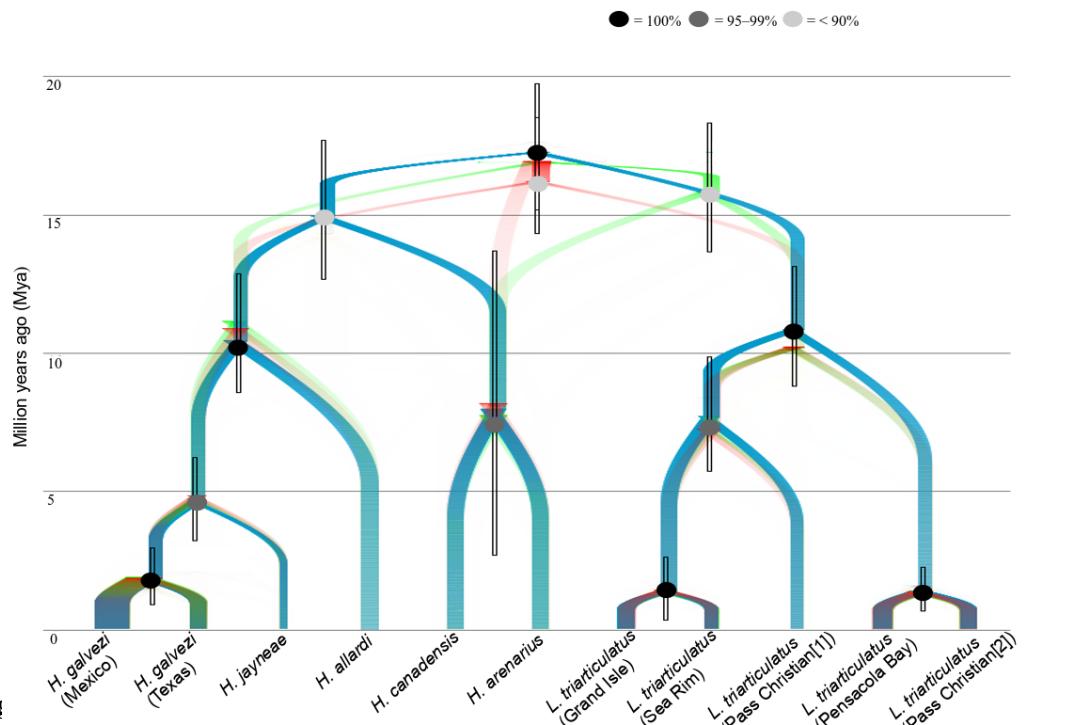
Species delimitation methods supported a minimum of five OTUs across both amphipod genera. Both the GMYC and bPTP methodologies identified nine OTUs using *COI* only, corresponding to each monophyletic clade produced by the ultrametric BEAST tree: four species within *Haustorius* (two clades within *H. galvezi* corresponding to Texas and Mexico, *H. jayneae*, and *H. allardi*) and five species within *Lepidactylus* corresponding to each collecting locality and the two lineages at Pass Christian (Sea Rim, Grand Isle, Pass Christian[1], Pass Christian[2], and Pensacola Bay; **Fig. 1.4**). However, the bPTP delimitation method differed when using the total evidence tree as input. Since the total evidence tree (**Fig. 1.3**) did not find support for monophyly of *Lepidactylus* from Grand Isle, the bPTP analysis supported each individual as independent species (PP = 0.71, 0.75, respectively) resulting in a total of eight putative species as opposed to

seven. However, I adopted a conservative approach and collapsed the two inferred species from Grand Isle into one as supported by the species-tree analysis (**Fig. 1.5**). Gaps were found in the ABGD at *COI* in both *Haustorius* and among *Lepidactylus* sites, corresponding to a maximal intraspecific distance for both genera of 3.6%. At this threshold, three distinct groups were identified in each genus; for *Haustorius*, this corresponded to the three nominal species, *H. allardi*, *H. galvezi*, and *H. jayneae*; for *L. triarticulatus*, this corresponded to lineage 1 from Pass Christian, specimens from Sea Rim and Grand Isle, and specimens from Pensacola Bay and lineage 2 from Pass Christian (**Fig. 1.4**). Using the distance threshold suggested by Lefébure et al. (2006), five putative species corresponding to *H. jayneae* and *H. galvezi*, *H. allardi*, and the same three *Lepidactylus* lineages identified by the ABGD method were delimited (**Fig. 1.4**).

#### *Divergence-time and species-tree estimation results*

For divergence-time estimation, the coefficient of rate variation was checked first to determine if the tree followed clock-like evolution. The coefficient was 0.077, indicating little rate variation across the tree, justifying the use of a strict clock for all subsequent analyses (a coefficient of rate variation <0.1 has been considered low enough to justify a strict molecular clock; Drummond & Bouckaert, 2015; Verheyen et al., 2017). Within *Haustorius*, the Gulf clade diverged from the Atlantic species around 14.2 Mya (95% HPD 11.5–16.8 Mya; **Fig. 1.3–1.4**). The divergence between the northern Texas and southern Texas and Mexico *H. galvezi* clades was estimated between 2.2–4 Mya (mean of 3.1 Mya), and the split between these *H. galvezi* clades and *H. jayneae* occurred ca. 6.2 Mya (4.6–7.8 Mya; **Fig. 1.3–1.4**). *Haustorius allardi* diverged from the other *Haustorius* Gulf taxa between 8.6–13.4 Mya (mean of 11 Mya). Within

*Lepidactylus*, a clade consisting of *L. triarticulatus* (Sea Rim) and *L. triarticulatus* (Grand Isle) were estimated to have diverged from *L. triarticulatus* (Pass Christian[1]) around 8.9 Mya (6.6–11.2 Mya), and the divergence between the western and eastern clades was estimated at 9.2–14.5 Mya (mean of 11.9 Mya).



**Fig. 1.5** Relaxed log-normal clock for the other three loci. Each of the three overall trees represents the three most common topologies in descending order from blue, green, and red. Width of the branches represents inferred effective population sizes. Clear bars on the nodes are 95% HPD and filled circles are posterior probabilities.

The species-tree analysis largely supported the divergence-time estimation above for the deepest divergences but lowered the most recent divergences by as much as 50% at the lower end of the 95% HPD (**Fig. 1.5**). The mean of the split between *H. jayneae* and *H. galvezi* was reduced to 4.6 Mya and the west-east split in the *L. triarticulatus* species complex was 5.4–13.2 Mya. The estimated divergence time for the western clades of *H. galvezi* was 1.7 Mya (0.79–2.9

Mya, 95% HPD). For *L. triarticulatus*, the sister group from Grand Isle and Sea Rim was estimated to have diverged 1.4 Mya (0.25–2.6 Mya) whereas the sister group from Pensacola Bay and Pass Christian[2] diverged 1.3 Mya (0.5–2.2 Mya). The coefficient of rate variation for *16S*, *18S*, and *28S* all exceeded 0.3, indicating a relaxed clock is the most appropriate for these loci. Most ancestral branches were inferred to have relatively constant population sizes with most expansions (*H. galvezi*, Mexico) or contractions (*H. jayneae*) occurring within the last 5 My. The three most common branching orders are displayed in **Fig. 1.5**, with the blue tree being the most common followed by green and red. The only clade with major incongruence is the Atlantic *Haustorius* group—85% of trees find *Haustorius* to be monophyletic, while the remaining trees are split between Atlantic *Haustorius* grouping with Gulf *Lepidactylus* or with Gulf *Haustorius* and *Lepidactylus* as sister clades with the Atlantic *Haustorius* as an outgroup (green and red trees, respectively; **Fig. 1.5**).

#### *Population genetic results*

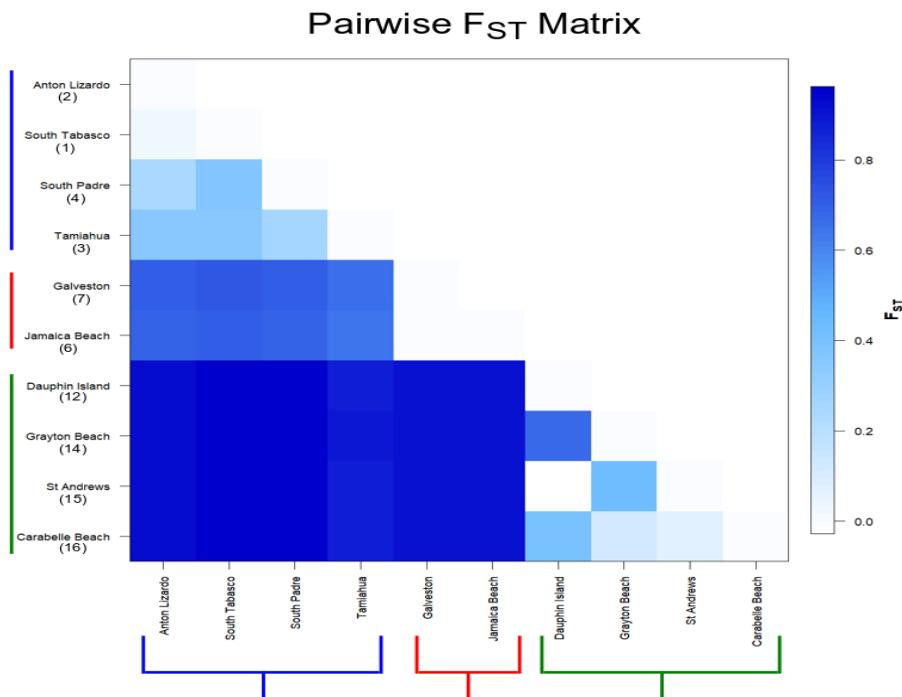
Population genetic summary statistics of *COI* revealed high haplotype diversity within each OTU identified using the tree-based species delimitation methods except *L. triarticulatus* (Pass Christian[1]) where all sequences were identical (**Table 1.1**). Nucleotide diversity ( $\pi$ ) was highest in *H. galvezi* (Texas) and *H. galvezi* (Mexico) among the *Haustorius* lineages, and comparable to *Lepidactylus* from Pensacola Bay; *H. jayneae* and the other *Lepidactylus* lineages had an order of magnitude less nucleotide diversity (**Table 1.1**). Pairwise  $F_{ST}$  revealed significant population structure among most sampled sites, including within identified OTUs (**Fig. 1.4**).  $F_{ST}$  was much higher between the *H. galvezi* OTUs than within, consistent with the phylogenetic

results and despite the close proximity of sampled sites along the Texas coastline. However, the Mantel test revealed that distance is a significant explanatory variable ( $P = 0.0160$ ;  $r^2 = 0.6149$ ;  $\beta = 0.0012$ ) of  $F_{ST}$ . The *COI* haplotype network revealed no shared haplotypes between *H. galvezi* in northern Texas and *H. galvezi* from southern Texas and Mexico, and the closest haplotypes were separated by no less than five inferred mutations (Fig. S1.5).

OTU	# Individuals	<i>S</i>	$\pi$ (SD)	<i>h</i> (SD)	Tajima's <i>D</i>	Fu's <i>F<sub>s</sub></i>	SSD	HRI
<i>H. galvezi</i> (Texas)	20	22	0.0116 (0.003)	0.9842 (0.000)	-1.2807	<b>-19.6246***</b>	<b>0.0257**</b> *	0.0443
<i>H. galvezi</i> (Mexico)	21	33	0.0160 (0.003)	1.0000 (0.000)	-1.4055	<b>-22.4360***</b>	0.0045	0.0107
<i>H. jayneae</i>	19	6	0.0040 (0.002)	0.7836 (0.004)	-0.3038	<b>-4.4631**</b>	0.0047	0.0387
LT (Sea Rim)	5	4	0.0046 (0.002)	0.9000 (0.026)	-	-	-	-
LT (Grand Isle)	5	5	0.0057 (0.003)	0.9000 (0.026)	-	-	-	-
LT (Pass Christian[1])	4	0	0	0	-	-	-	-
LT (Pass Christian[2])	6	5	0.0057 (0.003)	0.8667 (0.017)	-	-	-	-
LT (Pensacola Bay)	5	11	0.0152 (0.005)	1.0000 (0.016)	-	-	-	-

**Table 1.1.** Summary statistics. OTUs are pooled individuals from different sites identified via tree-based delimitation methods. *S* is the number of segregating sites,  $\pi$  is nucleotide diversity, *h* is haplotype diversity, SSD is the sum of squared deviations from the sudden expansion model, and HRI is Harpending's raggedness index. Significance is determined by bolded numbers; \* =  $P < 0.01$ , \*\* =  $P < 0.001$ , \*\*\* =  $P < 0.0001$ . LT = *Lepidactylus triarticulatus*.

Both *H. jayneae* and *H. galvezi* (Mexico) showed evidence of a recent population expansion (unimodal mismatch distribution, significant Fu's  $F_s$ , and the Bayesian skyline reconstruction; **Table 1.1**; **Fig. 1.6**). However, in *H. jayneae* the Bayesian skyline plot indicated a shallow coalescent tree, which is consistent with a population bottleneck, as was inferred from the species-tree analysis above (**Fig. 1.6–1.7**). For *H. galvezi* from northern Texas, both the Bayesian skyline and mismatch distributions supported a population expansion; however, a significant deviation from the sudden expansion model was detected by SSD ( $P < 0.0001$ ), but not Harpending's raggedness index (HRI). Additionally, Tajima's  $D$  was not significant ( $P = 0.096$ ).



**Figure 1.6.** Pairwise  $F_{ST}$  matrix. Numbers under localities refer to sample sites in **Fig. 1.2**.

## Discussion

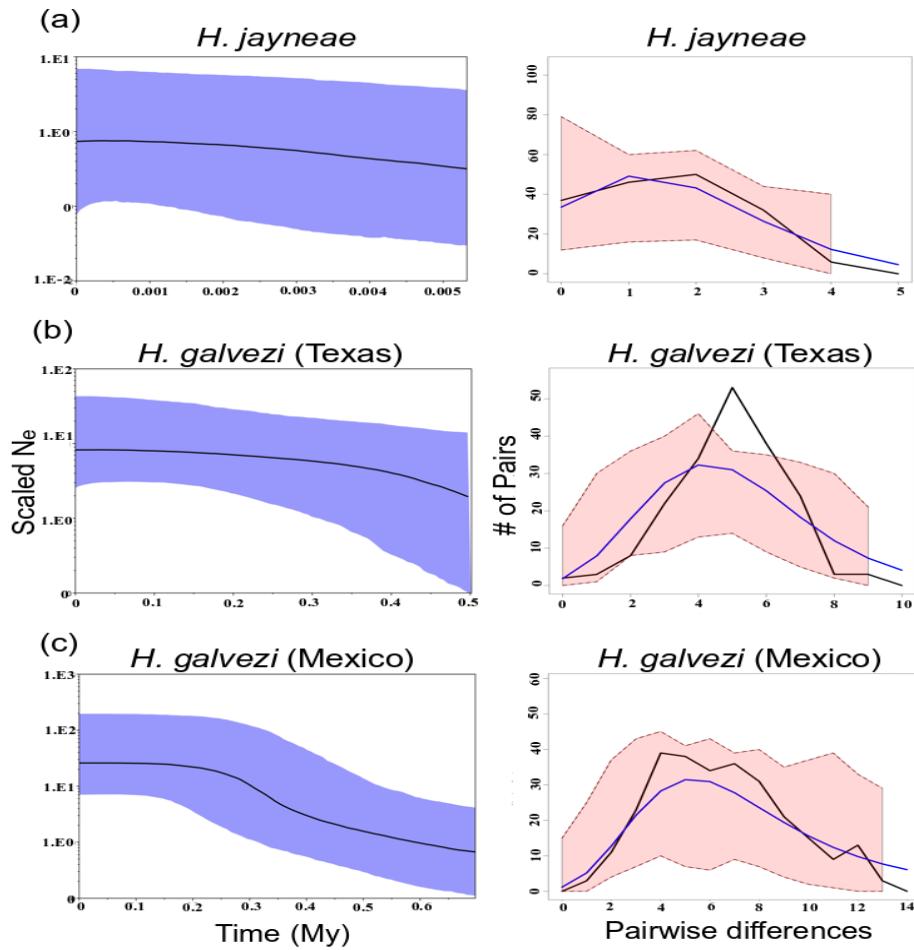
Analyses performed in this study suggest the presence of multiple cryptic lineages of *Haustorius* and *Lepidactylus* amphipods in the Gulf of Mexico. The Mississippi River is the major dividing feature of modern species distributions, creating eastern and western clades within both genera. All of the observed divergences between eastern and western clades of *Haustorius* and *Lepidactylus* are >4.5 My (**Fig. 1.3–1.4**), much too early to reflect the impact of glacial cooling during the Pleistocene (Portnoy & Gold, 2012), supporting an ancient suture zone in the Gulf of Mexico.

#### *How many species?*

Haustoriid amphipods have highly divergent body plans compared to other amphipods consisting of broadly expanded pereopods 5–7, fossorial gnathopods 1–2, powerful pleopods for creating a ventral food current, and spiny uropods modified for digging, all specializations for a burrowing lifestyle in sandy substrate. In addition, Bousfield (1970) proposed that the family represented a recent adaptive radiation, indicating that the group may be evolutionarily young. These two factors coupled together may explain the limited morphological divergence observed in the Gulf haustoriids, and therefore the widespread occurrence of cryptic diversity. A similar pattern has been seen in another amphipod group (the *Hyaellela azteca* species complex) where strong ecological constraints have been implicated as drivers of morphological stasis (Wellborn & Cothran, 2003; Witt et al., 2003). However, there remains the possibility of fine-scale microstructural, hormonal, or behavioral differences that could be used to delineate apparently morphologically cryptic species. Additional research would be necessary to determine if these or other traits may result in differentiation among haustoriid lineages.

A minimum of five and a maximum of nine “species” were identified with molecular species delimitation analyses. Incongruence among these methods on the number of OTUs likely is the result of differing underlying assumptions of each model. The tree-based methods found sufficient support in the ultrametric tree to identify a threshold that encompassed each unique clade (**Fig. 1.4**). However, since the total evidence tree did not support a monophyletic *Lepidactylus* from Grand Isle, the bPTP approach on this tree was skewed in favor of inferring more species (individuals from the same locality were determined as putative species). The topology of the total evidence tree may be driven by the two nuclear loci (*18S* and *28S*) that do not have sufficient variation to differentiate *Lepidactylus* from Grand Isle and Sea Rim overpowering the information at the mitochondrial loci (**Figs. S1.1–S1.4**). Thus, I collapsed these two inferred *Lepidactylus* species into one, resulting in a total of seven putative species. The distance-based methods found contradicting results depending on the *a priori* interspecific divergence ranges (six lineages for ABGD and five lineages for the distance method of Lefébure et al. 2006). The distance method of Lefébure et al. (2006) uses a *COI* substitution rate of 0.16 subs./site to delineate crustacean species, but this designation is likely overly conservative for amphipods. For example, Lagrue et al. (2014) examined mate discrimination in the freshwater amphipods *Gammarus pulex* and *G. fossarum* and found that individuals diverged by >0.04 subs./site at *COI* no longer mated randomly, indicating that mate discrimination occurs around that distance in at least this particular amphipod genus. If I were to use this substitution rate of a maximum intraspecific distance of 0.04 subs./site, the distance method of Lefébure et al. (2006) would produce six OTUs, which is in agreement with the ABGD method. Overall, the number of Gulf of Mexico *Haustorius* and *Lepidactylus* OTUs ranges from 6–8, which is 2–4 more species than are morphologically recognized in the Gulf of Mexico. Future work on haustoriid

amphipods will require a more integrative taxonomic approach (e.g., Padial et al., 2010; Pavón-Vázquez et al., 2018) including genetic data, environmental and ecological variables, fine-scale microstructural differences, and mate preference to adequately designate cryptic lineages (or OTUs) as species.



**Figure 1.7.** Effective population sizes ( $N_e$ ) over time inferred from *COI*. (Left) Bayesian coalescent skyline plots produced in BEAST2; filled blue boundaries are the 95% CI. (Right) Mismatch distributions produced in Arlequin. The black line is observed mismatches, blue line is simulated expectation based on a population expansion model, and shaded red area is the 95% CI. a) *Haustorius jayneae*, eastern Gulf of Mexico; b) *H. galvezi*, northern Texas, western Gulf; c) *H. galvezi*, southern Texas and Mexico, western Gulf.

### *Miocene sedimentation and divergence-dating*

The distribution of sister species of *Haustorius* and *Lepidactylus* on either side of the Mississippi River indicates that this fluvial axis likely played a major role in vicariant speciation within these genera. While the modern Louisiana coastline is relatively young (~7 Kya; Roberts, 1997), the Mississippi River became the dominant supplier of sediment into the Gulf of Mexico by the Late Miocene (~20 Mya; Galloway et al., 2011). By the Middle Miocene (~12.5 Mya), sedimentation levels had increased by twofold from levels on the Oligocene-Miocene boundary. This dramatic increase may have caused the coastline straddling the river's mouth to accumulate silt and mud, creating an inhospitable habitat for haustoriids (**Fig. 1.1d**). Given the deep divergences of sister clades of both *Haustorius* (2.9–6.2 Mya) and *Lepidactylus* (5.4–13.2 Mya), these splits were likely not caused by cold-water pulses down the Mississippi at the end of the last ice age (Portnoy & Gold, 2012), but rather increased sedimentation during the Middle and Late Miocene. Two additional fluvial systems likely contributed to this: 1) an ancient Red River, which flowed into the Gulf of Mexico directly, unlike its modern counterpart that has been captured by the Mississippi, and 2) the Tennessee River. Each of these fluvial axes was acting during the Miocene and into the Pliocene, and much of the sedimentation was concentrated along the Louisiana coastline (**Fig. 1.1d**; Galloway et al., 2011). A portion of the drainage system of the Tennessee River would eventually break-off and flow into Mobile Bay, but this deviation did not occur until the Pliocene (**Fig. 1.1a**; Galloway et al., 2011).

Discrepancy between the two dating methods demonstrates the importance of considering the impacts of ancestral population structure and deep coalescence. The molecular clock method

applied to *COI* is strictly dating gene divergence and likely inflates population divergence times as it fails to account for ancestral polymorphism (Edwards & Beerli, 2000). This is particularly striking in the most recently diverged clades, in which the median divergence time is between 25–50% lower in the multispecies coalescent (MSC) method compared to the molecular clock-only method. Despite this, the medians of both methods overlap in the 95% HPD, albeit in the lower bounds for the MSC. The MSC method pushes the east-west divergence closer to the Late Miocene–Pliocene boundary, but still within the hypothesized period of impact.

Miocene sedimentation would likely have little impact on highly mobile organisms not dependent on substrate type and capable of avoiding brackish water by swimming farther from the coast. This may explain why it has not been previously explored as a hypothesis causing suture zones in the Gulf of Mexico in the literature. However, this proposed ancient suture zone may be especially important for other organisms dependent on sandy substrate such as some polychaete worms (for example, *Scolelepis squamata*), mole crabs (*Emerita* spp.), and surf clams (*Donax* spp.), although pelagic larvae in these organisms may obscure more ancient vicariant patterns (Adamkewicz & Harasewych, 1996; Goulding & Cohen, 2014). Other organisms closely related to amphipods, such as other brooding crustaceans (e.g., Hurtado et al., 2016; Sutherland et al., 2010; Takada et al., 2018), may be best suited to evaluating the Miocene sedimentation hypothesis.

A major suture zone is not the only factor that may have contributed to present species distributions. Differences in physiological tolerance across OTUs may also explain the persistent distinction of clades even after the sedimentation subsided. In the western Gulf, the sand is rich with silt which limits the ability for water to percolate to deeper interstitial spaces causing the redox layer to be shallow (Revsbech et al., 1980). To the east of the Mississippi, the sand is

coarser allowing water to drain deeper and thereby increasing the habitable zone of burrowing amphipods. These differences in depth of the redox layer may impact which species can occupy the beach. For example, sympatric amphipods in South Carolina were found to partition the beach vertically by tolerance to anoxia (Grant, 1981). Variation in sand grain size has also been proposed to explain the over two-fold increase in body size in the eastern Gulf amphipods compared to those collected in the west (Hancock & Wicksten, 2018; LeCroy, 2002).

#### *Population connectivity and dispersal*

Significant population structure was found between almost all site comparisons (**Fig. 1.6**), which supports the hypothesis that haustoriid amphipods have limited dispersal abilities. However, dramatic differences in geographic range exist between the two genera. *Haustorius galvezi* (Mexico) ranges from at least South Padre Island, Texas, to southern Tabasco, Mexico (~1,200 km), whereas *Lepidactylus triarticulatus* contained entirely site-specific lineages. This could indicate behavioral differences associated with entering the water column and vertical beach migration. On beaches where *L. triarticulatus* and *H. galvezi* are found in sympatry, the former tends to occupy the high intertidal with the latter at the mid-to-low intertidal zones (Shelton & Robertson, 1981). In the eastern Gulf, LeCroy (2002) noted that *L. triarticulatus* is most dominant on protected beaches and is replaced by *H. jayneae* on open, wave-swept beaches. These differences in ecology likely expose *Haustorius* spp. to wave action that may dislodge them from the sediment, allowing for passive dispersal by the currents. *Lepidactylus triarticulatus*, alternatively, is largely protected from this in the higher intertidal and in calm, sandy bays.

Despite morphologically identifying all specimens from sites south of Sea Rim State Park in the western Gulf as *H. galvezi*, strong genetic divergence was uncovered along the southern Texas coastline. The northern and southern clades shared no haplotypes and were ~3% diverged at *COI* despite being separated by less than 200 km. These clades roughly correspond with defined ecoregions along the Texas coast, which may indicate a role of hydrological or geochemical differences on speciation (Mendelsohn et al., 2017). For example, the northern Texas coast has lower annual salinity than the south due to the high concentration of rivers flowing into the ocean (Zavala-Hidalgo et al., 2003). In addition, since haustoriid amphipods are passive dispersers, the current regime along the Texas coast may not favor migration northward from South Padre or southward from Port Aransas due to a counterclockwise flowing gyre between these sites (Mendelsohn et al., 2017; Zavala-Hidalgo, 2003). This gyre varies in intensity seasonally, but it may be prominent enough to act as a barrier to gene flow.

### *Conclusions*

The Gulf of Mexico has a complex geological history of episodic sedimentation, freshwater inflow, and sea level alterations that each may have impacted the distributions of coastal marine organisms in the past. Previous biogeographic work in the Gulf of Mexico focused on organisms with planktonic larvae or those that are highly mobile, which limited the ability of these studies to detect ancient divergences. By using a fossorial, filter-feeding organism without planktonic larvae, I was able to detect deep divergences between sister clades of haustoriid amphipods isolated on opposite sides of the Mississippi River. Divergence-time dating indicated that these splits occurred during the Miocene, and due to the distribution of

sister taxa I inferred massive sedimentation event along the Louisiana coastline was the most likely explanation.

Massive freshwater outflows have been shown to act as barriers to gene flow in coastal organisms in other systems; for example, the Orinoco River and the Amazon River freshwater plumes act as population breaks in the coastal sea urchin, *Tripneustes* (Lessios et al., 2003). Evaluating how widespread the impact of the Miocene sedimentation event was will require additional genetic analysis of coastal organisms, and I propose that other peracarids (isopods, mysids, and tanaids) will be most suited as, like amphipods, they are brooding crustaceans without pelagic larvae (e.g., Hurtado et al., 2016; Takada et al., 2018; Drumm & Kreiser, 2012). Notably, Haustoriidae is a relatively species-poor family in comparison with other amphipod families. However, this study has demonstrated widespread cryptic diversity exists in this group (possibly indicating the family as a whole has considerably more species than what are morphologically recognized), likely due, at least in part, to an ancient suture zone in the Gulf of Mexico.

# CHAPTER II

## RAPID GENOMIC EXPANSION AND PURGING ASSOCIATED WITH HABITAT TRANSITIONS<sup>2</sup>

### Introduction

Genome sizes vary by several orders of magnitude across the Tree of Life and lack any correlation to organismal complexity (known as the *C*-value paradox; Gregory 2005; Lynch 2007). For example, the fern *Tmesipteris obliqua* has a genome 50 times larger than ours (Hidalgo et al 2017). Several hypotheses have been proposed to explain this variation. The mutational-hazard hypothesis posits that genome size expansion is largely the result of the proliferation of parasitic DNA such as transposable elements (TEs), which is made possible by relaxed selection on genomes (Lynch & Conery 2003; Lynch 2007; Lefébure et al. 2017). Therefore, in organisms with low effective population sizes ( $N_e$ ), we would expect a general increase in genome size due to weakened selection against these mildly deleterious expansions. In addition, selection may act against large genomes as they are metabolically costly to replicate, slow the rate of replication and cell division (Kozlowski et al 2003; Gregory 2005) and increase cell size (Cavalier-Smith 1978). Genome size expansion may also reduce developmental rates and therefore expand generation times (Gregory & Johnston 2008), which may lead to an increase in body size. For example, in *Manduca sexta* (Sphingidae) faster developmental rates at higher temperatures led to a shortened interval of cessation of growth (ICG), which limited the

---

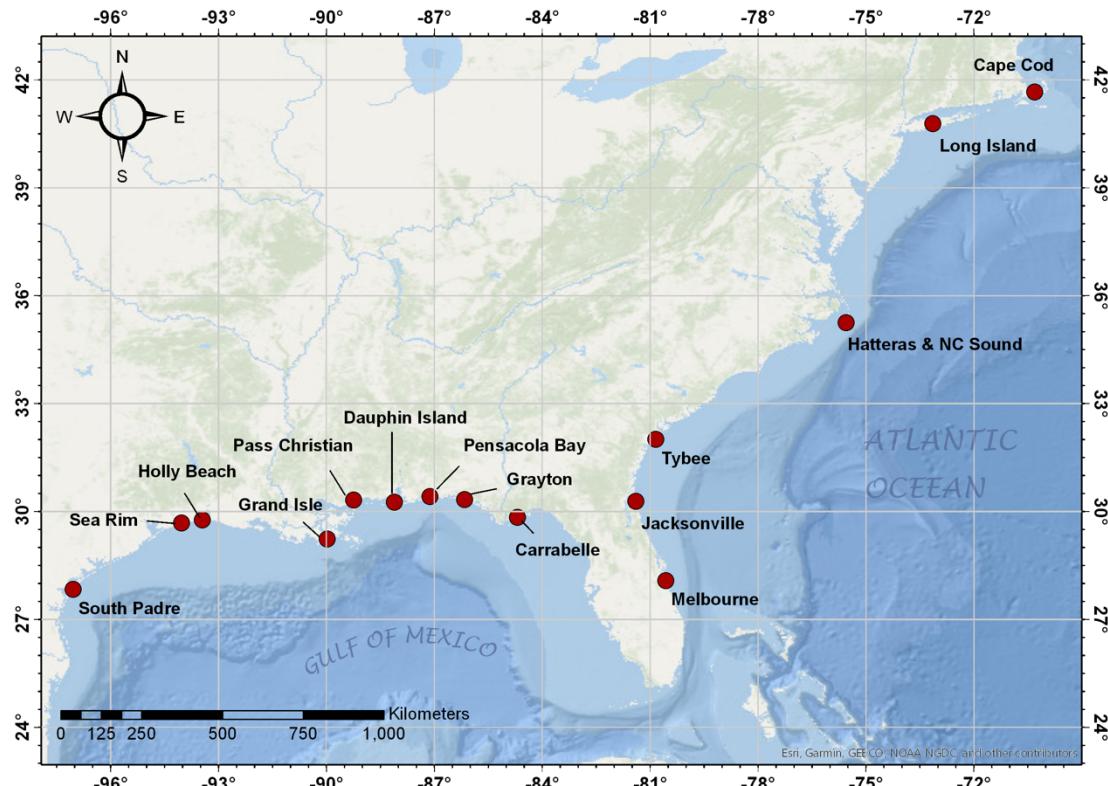
<sup>2</sup> Hancock, ZB, Hardin, FO, Murthy, A, Hillhouse, A, Johnston, JS. Rapid genomic expansion and purging associated with habitat transitions in a clade of beach crustaceans (Haustoriidae). *Genome Biology & Evolution*, in review.

amount of mass that could be accumulated and therefore decreased body size (Davidowitz & Nijhout 2004).

While genome size variation exists at deep macroevolutionary scales, in animals (but not plants, e.g. Hloušková et al 2019) it is largely constrained at smaller scales with closely related organisms having similar genome sizes (e.g., Tiersch & Wachtel 1991; Sessegolo et al 2016; Roebuck 2017; but see Lower et al 2017). Covariance between relatedness and genome size has been used as evidence against the mutational-hazard hypothesis (Whitney & Garland 2010; but see Lynch 2011). Alternatively, congruence between genome size and phylogenetic relatedness could be a product of the lag in evolutionary timescales over which TEs noticeably proliferate compared to reductions in long-term  $N_e$ . Therefore, a step toward disentangling the evolutionary rate of TEs and long-term  $N_e$  variation on phylogenetic genome size congruence would be to identify young clades of organisms with dramatic differences in genome size that may represent a kind of “upper-bound” to TE expansion rates in nature.

Spatial heterogeneity in selection may also explain genome size differences across a clade. Since genome sizes impact metabolic rates, environmental gradients that directly impact physiology, such as temperature or precipitation, may influence genome size variation. Alfsnes et al. (2017), in a study across Arthropoda, found that for Hexapoda there was a strong phylogenetic signal in genome size that appeared to be related to developmental strategies, whereas in crustaceans the relationship was less reliant on phylogeny and more a product of maximum latitude and water depth. Hultgren et al. (2018) found correlations between genome size and latitude in Amphipoda, and Jeffrey et al. (2016) found that genome sizes of Lake Baikal amphipods were correlated with water depth, body size, and diversification rates.

Crustaceans, and many ectotherms in general, often follow a “temperature-size rule” (Walters & Hassell 2005) in which body size increases with decreasing temperature, whether at deeper waters or higher latitudes. Given that previous work has found an association between genome size, body size, and latitude in some crustacean groups (e.g., Hultgren et al 2018; Jeffrey et al. 2016; Alfsnes et al. 2017), it is conceivable that there may be a general “genome-temperature-size rule”. Specifically, when body size is dictated by temperature, genome size should also be driven by temperature. This line of reasoning originates from the fact that at lower temperatures selection on metabolic rates is weaker, and therefore genome sizes are free to expand (Gregory 2005).



**Figure 2.1.** Map of sample sites.

Haustoriid amphipods are a family of beach-dwelling crustaceans that are widely distributed across the northern hemisphere. These amphipods are morphologically specialized to a fossorial lifestyle and display dramatic body size variation across their range (LeCroy 2002; Hancock & Wicksten 2018). In the Gulf of Mexico (GoM), haustoriids show an increase in body size west-to-east, with the largest in Florida and the smallest in Texas and Louisiana. A similar pattern of body size variation is seen in the Atlantic; however, body sizes increase south-to-north, with the smallest in Florida relative to the largest in Massachusetts. Therefore, haustoriid amphipods represent a natural system to test a possible “genome-temperature-size rule”. Since mean annual temperatures do not vary across the GoM but body sizes do, we do not expect to see a change in genome sizes if temperature is the constraining factor. Alternatively, in the Atlantic where mean annual temperatures vary dramatically from the southern to northernmost latitudes, we expect genome sizes to increase with increasing latitude. In this way, we can tease apart the correlation between body size and temperature seen in most amphipods (and ectotherms in general), and isolate the causal factor influencing genome size in this clade.

Previous work on this family has found strong population structure and widespread cryptic diversity (Hancock et al. 2019). In the GoM there are at least 6 species of haustoriids: *Haustorius galvezi*, *H. jayneae*, and a species complex of four distinct lineages of *Lepidactylus triarticulatus*. In the Atlantic, there exists only one “species” of *Haustorius*, *H. canadensis*, which ranges from at least Melbourne, Florida to Cape Cod, Massachusetts. This species shows incredible body size variation – at its lower range it is 4–5 mm in length, whereas at its northernmost range it can be as large as 18 mm (LeCroy 2002). Given the extent of cryptic

diversity in this group already identified in the GoM (Hancock et al 2019), I expect that similar diversity exists in the Atlantic and that *H. canadensis* likely represents a cryptic species complex.

In this study, I evaluate genome size variation within haustoriid amphipods and test for correlations between environmental and phenotypic variables such as latitude, temperature, salinity, and body size to evaluate the existence of a general “genome-temperature-size rule”. I provide a time-calibrated phylogeny of the Haustoriidae to determine the timescale over which genome size variation occurs, and use this phylogeny to reconstruct ancestral environments, body sizes, and genome sizes. Finally, I perform low-coverage next-generation sequencing to characterize repetitive content across the genome to identify which families of repeat elements might be driving genome expansion.

## Methods

### *Specimen collections and measurements*

Amphipods were collected across 9 sites in the GoM covering roughly 20° longitude and 6 sites in the Atlantic spanning 20° latitude (**Fig. 2.1**). Specimens were collected in the swash zone using a hand-shovel and 435 $\mu$ m sieve plate. Individuals targeted for genome size estimation were removed from the sieve plate with forceps and either kept alive in sand and seawater from their local environment or were fresh frozen on dry ice. For specimens in which genome sequencing was to be performed, they were removed from the plate and placed immediately into 95% EtOH. Specimens were identified to species in the lab using a Leica M205 FA dissecting scope, and I recorded body length for each specimen as distance (in mm) from rostrum to epimeron 3. At each site, I recorded both water temperature and salinity (using a refractometer).

In addition, I categorized sand grain size at each location as “fine”, “medium”, and “course” based on how freely it passed through sieve plates of 1.18mm, 0.6mm, and 0.3mm. Sand denoted “fine” passed through all plates without retaining any shells or rocks; “medium” sand passed through the first two without retaining shells or rocks but not the final; “coarse” sand passed through the first and second, but large amounts of rocks and shell hash were retained.

#### *Genome size estimation*

Genome size estimation via flow cytometry was performed on fresh frozen or live samples as described in Johnston et al. (2019), Arnqvist et al. (2015), and Harahan & Johnston (2011). In brief, the anterior portion (excluding reproductive tissues) of each amphipod was placed into 1ml of Galbraith buffer in a 2ml Dounce tube (Kontes) along with the brain tissue from two standards, a female cowpea weevil (*Callosobrucus maculatus*; 1C = 1,233 Mbp) and a male American cockroach (*Periplaneta americana*; 1C = 3,338 Mbp). Nuclei from the sample and standard were released with 15 strokes of the loose (B) pestle. Debris was reduced by filtering the ground solution through a 40U nylon filter. DNA was stained for 2 hours in the cold and dark with propidium iodide (25 ppm). Mean stain uptake in the 2C nuclei of the standards and the sample were determined using a CytoFlex Flow Cytometer (Beckman Dickenson). The genome size of each sample was calculated as the ratio of the mean fluorescence (output as a channel number) of the sample and the standard times the amount of DNA in the standard. Preliminary estimates found no sex-specific differences in genome size, and thus I pooled males and females. Unfortunately, sample degradation reduced our sample size to 1 for many sites, precluding us from evaluating within-site variation (**Table 2.1**).

**Table 2.1.** Summary of samples and their genome sizes. \*Sample sizes specifically for genome size estimates.

Species	Location	Sample size*	Genome size (Mb)	Coverage	Mean body length (mm)	Sand grain size
<i>N. schmitzi</i>	Tybee Island, GA	<i>n</i> = 1	7500	0.0221	2.0	Fine
<i>H. canadensis</i>	Melbourne, FL	<i>n</i> = 1	11050	0.0213	6.25	Middle
<i>H. canadensis</i>	Jacksonville, FL	<i>n</i> = 1	11690	0.0247	6.17	Fine
<i>H. canadensis</i>	Tybee Island, GA	<i>n</i> = 1	13080	0.0164	5.64	Fine
<i>H. canadensis</i>	Hatteras, NC	<i>n</i> = 1	12830	0.0193	6.42	Middle
<i>H. canadensis</i>	Long Island, NY	<i>n</i> = 1	11990	0.0216	9.0	Coarse
<i>H. canadensis</i>	Cape Cod, MA	<i>n</i> = 1	11390	0.016	13.78	Coarse
<i>H. galvezi</i>	Port Mansfield, TX	<i>n</i> = 5	7350	0.034	4.0	Fine
<i>L. triarticulatus</i>	Sea Rim, TX	<i>n</i> = 5	2200	0.036	3.2	Fine
<i>L. triarticulatus</i>	Grand Isle, LA	<i>n</i> = 1	2380	0.027	3.26	Fine
<i>H. allardi</i>	Holly Beach, LA	<i>n</i> = 5	2130	0.053	2.4	Fine
<i>L. triarticulatus</i>	Pass Christian, MS	<i>n</i> = 1	3000	0.011	3.6	Middle
<i>L. triarticulatus</i>	Pensacola Bay, FL	<i>n</i> = 1	2900	0.009	5.74	Middle
<i>L. dysticus</i>	North Carolina Sound	<i>n</i> = 1	7520	0.025	5.0	Middle

#### *DNA extraction and sequencing*

Whole genomic DNA was extracted from either the whole specimen or pereopods 6–7 depending on the size of the amphipod using an EZNA Tissue DNA kit (Omega Bio-tek Inc.)

following manufacturer's protocols. Mitochondrial cytochrome oxidase I (*COI*) was amplified using forward and reverse primers from Folmer et al. (1994), and nuclear 28S ribosomal RNA (28S) was amplified using primers from Hancock et al. (2019). Polymerase chain reaction (PCR) conditions followed Hancock et al. (2019) for both mitochondrial and nuclear loci. Amplicons were verified using gel electrophoresis and purified with ExoSAP-IT (Affymetrix Inc.). Sanger sequencing on forward and reverse strands was performed at the DNA Analysis Facility on Science Hill at Yale University. Sequences were cleaned and manually edited using Sequencher v4.10.1 (Gene Codes Corp.), and alignments were performed in MAFFT 7 (Katoh & Standley 2013). The *COI* sequence was visually checked in Mesquite v3.5 (Maddison & Maddison 2018) to ensure there were no premature stop codons.

To generate repeat profiles, I performed low-coverage Illumina sequencing on 85 individuals from 15 locations across the GoM and North Atlantic (**Table 2.1; Fig. 2.1**). DNA was quantified with the dsDNA high sensitivity Qubit Assay (ThermoFisher) and checked for quality with an Agilent TapeStation genomic DNA tape (Agilent Technologies). Sequencing libraries were generated using the Swift 2S Turbo DNA Library preparation kit with enzymatic fragmentation and combinatorial dual indexing following the manufacturer's instructions (Swift Biosciences). Libraries were pooled in equimolar ratios and sequenced in a single lane of 2 x150 paired-end sequencing on an Illumina NovaSeq 6000 (Illumina).

#### *Characterization of repeat elements*

Quality control was performed on FASTQ files in the RepeatExplorer Galaxy platform (<https://repeatexplorer-elixir.cerit-sc.cz/>). For each population, individual left-hand and right-

hand sequences of all individuals were first concatenated to increase coverage. I designated an acceptable Phred-score of 20 and a 95% cutoff on reads with bases less than this score. This was followed by trimming reads to a uniform 100 bp to allow straightforward calculation of coverage percent downstream. Reads shorter than 100 bp were discarded. Left- and right-hand reads were then interlaced and scanned for overlap. Next, I used RepeatExplorer v2.3.8, which uses a graphical clustering approach to identify repeat profiles using sequencing reads and has been shown to perform well with low-coverage sequencing (Novák et al 2010; Lower et al 2017; Hloušková et al 2019). I specified the Metazoa 3.0 protein database for repeat annotation. For each population, I ran RepeatExplorer for the maximum number of processible reads, which ensured coverage was  $\geq 1\%$  for all populations (**Table 2.1**). Only clusters that occurred in  $> 0.01\%$  of the sampled reads were annotated, the remaining fell into a “bottom clusters” category.

#### *Phylogenetic inference and comparative methods*

To determine whether populations of the same “species” could be considered as independent evolutionary lineages (i.e., represented cryptic diversity), I performed pairwise genetic differentiation tests in DnaSP v6 (Rosas et al 2017) with 5–10 individuals per population using the mitochondrial locus. I performed the three measures of differentiation from Hudson et al (1992), namely  $H_{ST}$ ,  $K_{ST}$ , and  $K_{ST}^*$ . The first,  $H_{ST}$ , is measured as  $1 - (H_S / H_T)$ , where  $H_S$  is the weighted average of haplotype diversity within the subpopulation and  $H_T$  is the total population haplotype diversity.  $K_{ST}$  is measured as  $1 - (\pi_{12} / \pi)$ , where  $\pi_{12}$  is the weighted average of nucleotide differences between site 1 and 2, and  $\pi$  is the average number of differences irrespective of locality. Finally,  $K_{ST}^*$  is identical to  $K_{ST}$  except  $\pi_{12}$  is changed to  $\log(1$

$+\pi_{12}$ ), which acts to downweight  $\pi_{12}$  when it becomes high (Hudson et al 1992). Significance was determined using a permutation test with 1000 replicates. Populations with significant differentiation values ( $\alpha = 0.05$ ) in at least two of the three tests were retained as “independent”; otherwise, one sample site was randomly discarded.

To estimate the age of the clade and to perform phylogenetic comparative methods, I first constructed a time-calibrated tree using BEAST2 (Bouckaert et al 2019). I specified a GTR model of substitution for *COI* and an HKY model for *28S*, as these were the indicated best models in PartitionFinder2 (Lanfear et al. 2017). I applied a molecular clock to *COI* as an exponential prior with a mean of 0.01 substitutions per site per Myr (Hancock et al 2019; Takada et al 2018; Knowlton & Weigt 1998). I applied a relaxed lognormal clock to *28S* following Hancock et al (2019). I then applied two calibration points based on known vicariant events: 1) the closure of the Okefenokee Trough ~1.75 Myr separating Atlantic and Gulf taxa (Bert 1986; McClure & Greenbaum 1999); and 2) the proposed interglacial Pleistocene colonization of *H. arenarius* to Europe (Bousfield 1970). For the first calibration, I applied an exponential prior offset to 1.5 Myr (a conservative lower bound for the closure of the trough) to the *Acanthohaustorius* sp. and *A. millsii* clade, GoM and Atlantic endemics, respectively. For the second calibration I applied a gamma distributed prior with  $\alpha = 1.5$  and  $\beta = 1.0$ , which has a mean of ~1.18 Myr, roughly in the middle of the Pleistocene. The MCMC was run for 50 million generations ensuring ESS values  $> 200$ . I generated a consensus tree using TreeAnnotator (Bouckaert et al 2019), which was visualized in FigTree v1.4.3 (Rambaut 2011).

The resulting tree was imported into the R platform and pruned down to only the taxa in which genome size data was available using the package *ape* v5.3 (Paradis & Schliep 2019). I next used the package *caper* v1.0.1 (Orme et al 2018) to perform phylogenetic generalized least

squares (PGLS) on associations between genome size and environmental variables, body size, and repeat element profiles. This allowed me to estimate Pagel's  $\lambda$ , which ranges from 0.0–1.0 and indicates how much phylogenetic signal is present in the data (Pagel 1999). I performed a combination of two-variable linear regressions (e.g., between body length and genome size) as well as models that included interactions between predictor variables. I performed the latter because there is expected to be more than one variable influencing body size. For example, body size may be a function of an interaction between temperature and genome size. In addition, I performed ancestral state reconstructions on both the continuous and discrete variables using the package *phytools* (Revell 2012). These included genome size, repeat content (specifically LINEs), body size, sand grain size, and habitat ("open" versus "brackish").

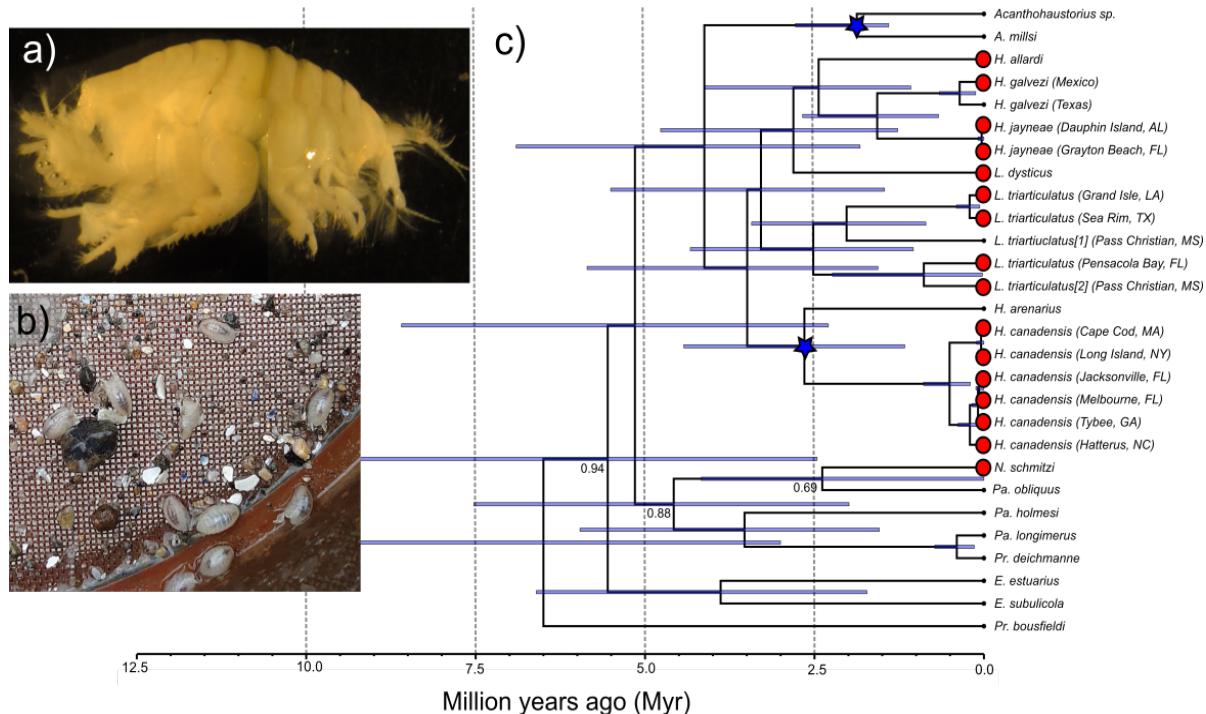
## Results

### *Genome size variation and age of the clade*

Estimates from flow cytometry indicated genome sizes ranged from 2,130 Mb to 13,080 Mb (**Table 1**). The amphipods with the smallest genomes were all restricted to warm, brackish water – namely, *H. allardi* and the *L. triarticulatus* species complex. The largest genomes belonged to the *H. canadensis* complex, with the largest (13,080 Mb) found at Tybee Island, Georgia.

All pairwise differentiation comparisons were significant for at least two measures except for between *H. jayneae* sampled at Carrabelle Beach, FL and Grayton Beach, FL / Dauphin Island, AL (**Table S2.1**). Carrabelle Beach was thus dropped from subsequent analyses.

The calibrated phylogenetic tree indicated that the entire haustoriid clade is less than 12 Myr old, and that the clade in which I performed genome size estimates is only ~3.5 Myr old (Fig. 2.2c); however, the range of the 95% highest posterior density (HPD) for the clade was large indicating wide uncertainty in the age estimates (1.67–6.25 Myr).

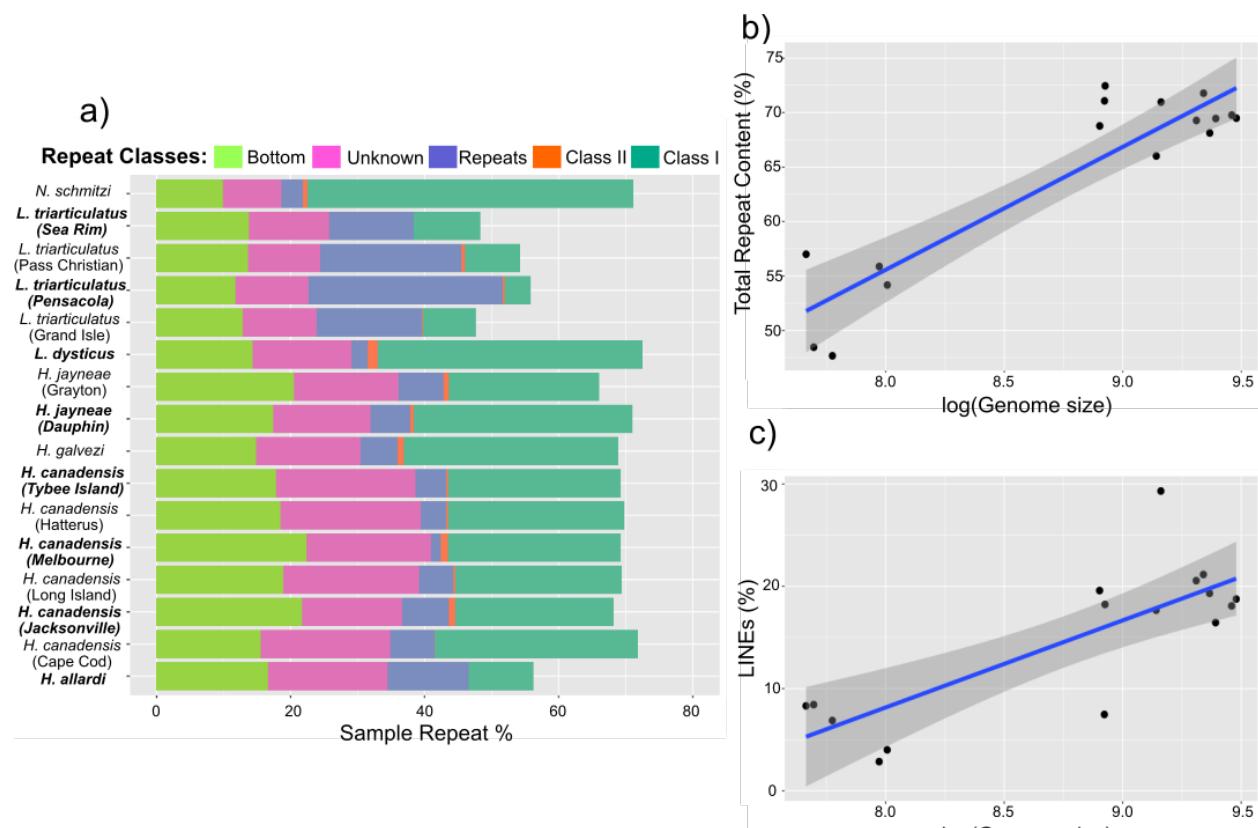


**Figure 2.2.** Calibrated phylogeny of the Haustoriidae. a) *L. triarticulatus* from Pensacola Bay, FL; b) haustoriids on a sieve plate; c) time-calibrated phylogeny. Red circles are samples with genome size estimates; blue stars are calibration points; node bars represent 95% HPD. All nodes have PP > 0.95 unless

### Repeat profiles

The most abundant annotated repeat elements were Class I repeats, the genomic proportion of which ranged from 3.9% (*L. triarticulatus*, Pensacola Bay, FL) to 48.6% (*Neohaustorius schmitzi*). Of the Class I elements, LINEs were the most abundant, accounting

for as much as 29% of the genome (*H. jayneae*, Dauphin Island, AL). LTRs were less prevalent, normally less than 6% (see **Table S2.2**), but two taxa showed dramatic LTR expansions: *N. schmitzi* (18.9%) and *L. dysticus* (17.7%). The only Class II repetitive element identified was the large transposon *Maverick*, which made up at most 1.5% of the genome (*L. dysticus*). Large satellite repeats were identified in all species, which ranged from 1.5% (*H. canadensis*, Melbourne, FL) to 28.5% (*L. triarticulatus*, Pensacola Bay, FL) of the genome. Unlike the previous repeat elements, satellite repeats appeared to be the most prevalent in the smallest genomes (**Fig. 2.3a; Table S2.2**).



**Figure 2.3.** Repetitive content. a) Barplot of repeat content proportion per sample, colored by repeat element family, every other species name bolded to aid in visualization; b) correlation between  $\log(\text{genome size})$  and total repetitive content; c) correlation between  $\log(\text{genome size})$  and LINE proportion.

I found no correlation between latitude and genome size with either the raw data ( $p = 0.08$ ,  $r^2 = 0.19$ ) or incorporating phylogeny ( $p = 0.20$ ,  $r^2 = 0.11$ ). I also found no correlation between genome size and salinity ( $p = 0.67$ ,  $r^2 = 0.01$ ,  $\lambda = 1.0$ ) when accounting for phylogeny, although a significant relationship was found in the raw data ( $p = 0.0004$ ,  $r^2 = 0.56$ ). However, I found a strong correlation between body length and genome size in both the raw ( $p = 0.01$ ,  $r^2 = 0.33$ ) and the PGLS comparisons ( $p = 0.01$ ,  $r^2 = 0.33$ ,  $\lambda = 0.0$ ). For the body length mixture models, I found the model with the lowest AIC score (60.18) was one that included genome size ( $p = 0.01$ ), temperature ( $p = 0.04$ ), and an interaction between genome size and temperature ( $p = 0.02$ ; see **Table 2.2**). In addition, I found a strong positive correlation between latitude and body size ( $p = 0.0007$ ,  $r^2 = 0.57$ ,  $\lambda = 0.0$ ).

I found a positive relationship between genome size and total repetitive content ( $p = 0.000047$ ,  $r^2 = 0.71$ ), and a ML estimate of Pagel's  $\lambda = 0.82$ , indicating that there was strong phylogenetic signal in total repeat content (**Fig. 2.3b**). Of the different identified repeat families, Class I repeats had the strongest correlation with genome size ( $p = 0.004$ ,  $r^2 = 0.44$ ,  $\lambda = 0.87$ ), with this pattern being driven by LINEs ( $p = 0.001$ ,  $r^2 = 0.53$ ,  $\lambda = 0.49$ ; **Fig. 2.3c**). An opposite relationship was found with satellite repeats due to the preponderance of these in the smallest genomes ( $p = 0.04$ ,  $r^2 = 0.26$ ,  $\lambda = 0.91$ ). There was no relationship between percent coverage and total genomic repeat content ( $p = 0.34$ ,  $r^2 = 0.06$ ).

The ancestral state reconstruction analysis revealed a general correlation between increases in genome size and body length, with the latter lagging behind the former (**Fig. 2.4**). In addition, I found that the ancestral environment was most likely an open coast with high salinity,

and that there have been at least four independent transitions into brackish, protected beaches: 1) *Eohaustorius* on the Pacific coast; 2) the *L. triarticulatus* species complex in the GoM; 3) *L. dysticus* in the North Atlantic; 4) and *H. allardi* along the Texas and Louisiana coastline (**Fig. 2.5**). Finally, I found a weak relationship between sand grain size and body length (**Fig. 2.4b**): the smallest haustoriids occur in fine sand and the largest in coarse, though there is considerable overlap in body sizes between fine and medium.

**Table 2.2.** Summary of genomic correlation models. Abbreviations: TR (total repeat %), GS (genome size), C1 (Class I repeats), C2 (Class II repeats), RE (satellite repeats), LTR (LTR-retrotransposons), LI (LINEs), BO (bottom clusters), Lat (latitude), BL (body length), SA (salinity), T (temperature).

Model	df	P value	$r^2$	Pagel's $\lambda$	$\lambda$ 95% CI	AIC
TR ~ log(GS)	14	<b>4.713e-05</b>	0.7052	0.824	0.127–0.961	-
C1 ~ log(GS)	14	<b>0.004834</b>	0.4439	0.872	0.481–0.970	-
C2 ~ log(GS)	14	0.1064	0.1754	0.000	0.000–0.813	-
RE ~ log(GS)	14	<b>0.04179</b>	0.2639	0.907	0.000–0.994	-
LTR ~ log(GS)	14	0.194	0.1174	0.995	0.938–1.000	-
LI ~ log(GS)	14	<b>0.001314</b>	0.5334	0.491	0.000–0.856	-
BO ~ log(GS)	14	0.1513	0.1413	0.514	0.000–0.914	-
log(GS) ~ log(Lat)	14	0.2013	0.1138	1.000	0.981–1.000	-
BL ~ Lat	14	<b>0.0007267</b>	0.5696	0.000	0.000–0.827	-
log(GS) ~ SA	14	0.6787	0.01262	1.000	0.958–1.000	-
BL ~ T + log(GS) + T*log(GS)	12	<b>0.0001575</b>	0.8027	0.000	0.000–0.870	<b>60.17677</b>
BL ~ T + log(GS)	13	<b>0.0004119</b>	0.6986	0.000	0.000–0.824	64.96170
BL ~ T	14	<b>0.0007938</b>	0.5644	0.490	0.000–0.945	67.21072
BL ~ log(GS)	14	<b>0.01929</b>	0.3329	0.000	0.000–0.744	75.67192

## Discussion

Is there a general “genome-temperature-size-rule” that explains haustoriid genome evolution? I found mixed evidence for this, indicating that the factors that influence genome size are complex and multifaceted. While I found evidence from mixed models of an interaction between genome size and temperature on body length, there was not a significant relationship between latitude and genome size ( $p = 0.08$ ). Instead, genome size variation in haustoriid amphipods may be the result of differing environments, with amphipods living on surf-swept, high salinity beaches having increased body length and genome size compared to those living in warm, brackish bays (Fig. 2.5). Further, I found that genome size and repeat element prevalence have strong phylogenetic signal (Table 2.2).

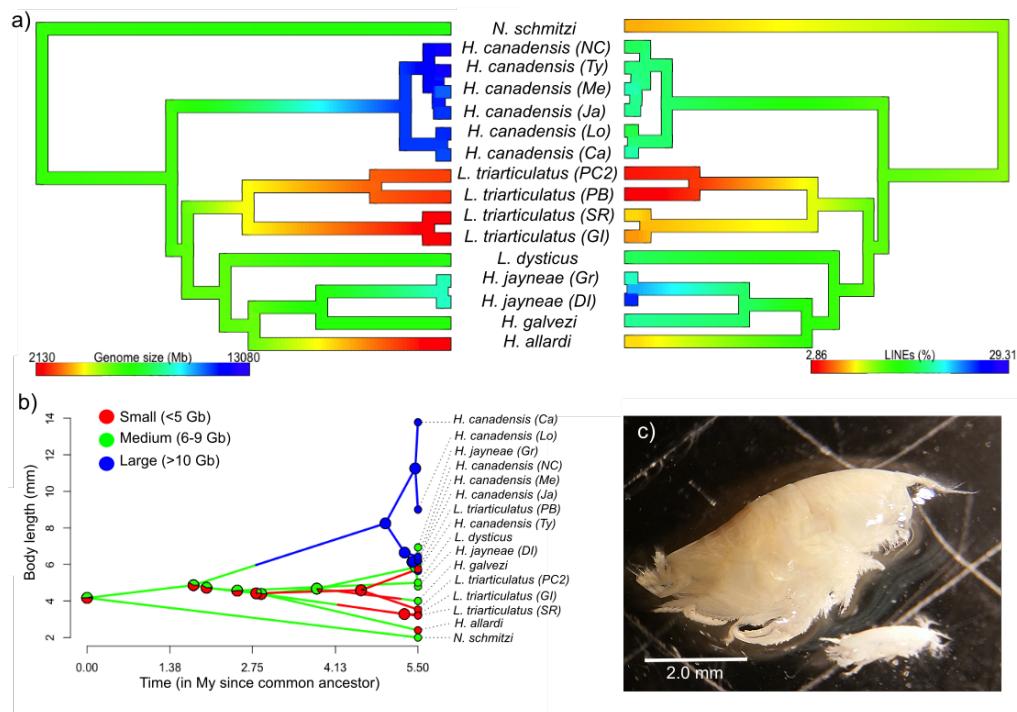
#### *The drivers of genome size evolution in Haustoriidae*

Haustoriid amphipods display incredible genome size variation across relatively recent evolutionary time (Fig. 2.4a). Most haustoriid genomes are large (>7000 Mb), but this may be an ancestral condition. The most closely related amphipod in the Animal Genome Size Database (<http://www.genomesize.com/>; Gregory 2020) are the pontopoerids *Monoporeia affinis* and *Pontoporeia femorata*, each of which have genome sizes of 8242 Mb. Therefore, *N. schmitzi*, *L. dysticus*, *H. galvezi*, and *H. jayneae* may represent cases of relative genomic stasis ( $\pm 1000$  Mb). Alternatively, the *H. canadensis* clade has experienced rapid genomic expansions within the last 3 Myr. It should be noted that *H. canadensis* and its sister species, *H. arenarius*, are some of the largest fossorial amphipods and are the most abundant at higher latitudes.

In stark contrast, the predominantly warm, brackish water clades of *L. triarticulatus* and *H. allardi* show evidence of massive genomic purging. If the ancestral genome size was ~8000 Mb, these clades have lost on average 5500 Mb. These genomes are largely devoid of LINEs, which occupy up to 30% of the genomes of their relatives. This may indicate that genomic purging has targeted transposable elements specifically, which has been found in other taxa (Kelley et al 2014; Barrón et al 2014; Lyu et al 2018; Misof 2019). Notably, I also found that these genomes have a much higher percentage of satellite repeats than their relatives, especially the *L. triarticulatus* species complex. Since these repeats are distributed throughout this complex, it likely represents an ancestral expansion.

Brackish bays present unique physiological stresses not present on the open coast. Bays tend to be on average warmer than the open ocean, and in the GoM can reach 34°C in the summer months (Ross & Behringer 2019; Holmquist et al 1989; **Table 2.1**). These higher temperatures reduce the capacity of water to retain dissolved oxygen, which may increase respiratory stress in some organisms compared to the highly oxygenated sands of the open coast (Bousfield 1970). Bays also have shallow anoxic layers due to densely packed organic matter (Revsbech et al 1980). Finally, due to freshwater runoff and poor mixing, bays are typically brackish and may have salinities as low as 5‰ (this study), which increases osmotic stress on cells. Haustoriid amphipods are in general tolerant of wide salinity fluctuations, but only a few clades (*Eohaustorius estuarius*; the *L. triarticulatus* species complex; *L. dysticus*; and *H. allardi*) have specialized to live in bays and freshwater inlets, possibly due to the reasons above. These species may also be found on the open coasts, but they tend to be less numerous and relegated to the subtidal zone (LeCroy 2002; Hancock & Wicksten 2018).

Temperature and salinity have been shown to have synergistic impacts on oxidative stress, such that some species may tolerate high temperatures or low salinity, but not in conjunction (Takolander et al 2017). This may explain why the two observed GoM shifts to brackish environments are accompanied by body size and genome size reductions, but not the Atlantic shift (i.e., *L. dysticus*; **Fig. 2.5c**). The North Atlantic consistently has lower sea surface temperatures (SST) than the GoM, and therefore, for much of its range, *L. dysticus* may be exposed to low salinity but not high temperature.



**Figure 2.4.** Ancestral state reconstructions. a) genome size reconstruction (left) and percent LINE content (right); b) phenogram of body length and genome size, node labels represent posterior state probabilities; c) body size comparison of a large haustoriid (*H. arenarius*) and a small one (*H. allardi*). Abbreviations: NC (Hatteras, North Carolina); Ty (Tybee Island, Georgia); Me (Melbourne, Florida); Ja (Jacksonville, Florida); Lo (Long Island, New York); Ca (Cape Cod, Massachusetts); PC2 (Pass Christian, Mississippi); PB (Pensacola Bay, Florida); SR (Sea Rim, Texas); GI (Grand Isle, Louisiana); Gr (Grayton, Florida); DI (Dauphin Island, Alabama).

Given the increased metabolic stress associated with transitioning to warm brackish bays, I hypothesize that strong selection on faster metabolism led to rapid genomic purging in *L.*

*triarticulatus* and *H. allardi*. In contrast, reduced metabolic stress at higher latitudes in *H. canadensis* may have provided a selectively permissive environment for transposable elements to proliferate.

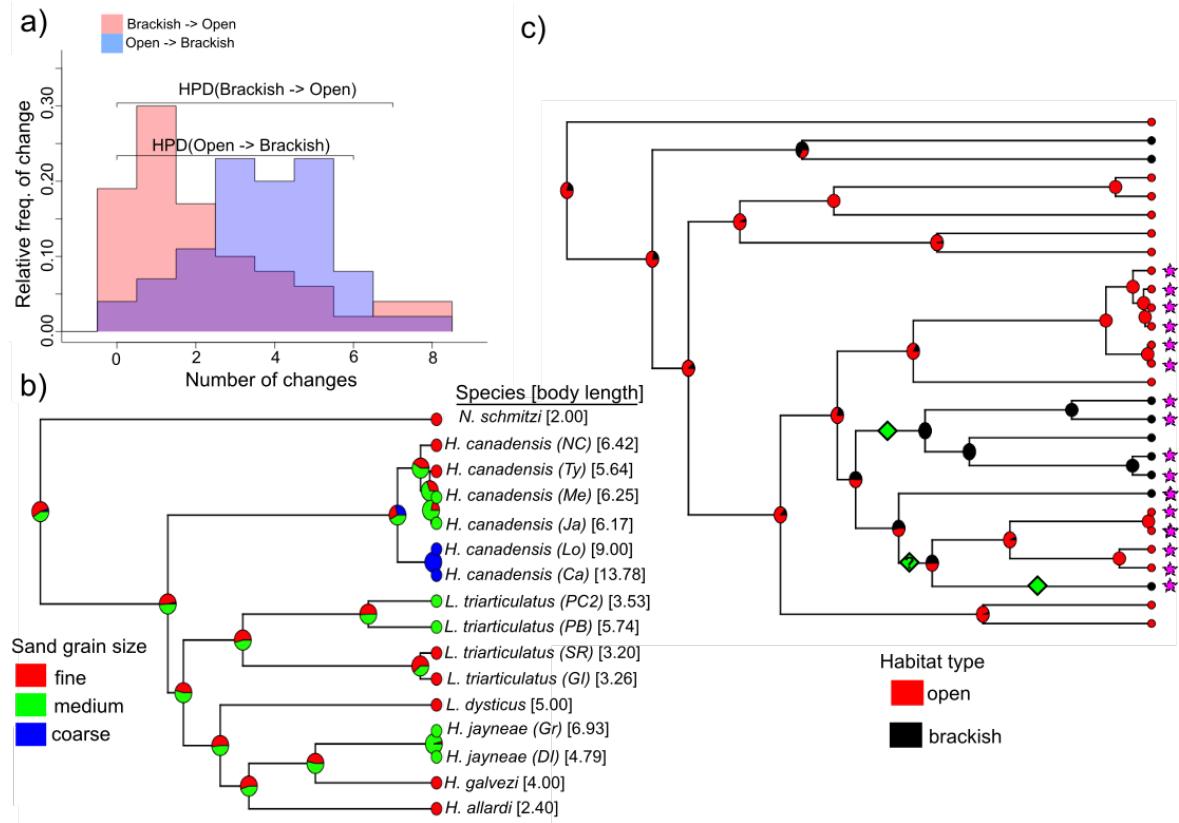
### *Body size evolution in Haustoriidae*

Haustoriid amphipods range from 2–18 mm in length from rostrum to telson, and even within a single species complex may range from 4–18 mm (i.e., *H. canadensis*). Three major trends of body size evolution emerged from this study: 1) influence of temperature and latitude; 2) influence of sand grain size; and 3) influence of genome size.

Previous work has established a relationship between body length and latitude in haustoriid amphipods (e.g., LeCroy 2002; Hancock & Wicksten 2018). I recapitulate these results, finding a strong correlation between body length and latitude (**Table 2.1**). Despite this correlation and the support from the mixed model of an interaction between genome size and temperature on body length, genome size is not correlated with latitude ( $p = 0.08$ ). This discrepancy may be the result of temperature variation at individual sites irrespective of latitude. This is related to the “bay versus open ocean” discussion above, in which the topography of a given sample site may influence its temperature profile. This influence helps to explain why there is a similar disparity in body length across the GoM, despite little differences in overall SST averages. However, the higher summer temperatures undoubtedly constrain body sizes in the GoM, as not even the largest sampled haustoriid approaches the size sampled at the northernmost latitudes (9 mm versus 18 mm).

Crustacean body sizes are highly plastic and often environmentally determined (Cheng & Chang 1994; Atkinson & Sibly 1997; Twombly & Tisch 2000). Hancock & Wicksten (2018) posited that variation in sand grain size at different sites may help explain body size differences within a single species range. Fine sand predominates in the western GoM and tends to be more compact, which would be difficult to move through for a large amphipod (**Table 2.1**; **Fig. 2.5b**). Alternatively, in the eastern GoM and Atlantic, coarse and medium sand is the most common. Coarse sand is loose and characterizes beaches with steep berms and strong surf. These factors together may promote larger body size to traverse the heavy quartz sand and reduce the chance of being dislodged from the sediment by wave action.

Outside of latitude, genome size was the strongest predictor of body size. Genome sizes are known to increase cell size (Gregory 2005), but this does not necessarily lead to a proportional increase in body size. Instead, the relationship between body size and genome size may be an emergent property of the previous two factors impacting body size, as well as a possible influence of selection on metabolism.



**Figure 2.5.** Transitions to brackish water. a) transition densities between brackish and open waters; b) ML estimate of ancestral sand grain size (“fine”, “medium”, “course”); c) ML estimate of ancestral habitat type (“brackish”, “open”). Green diamonds represent hypothesized transitions to warm brackish bays and lineages with genome size reductions; diamond marked with a “?” is a possible transition, though less parsimonious. Purple stars denote samples with genome-size estimates. Abbreviations follow Figure 2.4.

### Conclusions

Genome size variation is likely the result of a complex network of interactions that include cellular and metabolic processes, life-history, and the population genetic environment (Gregory 2005; Lynch 2007). I have shown that that variation can arise rapidly following shifts to new environments, whether to a more selectively permissive or constraining one. Future work on haustoriid amphipods should include more comprehensive genome size sampling across the

phylogenetic tree, especially in other species that have made the transition to brackish waters (such as *E. estuaricus*). In addition, laboratory studies examining the physiological tolerances of different haustoriids would be valuable to test the hypotheses proposed here.

## INTERLUDE

“It’s a dangerous business, Frodo, going out your door. You step onto the road, and if you don’t keep your feet, there’s no knowing where you might be swept off to.”

—J.R.R. Tolkien, *The Lord of the Rings*

“It is very easy to grow tired at collecting; the period at a low tide is about all men can endure.”

—John Steinbeck, *Log from the Sea of Cortez*

“Nothing in evolution makes sense except in light of population genetics.”

—Michael Lynch, *The Origins of Genome Architecture*

The most important book I’ve read since coming to graduate school is Michael Lynch’s *Origins of Genome Architecture*. It represents perhaps the last stage in my evolution from traditional systematics and biogeography to an interest in theory and methods in evolutionary biology. This transition began with Mariana Mateos’ course in quantitative phylogenetics and continued with a population genetics course with Charles Criscione. All of which coincided with my moving in with Dan Powell, who at the time was a senior PhD candidate in Gil Rosenthal’s lab. Dan had a house on George Bush Dr. only a few blocks from campus. Like myself, graduate school had come later in life and had eaten away his marriage.

Prior to moving in with Dan, I had made few actual connections with my fellow grad students. My life was entirely domestic when not working, and so I missed much of the first-year conversations, bar outings, and get-togethers. Two of my cohort members were gone by the second year – one decided to take a Master’s degree and the other transferred to the Galveston campus – and a third had a fellowship to work in Germany by our third year. In reflection, I deeply resent my lack of involvement my first year; my advice to all incoming graduate students has since been to get to know your cohort, participate, and enjoy their company. Confide in

them. This period in your life is so transitory; they will be gone from it before you realize you never knew them.



“Somewhere Over the Rainbow”. White sandy beach of Grayton, Florida. Photo by Thomas Strawn.

My involvement with the EEB community was entirely due to Dan’s friendship. He invited me to the bar, to gatherings, and eventually I became part of the community such that I didn’t need his invitations. My development as a scientist was also a product of my living with Dan – often soft spoken, careful with his words and Socratic in his conversations, there was always some knowledge to be imparted. This was clear to all who attended the EEB journal clubs during his tenure at A&M; he could be relied upon to provide thoughtful commentary to any article, no matter how obscure. He also, like his PI, always thought big – how does this contribute to theory, to our understanding of the universe, to the evolution of sexual systems, etc.

Unlike Gil, Dan tended to be more grounded in the practicality of said big ideas, and I think this sense of realism helped me be critical of my own pie-in-the-sky schemes.

Transitioning into theoretical work lays bare any shortcomings you may have regarding your knowledge of the literature. To develop an idea that tests fundamental model assumptions requires you to have some degree of mastery of the subject. Peer reviewers will quickly point out when you have failed at this. Despite my becoming interested in the subjects of Volume II by the second year of my PhD, these were not pursued until much later. Each new subject you attempt to break into restores the feeling of being an imposter – yeah, you understand amphipod taxonomy but who are you to question models in population genetics? Empirical work is easier in this respect; what you get is what you get. The data speak for themselves. When you’re testing or building models, it’s on you to do all the talking.

I blame Dan and Emma Lehmburg for this transition. In many ways, I was just peaking over the edge; their insistence that I could do it, their persistent faith in my abilities, I contend, was akin to a shove. I honestly don’t think I would have ever made that plunge without them. And once I started down it, I found I couldn’t go back. The literature was exciting, constantly updating and expanding, with new methods and tools coming out weekly. The field felt organic, whereas my experience with haustoriid amphipod taxonomy and systematics was mostly dredging up old forgotten books and obscure articles that likely hadn’t been read by anyone in half a century. I began to see the organisms I was working with as a means to an end – the biology and life-history of a critter determined how suitable it was to answer overarching questions I had about evolution, as opposed to the biology of the critter being the goal of the investigation itself.

In short, Volume II of this dissertation is the fault of Dan Powell, Emma Lehmberg, Michael Lynch, and some courses on genetics (and also Heath Blackmon, as mentioned in the Acknowledgements). To some readers, this transition may seem abrupt – Volume I and Volume II take dramatically different approaches to the study of evolution. But there is a common thread, and the investigations of Volume II are a result of my experience in the prior chapters.

My third year was a year of exams. I took my orals and writtens in the Fall and my proposal defense was the following Spring. These exams loom over every graduate student like a storm cloud, and no matter how much you prepare you never quite feel ready for them. One of my exam committee members described prelims as, “An academic hazing ritual”. I’m unsure if this was meant to comfort me. While I was prepared for most of my writtens, I’ll never forget the questions on community ecology from Mickey Eubanks. I had spent so much time preparing for quantitative phylogenetics and population genetics, I must have neglected some of the “big ideas” in community ecology. One of Mickey’s questions was akin to: *Provide the history of the idea “diversity begets stability”*. The phrase itself provided enough for me to try and elaborate, but the answer was mostly me attempting to string together disparate characters and subjects in the field that had little to no bearing on the actual history of the concept.

I will divulge this tidbit. In the week prior to my orals, the program coordinator at the time took me aside and told me that in a recent conversation with Mickey, he had gathered that Mickey was none-too-happy with my answers and that he was not going to go easy on any EEB students. Or something like that. Anyway, the program coordinator told me that I needed to be ready for Mickey and should really be studying my ass off. I was horrified. I must have read every major paper in community ecology in the days leading up to my orals. I spent hours rehearsing exactly what I would say, how I would try to assure Mickey that I knew the topic.

From conversations with other graduate students who had already passed their orals, one thing became clear: I needed to control the conversation.

My committee members, before we got started, chatted a bit about who would go first. I requested Mickey go first, saying, “I have a bit to clarify on my writtens”. He tilted his head a bit, folded his hands in front of him, and said, “Like what?”

I proceeded to regurgitate the entire history of community ecology. I took the opportunity to re-answer every question that he had asked in the writtens and then some. My orals took two hours and I passed unanimously.

The greatest perk of academia is the travel opportunities. I managed to go to Mexico for at least a week every year of my PhD; I saw the Pacific Ocean for the first time at San Francisco Bay at SICB in 2018; and I drove the entire Gulf and Atlantic coastlines for fieldwork. While my third year was a year of exams, it was also a year of extensive travel. Over Spring break, the Rosenthal lab, along with myself and Faith – who had moved in with Dan and I the previous year – drove down to the sleepy town of Calnali in the Sierra Madre Oriental mountains in Hidalgo where the lab has a field station. The days were spent driving to distant field sites and tossing minnow traps sprinkled with dog food into creeks and streams; I read a book of shorts from Hemingway that I had found on the shelf in the station, lounging on boulders or against trees while we waited for the swordtails to take the bait. I reread “The Snows of Kilimanjaro” for the first time since I was a teenager. At night, we’d recline on the roof of the field station overlooking the town. Concrete buildings lined either side of the gravel street, each decorated with vibrant murals courtesy of a traveling band of artists. The murals depicted scenes of coffee trees and dung beetles; abstract nature pieces of tree-like humans in embrace; a bizarre human-deer creature in a leather vest and sombrero with hideous fangs; delicate hands reaching up to

catch a beetle beneath the watchful, wrinkled eyes of an old woman. Many of these scenes were in the foreground of Cerro de la Aguja, a narrow spire that rose like a stone finger towards the heavens looming on the outskirts of the town. Faith and I would stay up late into the evening, drinking Victoria from the small shop across the street and watching the mist gather around Cerro de la Aguja; above us, the constellations reminded us of our insignificance amongst the grand scheme of the universe.

Sometimes I think I understand how the dying man in Kilimanjaro felt. So many details I have forgotten – all the laughs, the emotions, the moments that at the time had seemed so visceral. So many stories that I should have written and, never having been told, will simply cease to exist when I do.



“Last Taco Stand Before Nuevo Laredo”. Photo by Zach Hancock.

Here's one:

Dan, Faith, and I took a road-trip to Tamiahua in Veracruz at the end of the week. We had planned on visiting the same field site that I had collected amphipods at the year prior with Nuri. There's a funny thing to note about distance in Mexico. Driving, Google will inform you that it's ~120 miles from Calnali to Tamiahua; roughly a 4-hour drive.

It took us nearly 12 hours.

There had been heavy rains in the east, and we found ourselves down a narrow, flooded country road of thick muck. Flanked on either side by cow pastures or forest, we were hours from any town and none of us had cell signal. The truck struggled down the windy, muddy road, tires occasionally spinning-out and spraying the doors and bed with wet clay. We began to question if we should turn around when we came to a steep hill that dropped into a flooded curve below. Dan threw the truck into park and we climbed out, sinking up to our ankles in mud. The three of us stood at the precipice, leaned over, and stared down the slope of slick earth.

"If we go down, we may never get back up," I said.

Faith proceeded to dig the mud from the tire's tread with a stick while Dan stroked his chin, eyes going from the truck to the slope as if performing matrix algebra in his head. The sharp turn below made it impossible to see where the road went from here, and as far as we knew there was an even larger hill around the bend. We'd be sitting ducks in a valley of mud.

"I'll go see what's around the curve," I offered, gesturing down the hill. "Just wait for me. Don't go down until I come back." I yanked my shoes out of the mud, the suction making a sloppy *glop* as I freed myself. Carefully, I started down the hill, slipping a few times but managing, somehow, to not fall. I rounded the corner below and followed the road for another mile before I came to the highway. My shoes were soaked, my pants splattered with mud up to my thighs.

The mud road turned briefly into a paved overpass, but it continued into the forest on the other side. There was no way off this hell path.

By the time I made it back to the truck to warn them to turn around, to not go down the hill because there was no way to get back to the highway, they were halfway down it. The back tires were drifting wildly as Dan struggled to maintain control, the muddy path giving way under the weight of the truck; gravity was doing all the work. Upon seeing me, he managed to turn the truck sideways enough so that it sunk to a stop. I climbed up the hill, shouting, “No! No! Go back! Stop!”

It was too late. The truck was going down that hill, one way or the other. Dan later explained that he was worried something had happened to me – apparently, I’d been gone longer than I thought – and they were just going to bite the bullet and come find me. Faith claims that she stressed to him that I’d be back, to just wait, but Dan had insisted.

I climbed into the back of the truck. It only seemed fair that we all ride down this beast together. “Hold on,” Dan said, turning the wheel, putting us back in the hands of gravity. Faith took out her phone and pressed “record”. We’ve watched that video several times since. The audio is mostly just expletives.

Sometimes you just have to take the plunge.

VOLUME II  
PHYLOGENETICS IN SPACE

# CHAPTER III

## IMPACTS OF ANCESTRAL ISOLATION-BY-DISTANCE ON DIVERGENCE-TIME ESTIMATION<sup>3</sup>

### Introduction

A major goal in phylogenetic and phylogeographic studies is the estimation of species divergence times. The topic has a long and contentious history largely centered around questions of how to appropriately apply fossil calibrations (e.g., Heath et al. 2014; Brown and Smith 2018), rate heterogeneity (Pond and Muse 2005), rate of morphological evolution (Lynch 1990), and selecting an adequate clock model (Douzery et al. 2004; Lepage et al. 2007).

Beyond methodological concerns are those that emerge from the nature of the data itself. Most phylogenetic models assume that fixed differences between species are the result of genetic drift, and under the neutral theory of molecular evolution (Kimura 1968; King and Jukes 1969) the rate of evolution (or substitution rate) is equal to the per generation neutral mutation rate,  $\mu$  (Kimura 1983). For well-calibrated molecular clocks (e.g., Knowlton and Weigt 1998; Weir and Schluter 2008; Herman et al. 2018), we can estimate the time of divergence (usually in years) as  $\pi_{12} / 2\mu$ , where  $\pi_{12}$  is the pairwise sequence divergence between species 1 and 2. However, in general researchers are not interested in estimating the divergence time of specific genetic variants, but rather the time of population divergence ( $T_D$ ). For example, we might be interested

---

<sup>3</sup> Hancock, ZB & Blackmon, H. (2020) Ghosts of a structured past: Impacts of ancestral patterns of isolation-by-distance on divergence-time estimation. *Journal of Heredity*, in press.

in estimating the timing of a vicariant event that we suspect corresponds to a past geological upheaval.

There is a known discrepancy between the coalescent time of neutral genetic variants ( $T_{\text{MRCA}}$ ) and  $T_D$  (Nei and Li 1979; Nei and Takahata 1993). The degree of this discrepancy is determined by the ratio of  $T_D / N_e$ , where  $N_e$  is the effective population size (Edwards and Beerli 2000; Rosenberg and Feldman 2002). This is because lineages must first be within the same population, which occurs  $T_D$  generations in the past, followed by coalescence, which on average requires  $2N_e$  generations. Therefore, for a completely panmictic population:  $T_{\text{MRCA}} = T_D + 2N_e$ . The expected amount of pairwise sequence divergence is

$$E(\pi_{12}) = 2\mu[T_D + 2N_e] \quad (1)$$

(Wakeley 2000). When the ratio of  $T_D / N_e$  is large, the bias in coalescent time in the ancestral population is minimal compared to  $T_D$  (Edwards and Beerli 2000; Arbogast et al. 2002). However, as  $T_D / N_e$  becomes small,  $2N_e$  plays a major role in the overall sequence divergence between species. Rosenberg and Feldman (2002) evaluated the relationship between  $T_{\text{MRCA}}$  and  $T_D$  in a simple two population split model using coalescent simulations. They found that  $T_{\text{MRCA}}$  converged on  $T_D$  when the ratio of  $T_D / N_e \approx 5$ . Importantly, the  $N_e$  in these models is that of the ancestral population; therefore, the extent of overestimation is the result of demographic conditions present in the ancestor. Demographic conditions that inflate  $N_e$ , such as ancestral population structure or a bottleneck following the split, is expected to have a major impact on divergence-time estimation (Gaggiotti and Excoffier 2000; Arbogast et al. 2002; Angelis and Dos Reis 2015).

Wakeley (2000) demonstrated that in descendant species who share an ancestor whose population dynamics are characterized by an island model (Wright 1931) with free migration

between demes, overestimation of divergence-times are on the order of  $2N_eD[1 + 1/(2M)]$  where  $M = 2N_e m D / (D - 1)$ ,  $m$  is the migration rate and  $D$  is the number of demes. The expected amount of pairwise sequence divergence is therefore

$$E(\pi_{12}) = 2\mu \left[ T_D + 2N_e D \left( 1 + \frac{1}{2M} \right) \right]. \quad (2)$$

Population subdivision initially leads to shallow coalescent times where individuals within a shared deme rapidly find ancestors (the “scattering phase”; Wakeley 1998). However, since ancestral lineages must be in the same deme to coalesce, the rate in the “collecting phase” is characterized by the migration rate that shuffles ancestors around the range, reducing the probability that lineages coalesce (Wakeley 1998; 1999).

In the context of real populations, the island model of migration rarely applies (Meirmans 2012; Whitlock & McCauley 1999). Instead, population structure is the product of the spatial distribution and dispersal potential of the organism in question. Often this structure is in the form of isolation-by-distance (IBD). IBD is a widespread pattern in natural systems, characterized by a reduction in the probability of identity by descent (Wright 1943) or genetic correlation (Malécot 1968) with geographic distance. Patterns of IBD are most pronounced in stepping-stone models (Kimura 1953; Kimura and Weiss 1964) in which migration is restricted to neighboring demes. In this way, demes in close proximity share a greater proportion of migrants than they do with more distant demes. Distributions of coalescent times in stepping-stone models have been studied both in the context of one dimensional and two-dimensional models that are circular or toroidal (Maruyama 1970a; 1970b; Slatkin 1991), and in continuous models with joined ends (Maruyama 1971) or with discrete edges (Wilkins and Wakeley 2002). Slatkin (1991), using a circular stepping-stone model, showed that the probability for two genes sampled  $i$  steps apart have an average coalescent time:

$$T_{MRCA} = 2N_e D + \frac{(D-i)i}{2m} \quad (3)$$

Therefore, the amount of expected pairwise sequence divergence is:

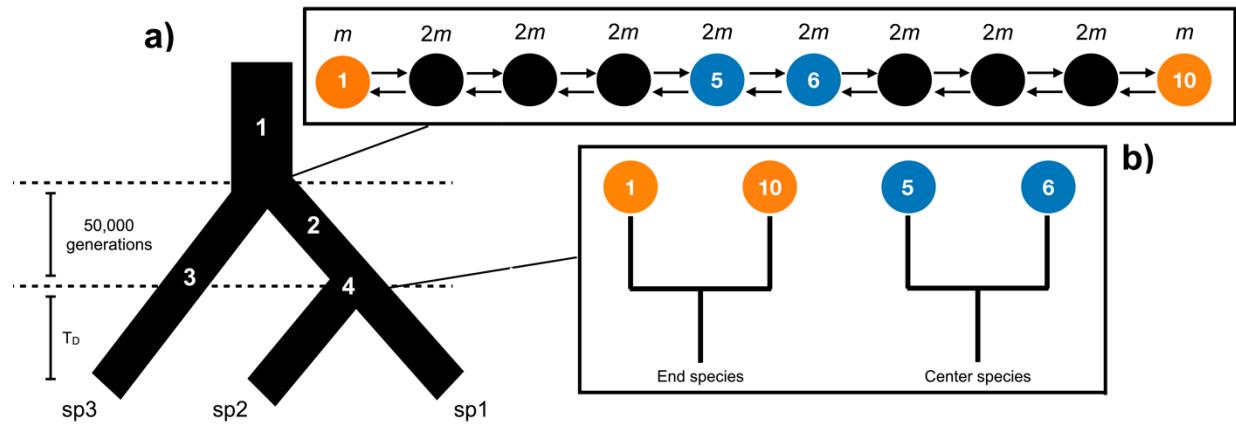
$$E(\pi_{12}) = 2\mu \left[ T_D + 2N_e D + \frac{(D-i)i}{2m} \right] \quad (4)$$

The circular stepping-stone model should overestimate  $T_D$  more dramatically as the number of demes becomes large and the distance between them increases. However, like the island model of free migration, circular ranges are likely rare in nature. Instead, natural populations are characterized by discrete range edges where end demes may only receive migrants from one direction (e.g., Peterson and Denno 1998; Broquet et al. 2006; Aguillon et al. 2017). Hey (1991) showed analytically in the case of a linear stepping-stone model that the distribution of coalescent times of two alleles from demes at the extremes of the range should coalesce much deeper than any two alleles chosen randomly from the population.

Vicariant speciation is considered one of the most common forms of allopatry (Coyne & Orr 2004), and results from the cessation of gene flow at some discrete barrier in a species range. This form of speciation has been invoked across many empirical systems (e.g., Riddle et al. 2000; Bocxael et al. 2006; Hancock et al. 2019). Vicariant speciation in organisms with low dispersal abilities may maintain strong allelic differences at the range edges ancestrally consistent with a pattern of IBD. For example, Hancock et al. (2019) found that sister species of beach amphipods in the Gulf of Mexico with large ranges showed patterns of IBD within species. In addition, they identified a distinct barrier to gene flow (the Mississippi River) and posited that this resulted in vicariant speciation. Since both IBD and vicariant speciation are presumed common in nature, biases in divergence-estimation based on  $\pi_{12}$  could be widespread.

Ultimately, the degree to which  $T_{MRCA}$  impacts phylogenetic inference and divergence-time estimation is dependent on its impact on  $\pi_{12}$ . Given that lower migration rates lead to

greater  $T_{\text{MRCA}}$  (Hey 1991), we expect that differentiation ( $\pi_{12}$ ) between end demes compared to center demes will become more pronounced at smaller  $m$ . If the difference between the  $T_{\text{MRCA}}$  of central demes and end demes is dramatic enough, we expect that divergence dating of species that arose from ancestral end demes may significantly overestimate  $T_D$ .



**Figure 3.1.** Population model for SLiM simulations. A) Three-taxon species tree: 1) coalescent simulations in *msprime* with  $N = 2000$ ; 2) ancestral stepping-stone conditions begin (see B); 3)  $N = 1000$ , panmictic; 4) population split, leaving end or center species surviving as sp1 and sp2. B) Ancestral population dynamics. Circles designated “1” and “10” are end species; center species are “5” and “6”.

In this study, I estimate mean  $T_{\text{MRCA}}$  for two genes sampled in descendant species (either from the ends or the center of the ancestral range) in which the ancestral population is characterized by a stepping-stone model with discrete ends using a simulation approach. In particular, I am interested in what value of  $T_D / ND$  I expect  $T_{\text{MRCA}}$  to converge on  $T_D$ . I use  $ND$  (the product of the census size and deme number) as our expected  $N_e$  under panmixia (Wakeley 2009). Next, I examine the distribution of  $\pi_{12}$  across the genome under different simulated migration conditions to compare with expectations under a panmictic model. I then test the performance of the phylogenetic inference program SNAPP (Bryant et al. 2012) on simulated single nucleotide polymorphism (SNP) data to evaluate how these trends may bias inference of

species divergence times. SNAPP is a BEAST (Bouckaert et al. 2014) package that operates under an explicit coalescent framework, inferring gene trees from individual SNPs. The program is ideal for phylogeographic studies that utilize RADseq and other genotype-by-sequencing (GBS) technologies to generate thousands of SNPs across the genome (e.g. Manthey et al. 2015; Dowle et al. 2017; Manthey et al. 2017; Leslie and Morin 2018), and has been used explicitly in divergence-time dating previously (Strange et al. 2018; Spalink et al. 2019; Fang et al. 2020). Finally, I illustrate how ancestral IBD can inflate divergence-time estimates on an empirical phylogenomic dataset of lizards (Domingos et al. 2017).

## Methods

In the following methods, I use the term “deme” to represent a subpopulation of randomly mating individuals within a broader collection of demes that I refer to as the “population”. To be a part of the population, a deme must be able to share migrants with other demes within the population. I use the term “species” to represent an isolated randomly mating unit that no longer shares migration with other populations or demes. This is not meant to reflect any species definition. Finally, I use the term “end species” and “center species” to refer to a set of sister species that either descend from demes on opposite ends of the ancestral range (end species) or descend from neighboring demes in the range center (center species). Our focus below is on evaluating the coalescent times, pairwise differences, and estimated divergence times between these sister species (i.e., the outgroup – “sp3” in **Fig. 3.1A** – is only meant to root the tree).

For all simulation models, I used the product of the census population size ( $N$ ) and the deme number ( $D$ ) to evaluate the relationship of  $T_D / N$  (Rosenberg and Feldman 2002). Since sp1 and sp2 (**Fig. 3.1**) transition to panmixia following the split at time  $T_D$ ,  $ND$  should approximate  $N_e$ , though there will be a period of nonequilibrium immediately following divergence. Therefore, the contribution to  $\pi_{12}$  from  $T_D$  is on the order of  $ND$ . The ancestral contribution of  $\pi_{12}$  will necessarily be some value greater than  $ND$  due to population structure (i.e.,  $N_e > ND$ ; Wakeley 2000). My interest here is explicitly on how much greater this contribution is relative to  $T_D$ .

### *Coalescent simulations*

To evaluate the relationship between  $T_D / N$  and  $(T_{\text{MRCA}} - T_D) / T_D$  when the ancestral population is characterized by a stepping-stone model of migration, I used *fastsimcoal2* (Excoffier et al. 2013) simulations over a wide range of  $T_D / N$  values. Specifically, starting at time 0 and going backwards, each simulation consisted of initially two species with no migration between them until time  $T_D$  in the past. At  $T_D$ , these two species merge into an ancestral population with 10 demes ( $D$ ) following a linear stepping-stone migration model. For simulations of end species, the ancestral deme of each species was on opposite ends of the range (i.e., demes 1 and 10 in **Fig. 3.1A**). For the center species, the ancestral demes were neighboring and in the range center. For each of these two models, I sampled  $k = 2$  individuals to coalesce, and each simulation terminated upon coalescence.

In the ancestral population, center demes received migrants from neighboring demes at rate  $2m$ , whereas demes at the end of the range received migrants at rate  $m$ . This is due to the fact

that end demes have only a single neighbor, whereas all center demes have two neighbors (**Fig. 3.1A**). The ancestral population was simulated for migration rates of 0.1, 0.01, and 0.001, and a range of  $T_D / ND$  values from 0.01–10. In addition, I simulated an island model of migration for comparison with the stepping-stone model. In the island model, the ancestral population consisted of 10 demes with free migration between each at rate  $m$ . This resulted in a total of 84 distinct simulation scenarios, and each were replicated 1,000 times. I did not explicitly model chromosomes; instead, replicates were treated as independent loci.

To statistically compare between the three models (end species sampled in stepping-stone, center species in stepping-stone, and the island model), I subset ratios of  $T_D / ND$  to values of 10, 5, 2, 1, 0.5, and 0.1. Resulting  $T_{MRCA}$  distributions for each population model were compared using a pairwise Wilcoxon test in the R platform (R Core Team 2019), as the resulting distributions were non-normal.

### *Genome simulations*

To evaluate how ancestral IBD impacts pairwise sequence divergence ( $\pi_{12}$ ), genome-wide coalescent times ( $T_{MRCA}$ ), and divergence-time estimation, I performed hybrid simulations that combined the coalescent simulator *msprime* (Kelleher et al. 2016) and the forward-time simulator SLiM v3.3 (Haller and Messer 2019). Since forward-time simulators begin with individuals that are completely unrelated, often a neutral burn-in period is required to allow coalescence or mutation-drift equilibrium to occur (Haller et al. 2019). This can be computationally costly and time consuming; however, using tree-sequence recording methods in SLiM (Haller et al. 2019) I can bypass the need to equilibrate during the forward-time

simulation. To generate a panmictic ancestral population with a coalescent history, I simulated 2000 individuals ( $N_e = 4000$ ) using *msprime* with genome sizes of 10 Mb and a recombination rate of  $10^{-8}$  ( $\sim 0.1$  recombination events per individual per generation). The resulting coalescent trees were then imported into SLiM as the basis for the starting population.

In SLiM, the initial population was split into two populations of  $N = 1000$ : 1) an outgroup that remained panmictic (“sp3” in **Fig. 3.1A**) and 2) the ancestral population of the sister species “sp1” and “sp2”, which was subdivided into 10 demes ( $N = 100$  per deme) in a linear stepping-stone model (**Fig. 3.1A**). These dynamics persisted for 50,000 generations after which the ancestral population was split into either end species or center species (see **Fig. 3.1A**). This was done by removing the intermediate demes and instantaneously adjusting  $N$  for the two species to 1000 so that the census size remained constant. The resulting 3-species were then allowed to evolve for  $T_D$  generations before the simulation was terminated. Five different  $T_D$  values were simulated which correspond to  $T_D / ND$  ratios of 50, 25, 10, 5, and 1 ( $T_D$  values of 50,000, 25,000, 10,000, 5,000, and 1,000 generations). These values of  $T_D / ND$  were chosen based on the results from the coalescent simulations in *fastsimcoal2* (see **Results**); for values  $> 10$ ,  $T_{MRCA}$  is expected to converge on  $T_D$ , whereas values  $< 10$  are expected to overestimate  $T_D$  regardless of migration rates.

The resulting tree-sequences from the SLiM simulation were imported into Python 3 using *pyslim*, and I overlaid neutral mutations ( $\mu = 10^{-7}$  per base per generation) onto the trees using *msprime*. Pairwise divergence ( $\pi_{12}$ ) was then estimated across the genome in windows of 100 kb for both end demes and center demes. These values were also converted into generations using  $\pi_{12} / 2\mu$ , which gives a rough estimate of divergence time per window.

Equation 1 primes our expectation for the amount of sequence divergence expected given some value of  $T_D$  and ancestral  $N_e$ . By rearranging equation 1, I can naively calculate the ancestral  $N_e$  from genome-wide  $\pi_{12}$  as:

$$N_e = \frac{\pi_{12} - 2T_D\mu}{4\mu} \quad (5)$$

From this, I plot estimated ancestral  $N_e$  within 100 kb windows across the genome to compare with the known census population size ( $N_c = 1000$ ), and to evaluate the relationship between  $N_e$  and  $N_c$  in the presence of IBD.

Next, I plotted the distribution of coalescent times ( $T_{MRCA}$ ) across the genome to visualize differences between  $T_{MRCA}$  of end and center species. Median  $T_{MRCA}$  for each ratio and migration rate was compared via a Kruskal-Wallis test and a pairwise Wilcoxon rank test in R due to the data violating normality.

Each simulation produced >200,000 SNPs. For divergence-time analysis, I randomly sampled 3000 SNPs—a number found by Strange et al. (2018) to optimally perform in SNAPP. Each run consisted of 10 individuals from species sp1 and sp2, and 1 individual from the outgroup population, sp3 (Fig. 3.1). Unlike other fully coalescent models, SNAPP does not sample from gene trees directly to estimate the species tree, but instead integrates over all possible gene trees using biallelic SNPs. The method has been found previously to perform well on both simulated and empirical data (Bryant et al. 2012; Strange et al. 2018). I designated a gamma-distributed prior on  $\theta (=4N_e\mu)$  with a mean equal to the expected  $\pi_{12}$  (equation 1). Forward ( $u$ ) and backward ( $v$ ) mutation rates were estimated within BEAUTi (Bouckaert et al. 2014) from the empirical SNP matrix using the tab *Calc\_mutation\_rates*, and these values were sampled during the MCMC. The rate parameter  $\lambda$ , which is the birth-rate on the Yule tree prior,

was gamma-distributed with  $\alpha = 2$  and  $\beta = 200$ , where the mean is  $\alpha / \beta$  (Leaché and Bouckaert 2018).

SNAPP is designed to handle incomplete lineage sorting (ILS), but to minimize its effects—since I am not interested in the program’s ability to estimate topology but rather branch-lengths—I applied a fixed species tree. Branch-lengths in SNAPP do not scale to time, but instead are measured in number of substitutions. Given a fixed mutation rate, I converted the number of substitutions separating sp1 and sp2 to the number of generations as  $g = s / \mu$ , where  $s$  is branch-lengths in units of substitutions (Bouckaert and Bryant 2015). The MCMC chain length was 10–50 million sampling every 1000 with a burn-in of 10%, ensuring that ESS values of interest were all  $>200$ . Runs were performed on the high-performance computing cluster CIPRES ([www.phylo.org](http://www.phylo.org); Miller et al. 2010).

MCMC log files were then downloaded and analyzed in R. The performance of SNAPP was evaluated by comparing traces of end and center species across migration rates and  $T_D / ND$  values. Results were evaluated by first randomly sampling 1000 rows for each migration rate, and then performing a Kruskal-Wallis test. Trees from the MCMC were summarized in TreeAnnotator v.2.6.0 (Bouckaert et al. 2014) and visualized in R using the package *ggtree* (Yu et al. 2017). Branch colors were scaled by estimated median  $\theta$  per branch.

To ensure the trends observed were the result of inflated  $\pi_{12}$  when  $T_D / ND$  and migration rate is low and not an issue unique to SNAPP, I performed pairwise  $F_{ST}$  tests of end and center demes that were used in the divergence-time estimation (see *Appendix B*; **Fig. S3.10**). These tests were performed using *tskit*. All SLiM recipes, python and R code, and xml files can be found at <https://github.com/hancockzb/ancestralIBD>.

### *Empirical dataset*

To evaluate how ancestral patterns of IBD may impact divergence-time estimation in practice, I analyzed the phylogenomic dataset of lizards from Domingos et al. (2017). The dataset consists of 12 species (including the identified cryptic lineages) sampled broadly across the geographic range of *Tropidurus itambere*, which is native to the Cerrado, a tropical savanna in Brazil that stretches across the states of Goiás, Mato Grosso do Sul, Mato Grosso, Tocantins, Minas Gerais, and the Federal District. Domingos et al. (2017), using anchored hybrid enrichment, identified five cryptic species they designated A–E within *T. itambere*. Each of these species were sampled across multiple localities with some locations spatially nearer to their close relatives than others (see Fig. 1 in Domingos et al. 2017).

This geographically broad sample scheme is ideal to test the impact of IBD on divergence-time estimation. If ancestral IBD has influenced  $\pi_{12}$ , we expect species localities more distant to one another to be more deeply diverged than when two species are sampled from locations nearby. Importantly, this pattern should only hold if the range was once continuous (as in my simulations above); otherwise, there should be no difference in  $\pi_{12}$  between species even if there are current patterns of IBD within species. Ongoing gene flow between the spatially close localities could also generate this pattern, but Domingos et al. (2017) found no evidence for this.

The alignments used in the coalescent species delimitation in Domingos et al. (2017) were downloaded from <https://datadryad.org/stash/dataset/doi:10.5061/dryad.1hs2m>. I randomly selected 10 loci conditional on them containing samples from all locations (total length = 15,303 bp). I generated two separate alignment files based on the sampled locality's distance from their nearest relative. Preliminary analyses including the entire dataset showed that species (DE) were

sister to A, and B was sister to C (which was also found from the entire concatenated dataset in Domingos et al. 2017). Therefore, the first dataset I designated as “far”, and it included only the location of E that was most distant from A (i.e., São João D’Aliança), and the location of C that was the most distant to B (i.e., Ribeirão Cascalheira). The second dataset was designated “near” which included localities of E geographically closest to A (Brasilia and Pirenópolis) and C to B (Barra do Garças).

Phylogenetic inference was performed using the fully coalescent software \*BEAST (Heled & Drummond 2010). For each locus, I applied the HKY model (Hasegawa et al. 1985) and a strict molecular clock. I also performed runs with a relaxed lognormal clock for each locus to ensure age differences across the tree were not the result of restricting substitution rate variation. Domingos et al. (2017) only focus on topology and time is not considered; therefore, I estimate branch-lengths in units of substitutions per site. I use an arbitrary rate of mutation ( $10^{-8}$ ) to convert substitutions to generations. This analysis is not meant to be a rigorous evaluation of the true time of divergence, but merely a demonstration of the impacts of ancestral IBD on empirical data. Each analysis was run for 100 million generations with a burn-in of 10%. Estimated divergence times between the two models were compared using a two-way ANOVA in R. Differences in the densities of estimated gene trees from the posterior were visualized using DensiTree (Bouckaert 2010).

## Results

### *Coalescent simulation results*

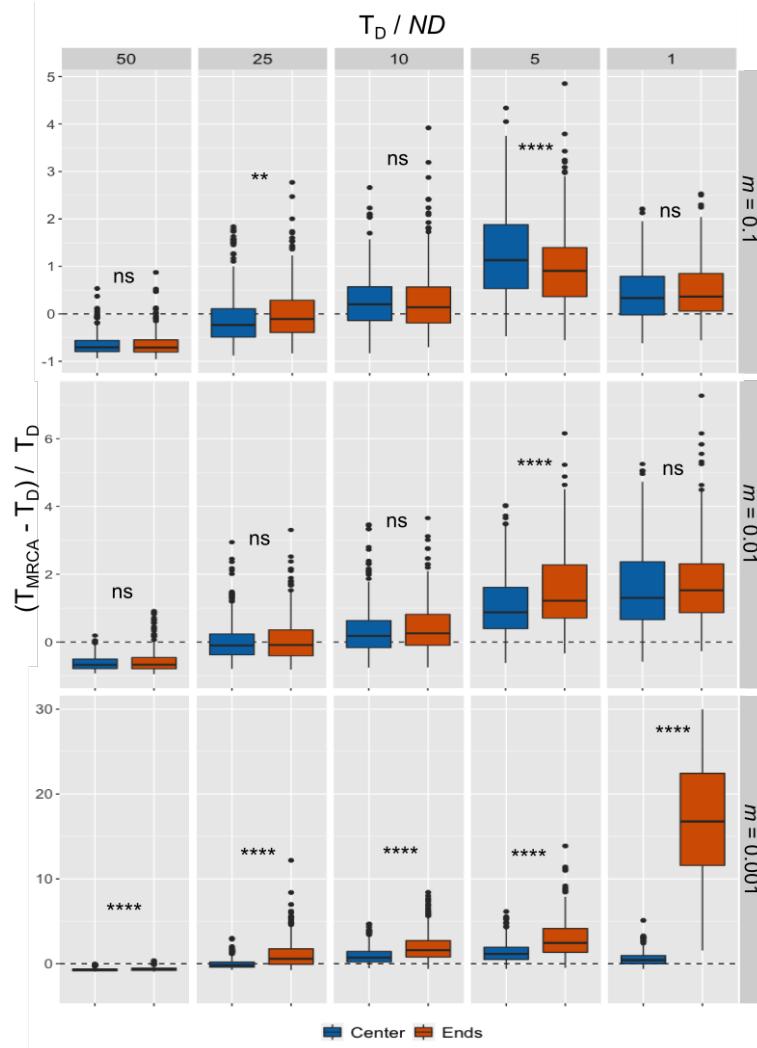
The coalescent simulations produced trends superficially similar to those found by Rosenberg and Feldman (2002). At the lowest  $T_D / ND$ , the proportion of deep coalescence was dramatically greater than at higher values with the curve producing a similar logarithmic relationship (**Fig. S3.1**). However,  $T_D$  and  $T_{MRCA}$  did not necessarily converge when  $T_D / ND = 5$ . Instead, the rate of convergence was dependent on both the deme sampled and the migration rate.

When migration was high ( $m = 0.1$ ) and  $T_D / ND$  was less than 0.5, there was no significant difference between center or end species in the stepping-stone model or the island model. However, for values of  $T_D / ND > 0.5$ , the  $T_{MRCA}$  of end species became significantly different from both island ( $p < 0.02$ ) and center species ( $p < 0.01$ ; see **Table S3.1**). When migration was reduced below 0.1, this pattern became more extreme. End species were significantly different in all pairwise comparisons of models ( $p < 0.000001$ ), and center species differed from the island model at  $T_D / ND$  ratios of 0.5, 2, and 10 ( $p < 0.03$ ) when  $m = 0.01$ . At the lowest migration rate simulated ( $m = 0.001$ ), all pairwise model comparisons were significantly different when  $T_D / ND > 0.5$  ( $p < 0.001$ ; see **Table S3.1**).

#### *Genome simulation results*

Results from the genome simulation approach corroborated those found with *fastsimcoal2*. Regardless of  $T_D / ND$ , when  $m = 0.1$  the difference between center and end species was less severe relative to when  $m < 0.1$  (**Table S3.2**). Across the simulated genomes,  $T_{MRCA}$  became dramatically deeper between end than center species as migration fell below 0.01. For the genome-wide divergence estimates, the degree of overestimation depended on the ratio

of  $T_D / ND$ . While all scenarios where  $m = 0.001$  overestimated the true  $T_D$ , when  $T_D / ND < 10$  end species were 5–60 times more diverged than expected (Fig. S3). This is a direct result of the deeper coalescent times between end species when  $m < 0.1$ , as these longer branches provide more time for mutations to occur and accumulate (Fig. S3.2).



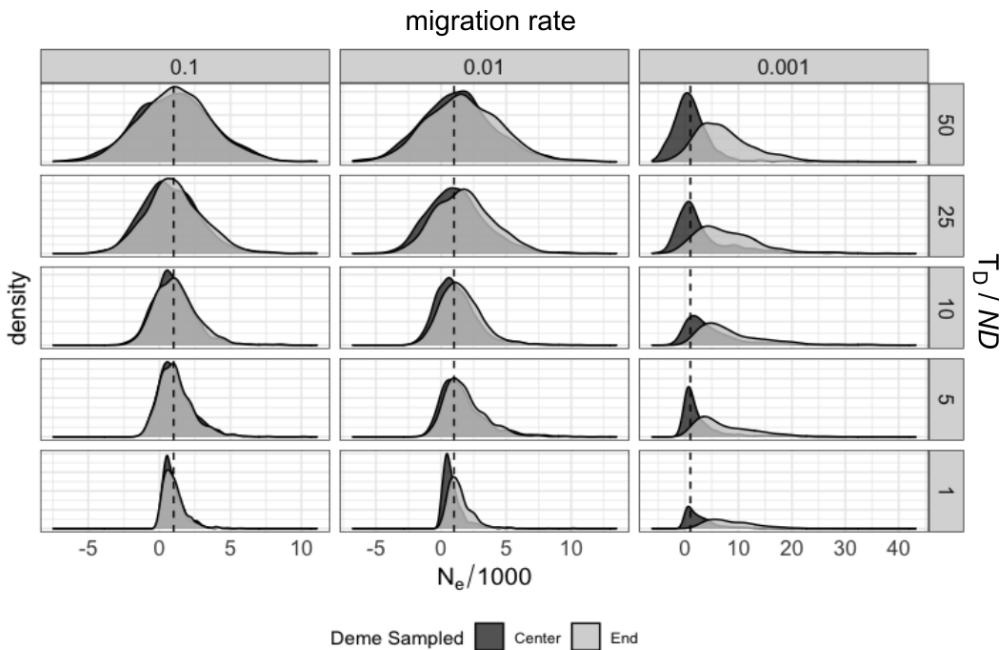
**Figure 3.2.** Box plots of the estimated  $T_{MRCA}$  by SNAPP; ns = “not significant”,  $p < 0.05$  (\*),  $p < 0.001$  (\*\*),  $p < 0.0001$  (\*\*\*),  $p < 0.00001$  (\*\*\*\*). Dashed lines represent when the estimated age converges on the true age (i.e., at 0). Note that the y-axis is different between the panels. Center species are on the left, end species on the right.

Genome-wide coalescent times ( $T_{MRCA}$ ) are shown in Fig. S3.2. When  $m = 0.1$ , only  $T_D / ND = 25$  and 10 were significantly different between end and center species ( $p < 0.005$ ).

Regardless of  $T_D / ND$ , the variance in  $T_{MRCA}$  steadily increased with decreasing  $m$ . Indeed, the increase in mean  $T_{MRCA}$  when  $m = 0.001$  appears largely driven by an increase in the variance at this lower rate. Due to this, I find that ancestral  $N_e$  dramatically exceeds  $N_c$  when  $m = 0.001$  (**Fig. 3.3**).

Despite the potential for divergence-time overestimation to be extreme, SNAPP was relatively resilient when  $T_D / ND > 10$  and when  $m > 0.001$ . When  $T_D / ND = 50$ , SNAPP was overly conservative and underestimated the number of substitutions expected to occur (**Fig. 3.2**). When  $T_D / ND = 25$ , the mean estimate of both center and end species when  $m > 0.001$  either underestimated the true age or was within 5%. However, for end species where  $m = 0.001$  the estimated divergence time exceeded the true age by ~80% (**Table S3.3**). A similar trend occurred when  $T_D / ND = 10$  and 5. Here, both center and end species overestimated the true age, but the end species did so more dramatically (138% the true age versus 81% for 10; 184% versus 67% for 5). The most dramatic overestimation occurred between end species when  $T_D / ND = 1$  at ~700% the true age. Importantly, this was not merely the result of a low  $T_D / ND$  ratio, as the other migration regimes performed well. In fact, most were closer to the true  $T_D$  than the expected  $\pi_{12}$  accounting for  $2N$  (**Table S3.3**).

Estimated  $\theta$  for each branch is shown in Fig. 4 for  $T_D / ND = 10$ , and in Figs. S4–S7 for the remaining ratios. For all  $T_D / ND$  values except 1, the median ancestral  $\theta$  was higher for end species than center when  $m = 0.001$ , and the estimated  $\theta$  for the descendant species (sp1 and sp2 in **Fig. 3.1**) was considerably lower than for the ancestor or the outgroup, sp3 (**Fig. 3.4**; Figs. S3.4–S3.7). These patterns are consistent with a population bottleneck, despite  $N$  being maintained throughout the simulation.



**Figure 3.3.** Density plot of scaled ancestral  $N_e$  (/1000) based on mean  $\pi_{12}$  across genomic windows of 100kb. Dashed line is when  $N_e / N_c = 1$ .

### Empirical results

The estimated divergence-time for the clades ((DE)A) and (BC) were significantly older when samples were from geographically distant localities as opposed to those nearby ( $p < 0.00001$ ; **Fig. 3.4c**). For the ((DE)A) clade, the “far” dataset inferred an age of divergence of 270,000 generations, which was 40,000 generations (or ~14%) higher than the “near” estimate. The (BC) divergence was even more extreme, with the “far” being ~24% older than the “near” (**Fig. 3.4c**). Interestingly, despite the fact that all other samples were included in both analyses, the “far” dataset estimated older ages for most of the other nodes in the tree as well (**Fig. 3.4**). The total tree height of the “far” was 220,000 generations deeper in time than the “close” (or ~

9%). The relaxed clock estimates were more extreme, with all node heights being higher in the “far” versus “near” datasets (**Fig. S3.11**).

## Discussion

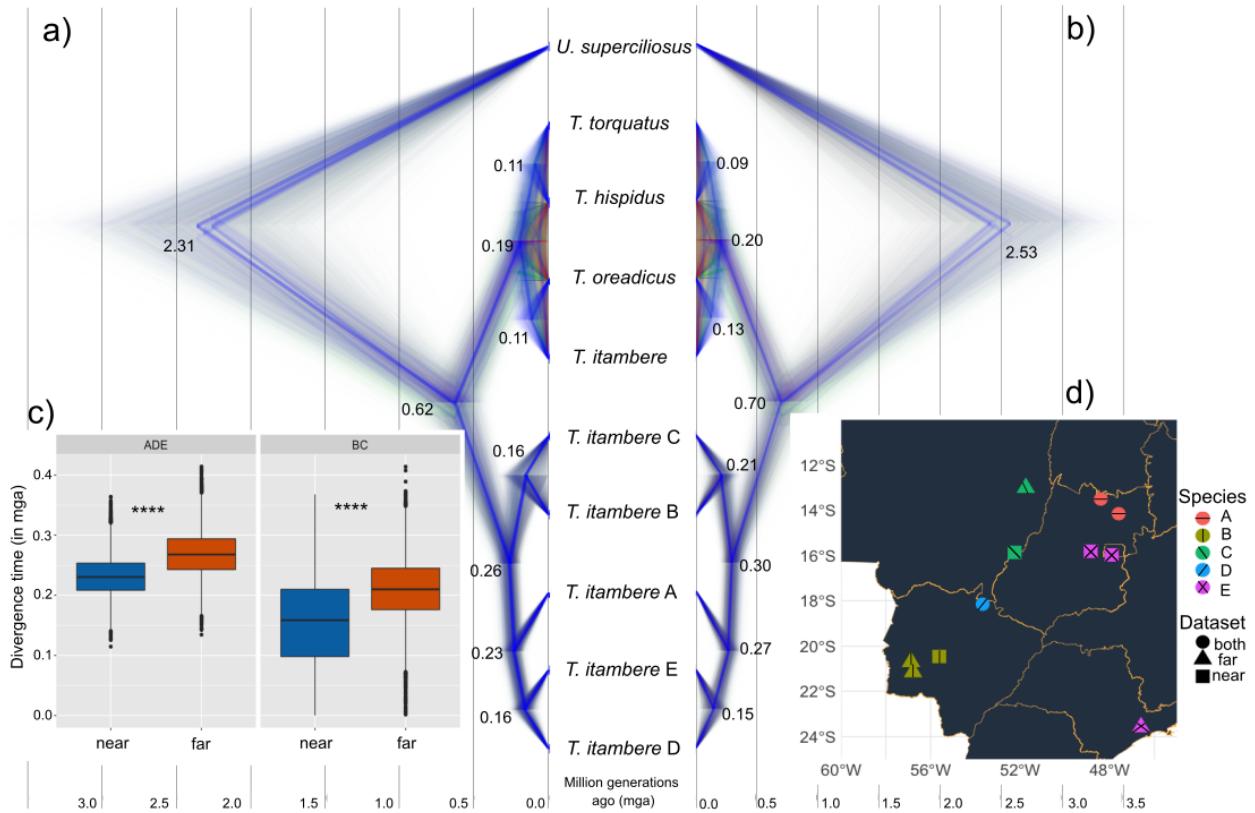
Macroevolutionary patterns are ultimately governed by microevolutionary processes (Li et al. 2018), an observation Lynch (2007), extending Dobzhansky’s (1973) maxim, summed up as “nothing in evolution makes sense except in light of population genetics”. In this light, I have demonstrated that the population genetic environment of the ancestor shapes the genetic landscape of descendant species. This has been known to impact tree topology when ILS is common (Kubatko and Degnan 2007) and overestimate divergence times in the presence of population structure caused by an island model of migration (Edwards and Beerli 2000; Wakeley 2000). Extensive prior work has shown that the stepping-stone model of migration reduces genetic correlation between demes (Kimura and Weiss 1964; Maruyama 1970a) and that demes farther apart should coalesce deeper in time than those geographically closer (Slatkin 1991; Hey 1991). However, to my knowledge, the impact of ancestral IBD has not been evaluated in the context of divergence-time estimation previously.

Rosenberg & Feldman (2002) found previously that when  $T_D / N = 5$ ,  $T_{MRCA}$  and  $T_D$  largely converged in a simple population split model. However, when in the presence of ancestral IBD I found that convergence was dependent on the migration rate (i.e., the strength of ancestral IBD) and whether surviving species neighbored each other or were at the range ends in the ancestral population.

When  $T_D / ND > 10$ , the ancestral dynamics contribute little to the divergence-time estimate differences between center and end species. However, as this ratio decreases the

contribution of  $2N_e$  to overall sequence divergence becomes non-trivial. The probability that genetic variants share an ancestor just prior to the population split is higher between species that were geographically closer than those more distant in the ancestral population. This is mediated by the migration rate, which, when high enough, can largely erase the differences between center and end demes. When migration is high (10%, or  $m = 0.1$ ), individuals move well between demes and the coalescent times largely converge (though deeper in time depending on the ratio of  $T_D / ND$ ). However, as  $m$  falls below 1% ( $m = 0.01$ ), or less than one migrant per generation being shared between demes, dispersal cannot keep up with genetic differentiation. Despite all migration regimes producing similar patterns of IBD (**Fig. S3.9**),  $F_{ST}$  becomes dramatically higher as migration drops below 1%. This differentiation in the ancestor contributes to the overall sequence divergence ( $\pi_{12}$ ) between species, which drives an overestimation of the time of the population split ( $T_D$ ) when surviving lineages descend from demes on the opposite ends of the range.

As expected, ancestral IBD skews  $\pi_{12}$  and  $T_{MRCA}$  away from expected values in a panmictic population, and this caused an inflation in  $N_e$  relative to  $N_c$ . For  $T_D / ND = 50$  and  $m = 0.1$ , the mean  $\pi_{12}$  for end species was 0.010459 and 0.010419 for center species. Using equation 5, ancestral  $N_e$  was 1147.5 for end species and 1047.5 for central. However, when  $m = 0.001$ ,  $\pi_{12}$  for end species was 0.012948, an ancestral  $N_e = 7370$ . Center species, on the other hand, only increased to  $N_e = 1255$ . As with the coalescent times, at lower migration rates the variance in  $N_e$  becomes exceedingly large, driving up the mean. Importantly, mean genome-wide  $N_e$  always exceeds  $N_c$  in the presence of ancestral IBD at a level dictated by the migration rate.



**Figure 3.4.** Estimated divergence-times for the Domingos et al. (2017) dataset. A) the “near” trees from the posterior distribution; B) the “far” trees; C) boxplot of the two nodes of focus, where ADE is for clade ((DE)A) and BC is (BC); D) map of sample sites from Domingos et al. (2017). Significance is as in Fig. 3.2. Numbers at nodes represent the median height in units of millions of generations ago (mga).

This feature of ancestral IBD has important consequences for conservation genetics.

Many studies use  $N_e$  as a rough biological measure of population size (Turner et al. 2002; Rieman and Allendorf, 2011; Hare et al. 2011), and therefore a metric of the health of a population. However, a common phenomenon in range contractions is fragmentation and isolation (Ceballos et al. 2017), which may result in IBD. If many of the demes once contributing to the connectivity of the population have become extinct, and  $N_e$  is estimated based on the surviving demes, it will overestimate the actual number of individuals within the population (i.e., the census size,  $N_c$ ). Thus, we might incorrectly conclude that a species has a larger population size than it actually does, which may lead to mismanagement.

Since  $N_e$  is inflated in the ancestral lineage, the descendant species appear to pass through a bottleneck despite  $N$  remaining constant (**Figs. S3.4–S3.8**). Estimated  $\theta$  in SNAPP captured this dynamic with more extreme differences in  $\theta$  (i.e., more dramatic bottlenecks) being inferred between end species and when  $m = 0.001$ . Population bottlenecks have been found to cause divergence-time overestimation due to random differential survival of ancestral alleles into the descendant species (Gaggiotti and Excoffier 2000). In the presence of IBD, this differential allelic persistence between species is mimicking a bottleneck—when demes were far apart this pattern is more extreme as they already maintain different allelic patterns ancestrally. However, because this pattern is recognizable (**Figs. S3.4–S3.8**) it can be used to signal when ancestral IBD may be impacting our divergence-time estimation. Unfortunately, without prior range-size knowledge it may be impossible to differentiate between ancestral IBD and a bottleneck since these produce virtually identical genetic patterns. However, it may not be necessary to do so for simple divergence estimates.

Demes need not necessarily go extinct for ancestral IBD to still influence  $\pi_{12}$ , as seen from the results of the empirical dataset; however, the persistence of demes into the present allows for a geographically aware sampling scheme. Since Domingos et al. (2017) sampled broadly across the range of *T. itambere*, they would be well-positioned to identify inflated  $\pi_{12}$  resulting from ancestral structure. For example, as I have done here, by subsetting the dataset by geographic proximity one can explicitly test for ancestral IBD. For ancestrally structured populations, geography should dictate the degree of  $\pi_{12}$ . Importantly, this also requires that species diverged via vicariance (Coyne & Orr 2004) – the splitting of a once larger range by a discrete barrier – and not some other means, such as population expansions following divergence from a more restricted habitat.

The broader impact of ancestral IBD on divergence-time estimation when in the context of large phylogenies is beyond the scope of this work, but it is conceivable that the longer than expected branches between sister species might bias rate estimation (Aris-Brosou and Excoffier 1996; Magallón 2010). In the case of ancestral IBD, the inflated  $N_e$  is mimicking a pattern of substitution rate increase. Under neutrality, the rate of substitution is equal to the per generation mutation rate,  $\mu$  (Kimura 1983); however, in the presence of population structure, substitutions may occur in the ancestral lineages between demes separated by large geographic distances. If the true age of the sister taxa is known but ancestral structure is not accounted for, the substitution rate will be upwardly biased. I found some evidence for this in the *T. itambere* dataset, in which I found higher estimates of  $\pi_{12}$  in the “far” versus the “close” dataset even for nodes more distantly related to the focal clades (**Fig. 3.4**).

Ancestral structured populations leave their imprint on descendent species in the form of greater coalescent times, and therefore larger than expected pairwise divergences between species. Further, these patterns cause inflated  $N_e$  relative to census sizes. Since ancestral IBD mimics the signature of a population bottleneck, coalescent methods that co-estimate  $\theta$  along with the topology and  $\pi_{12}$ , such as SNAPP and \*BEAST, may be the best suited to reveal this potential source of bias. However, fully coalescent models such as these are infamously computationally costly and not presently used for whole-genome sequence data or for phylogenies with large numbers of tips. Indeed, SNAPP becomes prohibitively slow when the number of tips is ~30 (Leaché and Bouckaert, 2018).

In the context of larger phylogenies or organisms in which little is known about their ancestral range, it may be impossible to know if extant species descend from range centers or ends, or the level of IBD present in the ancestor. The genetic consequences of ancestral structure

therefore behave much like “ghost” populations (Slatkin 2005); despite being extinct, their influence haunts our ability to adequately assess the phylogenetic history of their descendants.

# CHAPTER IV

## PHYLOGENETICS IN SPACE: HOW CONTINUOUS SPATIAL STRUCTURE IMPACTS TREE INFERENCE

### Introduction

Humans have an innate tendency to discretize nature. Taxonomy epitomizes this instinct: it searches for distinct characters, markers, or behaviors that differentiate organisms. Binomial nomenclature, the dichotomous key, all characterizes our desire to categorize biology. These categorizations make superficial sense; a cursory examination of biological diversity appears to support distinct species boundaries (Coyne and Orr 2004). Dawkins and Wong (2016) referred to this human urge as the *tyranny of the discontinuous mind* – a mental state of incongruity at the continuous nature of biology. As alluded to above, an example of this is the “speciation continuum”, which has received increased attention in recent years (e.g., Etienne and Rosindell 2012; Li et al. 2017). For example, Etienne et al. (2014) introduced the “protracted speciation model” that incorporates probabilities of lineage splitting, extinction, merging, and completion. Li et al. (2017) use the protracted speciation model to demonstrate how the microevolutionary processes of lineage extinction and splitting can drive macroevolutionary patterns of species diversity along latitudinal gradients.

In population genetics, most models assume some level of discretization out of mathematical convenience. Discrete population models have no spatial component and are composed of randomly mating individuals. These discrete populations are the foundation of the classical Wright-Fisher (Wright 1931; Fisher 1923) and Moran (Moran 1958) models. Even

when models subdivide the population into a series of demes, such as in the stepping-stone model (Kimura and Weiss 1964), each deme still consists of randomly mating individuals. Once a migrant finds their way into a deme, regardless of their deme of origin, they have the same probability of finding a mate as any other deme constituent.

Phylogenetic inference using the multispecies coalescent model likewise assumes individuals exist independent of space. There are only two dimensions to the model: the width, dictated by the effective size of the population ( $N_e$ ), and the length, the temporal aspect (**Fig. 4.1b**). Bifurcations in the tree implicitly assume that the descendant  $N_e$  is immediately in equilibrium and can be represented by a single value ( $N$ ,  $\tau$ , or  $\theta$ ). In reality,  $N_e$  progresses through a period of nonequilibrium the duration of which is dictated by the degree of structure present in the ancestral population (Poissant et al. 2005; Pearse et al. 2006). Furthermore, these divergence events are assumed to occur evenly – lineages are equally likely to end up in either of the two descendant species.

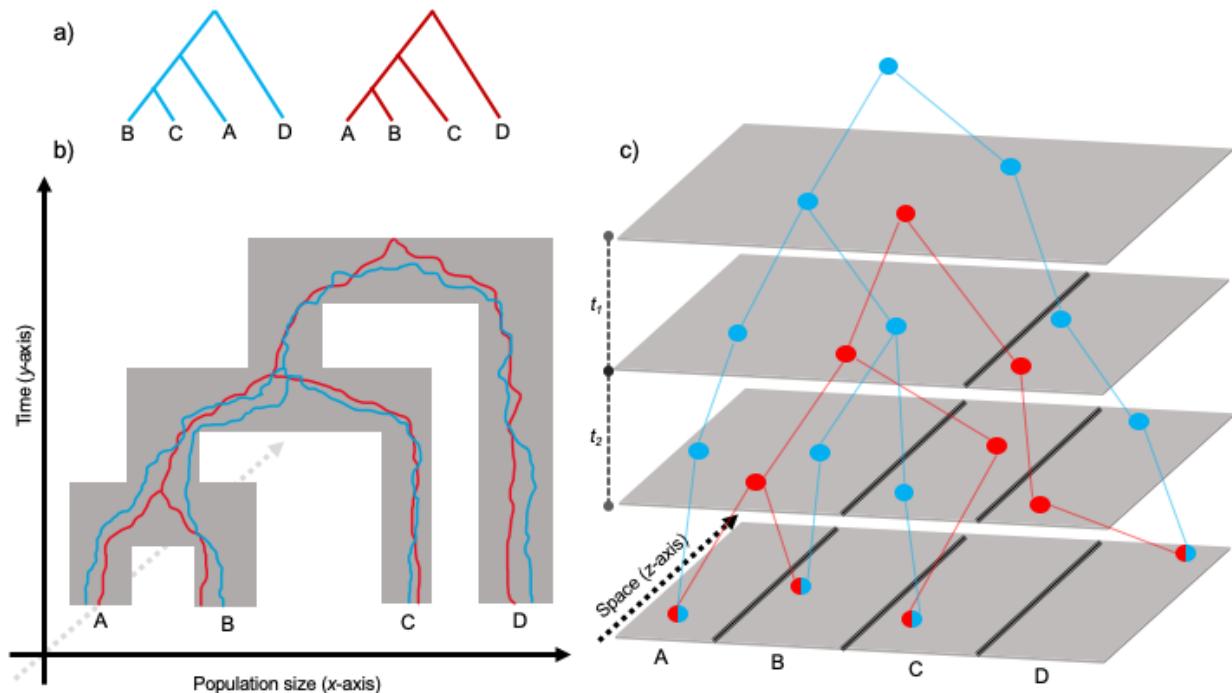
The space-independence or “missing z-axis” typical of the models described contribute to their poor performance when populations are actually continuous (**Fig. 4.1**; Bradburd and Ralph 2019). Battey et al. (2020) found that many summary statistics (such as  $\theta_\pi$ , Tajima’s  $D$ , and  $F_{IS}$ ) behaved significantly differently in continuous populations versus a random mating population of equal size. While some of the trends were captured in stepping-stone models, strange artifacts emerged; for example, observed heterozygosity and Tajima’s  $D$  were far higher at low neighborhood sizes in the stepping-stone models (neighborhood size =  $4\pi\rho\sigma^2$ , where  $\pi$  is the mathematical constant,  $\rho$  is the population density, and  $\sigma$  is the dispersal distance; Wright 1943). These may be a feature of both discretizing the habitat and the sample size approaching the local  $N_e$  when deme size is small (Battey et al. 2020). Wright (1943) noted that for continuous

populations to be indistinguishable from panmixia, neighborhood sizes must exceed 1000. This neighborhood size is likely rarely achieved in nature. For example, Jasper et al. (2019) estimated the neighborhood size of the mosquito *Anopheles aegypti* to be only 268. Clearly, space matters for the inference of population genetic data. But what about phylogenetics?

Phylogenetic inference rarely incorporates spatial structure as an explanation for gene tree discordance or node height variance. Many phylogeographic studies incorporate population genetic analyses, including analyzing population structure, but this tends to be a separate analysis that doesn't directly influence the phylogenetic interpretation (e.g., Waldrop et al. 2016; Domingos et al. 2017; Takada et al. 2018). Instead, most of the variance in tree topology is attributed to incomplete lineage sorting (ILS) or hybridization (Cui et al. 2013; Árnason and Halldórsdóttir 2019; MacGuigan and Neer 2019) and branch-length variation to the action of natural selection or substitution rate heterogeneity (which may be influenced by population structure, though this is rarely directly invoked; Bromham et al. 2018). This is not due to a lack of theoretical or empirical work that demonstrates the importance of spatial structure (e.g., Wakeley 2000; Slatkin and Pollack 2009; DeGiorgio and Rosenberg 2016). In phylogenetics, the issue may instead be a kind of *tyranny of discontinuous timescales* – that is, the mental state that posits the processes of microevolution act on such different timescales that their influence on macroevolution should be negligible.

Most simulation studies on ILS or species tree inference begin by simulating a species tree under either a pure birth model or a birth-death model, followed by coalescent simulation of gene trees (e.g., Maddison and Knowles 2006; Leaché and Rannala 2011; Fujisawa et al. 2016). As mentioned above, this method assumes panmictic ancestral populations, and therefore all genetic lineages have equal probability of being split into descendent species (i.e., lineages exist

independent of space). This space-independence simplifies ranges by failing to capture how realistic landscape dynamics can shape lineage diversity (Bradburd and Ralph 2019). For example, Wilkins and Wakeley (2002) found that coalescent times in a continuous space population depended not only on the distance between samples, but also the distance the samples were from the range center. Lineages within the range center coalesced deeper in time than peripheral lineages separated by the same distance. This feature of ranges – higher diversity in the center relative to the periphery – is not merely a feature of population genetic models but has been discussed for decades in the rangeland and conservation biology literature (e.g., Hengeveld and Haeck 1982; Lomilino and Channell 1995; Donald and Greenwood 2001).



**Figure 4.1.** Phylogenetics in space. A) Gene tree topologies; B) the  $x$ - and  $y$ -axis of MSC phylogenetic inference, with the  $z$ -axis missing; C) making sense of gene tree incongruence by considering the  $z$ -axis.

The specific mode of speciation may contribute to gene tree asymmetry and node height variance as well. Allopatric speciation involves a split following a range expansion into a previously uninhabited territory or some form of range fragmentation. Coyne and Orr (2004) define these two modes of allopatry, originally identified by Mayr (1954), as *peripatric speciation* and *vicariant speciation*, but note that their main difference is in the size of the resulting populations. Each of these modes impact gene trees differently. For example, peripatric speciation is biased towards lineages present in the periphery of the population. Since these lineages tend to be less diverse, this may intensify the founder effect as the population expands into a new territory, especially if dispersal is low (Coyne and Orr 2004; Malay and Paulay 2010; Castellanos-Morales et al. 2016). In a model of vicariant speciation, an initially large population characterized by patterns of isolation-by-distance (IBD) is cut-off from its neighbors. This may lead to species with ancestors historically on opposite ends of the range to be more deeply diverged than expected in a panmictic model (Hancock and Blackmon 2020).

I contend that discounting the spatial component of populations (the *z*-axis; **Fig. 4.1b**) can lead to spurious conclusions about both the underlying relationships of species (topology) and their history of divergence (node heights). In this study, I review several key findings in the literature and use continuous population models to illustrate the importance of spatial structure in three aspects of phylogenetics: 1) gene tree asymmetry and inference of hybridization; 2) species tree inference and divergence estimation; and 3) species delimitation. Finally, I propose how these findings can aid in the identification of spatial structure in a dataset, which, I hope, will promote a more thorough investigation of phylogenetics in an explicitly spatial context.

## Trees in Space: Model

To illustrate the impacts of spatial structure on phylogenetic inference, I constructed a series of continuous-space models for 3-, 4-, and 6-taxon species trees using the forward-time simulator SLiM v3.3 (Haller and Messer 2019). Habitats were modelled as a grid with dimensionality ( $x, y$ ), and were subset by how speciation (or population divergence) occurred: 1) vicariant speciation or 2) peripatric speciation (Coyne and Orr 2004). Vicariant speciation begins with a population distributed across the entire range, and proceeds through serial fragmentation events (**Fig. S4.1a**). Peripatric speciation, alternatively, begins with a single population that expands into a new territory before being cut-off from its source population (**Fig. S4.1b**). Due to the nature of population divergence, population size per population fluctuates, either progressing through a series of expansions or contractions. All simulated species trees are asymmetric or comb-like:  $((AB)C)$ ,  $((((AB)C)D)$ , and  $(((((AB)C)D)E)F$ .

Population size is an emergent property of the local population density, which is governed by the spatial interaction distance,  $\sigma_I$ . The strength of  $\sigma_I$  is drawn from a Gaussian distribution with a maximum of  $1 / 2\pi\sigma_I^2$ , where  $\pi$  is the mathematical constant, and is experienced within a maximum distance of  $3\sigma_I$ . Mate choice is dictated by  $\sigma_M$ , the strength of which, like  $\sigma_I$ , is drawn from a Gaussian distribution with max  $1 / 2\pi\sigma_M^2$  within a max distance of  $3\sigma_M$ . After mates are chosen, the number of offspring produced are drawn from a Poisson distribution with  $\lambda = 2$ , which on average replaces the parents. Dispersal occurs immediately after offspring generation. Dispersal distance ( $\sigma_D$ ) is drawn from a uniform random distribution with a minimum and maximum of  $-3\sigma_D$  and  $3\sigma_D$ , respectively. Following dispersal, the parental generation dies. Simulated individuals in all models are diploid with 1000 Mb haploid genomes with a recombination rate of  $10^{-9}$ . Tree-sequences produced from each simulation are uploaded into *tskit* via *pyslim* (Kelleher et al. 2018). I then use the recapitation method in *pyslim* to

simulate a coalescent history across all genomic intervals in which there exists multiple roots (i.e., coalescence did not occur during the run; Haller et al. 2019). Using *msprime* (Kelleher et al. 2016), I overlay neutral mutations onto the trees at a rate of  $10^{-8}$ . Further details on each analysis performed is presented in the section in which it appears, and all SLiM recipes can be found at <https://github.com/hancockzb>. Detailed discussion of the model can be found in the *Appendix A*.

### Slatkin's Skew and Gene Tree Asymmetry

The multispecies coalescent (MSC) provides an intuitive way to investigate topological incongruence between gene trees. The standard MSC assumes that all gene tree incongruences are the result of ILS, and the pervasiveness of ILS can be predicted given we know something about the intervals between speciation events. This interval is measured in units of  $2N_e$  (or “coalescent units”). For a 3-taxon tree with topology ((AB)C), the probability of a gene tree being congruent with the species tree is  $1 - \frac{2}{3}e^{-t}$ , where  $t = T_D / 2N_e$ , and  $T_D$  is the time of divergence in generations. For a tree to be incompatible with the species tree – topologies ((AC)B) and ((BC)A) – coalescence must occur in the ancestral lineage of all three species. Since the MSC assumes panmixia and is space-independent, once all lineages are in the same population, they are equally likely to coalesce. Therefore, topologies ((AC)B) and ((BC)A) each occur with probability  $\frac{1}{3}e^{-t}$ .

The same logic can be applied to a 4-taxon tree. In the case of four tips, there are 15 possible topologies, and the probability of any one topology is a function of  $t_1$  and  $t_2$  (**Fig. 4.1c**). For the species tree topology (((AB)C)D), we can predict its probability as

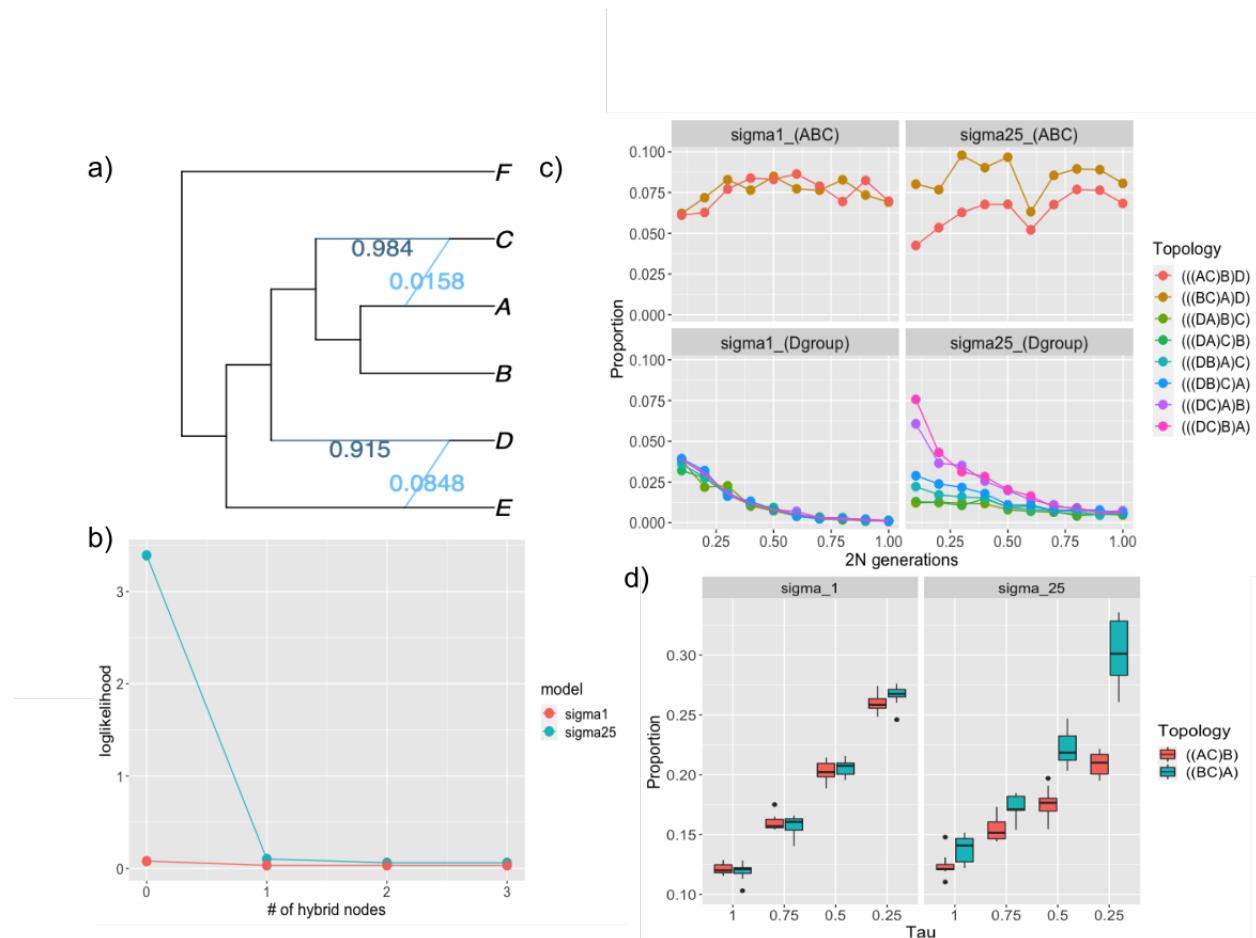
$$f(t_1, t_2) = 1 - \frac{2}{3}e^{-t_1} - \frac{2}{3}e^{-t_2} + \frac{1}{3}e^{-(t_1+t_2)} + \frac{1}{18}e^{-(3t_1+t_2)}.$$

When  $t_1 = t_2 = 1$ , we expect ~55% of the gene trees to match the species tree topology (Rosenberg 2002; Degnan and Rosenberg 2006). For gene trees that are discordant with the species tree, the presence of each topology is related to the length of intervals  $t_1$  and  $t_2$ . Focusing only on asymmetric trees, those with topology (((AC)B)D) and ((CB)A)D result from the lack of coalescence between lineage A and B during the interval  $t_2$ , and therefore should occur with equal probability (Rosenberg 2002). The same is true for any tree in which D is an ingroup – they can only result from the lack of coalescence during interval  $t_1$ , and therefore all have equal probability. The tree (((AB)D)C) occurs with slightly higher probability because it only relies on lineages (AB) and C not coalescing during the single interval  $t_2$ , as opposed to all three lineages (A, B, and C) not coalescing during either interval  $t_1$  and  $t_2$ .

These underlying gene tree stoichiometries are the basis for many common tests of hybridization, such as the ABBA-BABA test (Green et al. 2010; Durand et al. 2011) and forms the basis for network approaches to resolving gene tree conflict (Solís-Lemus and Ané 2016). The logic is straightforward; for example, in the case of a 3-taxon tree, when there is a significantly greater proportion of ((BC)A) topologies compared to ((AC)B) topologies, one might infer that gene flow has occurred at some point in the past between lineages B and C.

Slatkin and Pollack (2009) were the first to point out that gene tree asymmetry may not be the result of hybridization, but instead can emerge due to underlying population structure. They imagined a simple 3-taxon tree in which there was a barrier to gene flow in the ancestral lineage of species A and species B+C. They found that the stronger the barrier to gene flow, the greater the gene tree asymmetry became. I refer to this asymmetry due to ancestral population structure as “Slatkin’s skew”.

To illustrate Slatkin's skew, I simulated 3- and 4-taxon trees in continuous space. These populations were modelled as “clustered” or “unclustered”. Clustered models had low  $\sigma_D$  and  $\sigma_M$ , but high  $\sigma_I$  – this resulted in clumping patterns of ancestry across the range. The unclustered model was the opposite – high  $\sigma_D$  and  $\sigma_M$  with low  $\sigma_I$  – and was meant to approximate a random mating population. For these simulations, we used both the vicariant and peripatric speciation modes to demonstrate the differential impact on gene tree asymmetry. Gene trees were parsed in *tskit* and imported into the R platform (R Core Team 2020) where I used the package *evobiR* (Blackmon and Adams 2015) to count the different topologies among the trees.



**Figure 4.2.** Demonstrating Slatkin's skew. A) Inferred hybridization using SNaQ; dark blue is the CF for the bifurcating tree and light blue is the discordant proportion; B) change in pseudolikelihood at increasing number of hybrid nodes; C) 4-taxon gene tree distributions of discordant topologies; D) 3-taxon tree distributions of discordant topologies.

Results from these simulations are shown in **Fig. 4.2**. For the case of the 3-taxon tree previously investigated by Slatkin and Pollack (2009), I show that the same gene tree asymmetry will arise in continuous-space given that the mode of speciation is vicariance (**Fig. 4.2d**). The asymmetry becomes more skewed by geography as  $t$  approaches 0. As expected, similar trends emerge with the 4-taxon tree: when D is an ingroup, the skew becomes more dramatic as  $t_1$  and  $t_2$  in units of  $2N$  fall below 0.75. Interestingly, for the (ABC) coalescence, topology (((BC)A)D) occurs with greater frequency at *all* values of  $t_2$  in the clustered models, though the difference becomes more extreme at lower  $t_1$  values (**Fig. 4.2c**). When the mode of speciation is peripatry, no skew occurs, and we find that the proportion of ((AB)C) trees are much higher at all  $t$  than expected under a neutral MSC model.

Slatkin's skew has mostly been investigated in the context of human population genetics where there is interest in distinguishing ancestral structure from introgression with Neanderthals (e.g., Krings et al. 1997; Nordborg 1998; Wall 2000; Yang et al. 2012; Schumer et al. 2018). Eriksson and Manica (2012), for example, extended Slatkin and Pollack's (2009) model to an ancestral stepping-stone and found that the polymorphisms shared between Eurasians and Neanderthals was compatible with a scenario in which no hybridization had occurred. Theunert and Slatkin (2017) replicated their results and proposed using the unconditional site frequency spectrum as a tool to distinguish ancestral population structure from hybridization. However, outside of human population genetics there are few examples in which the impact of ancestral structure is even mentioned as a potential cause of gene tree discordance (but see Degnan 2018; Huynh et al. 2019).

Phylogenetic networks have become a popular tool to visualize gene tree incongruence. Whether the network is implicit (SplitsTree; Huson 1998) or explicit (SNaQ; Solís-Lemus et al.

2017), reticulations are interpreted as evidence for historic gene flow (Degnan 2018).

Phylogenomic studies often leverage species tree inference programs to generate a guide-tree for network inference (Cui et al. 2013; Árnason and Halldórsdóttir 2019; MacGuigan and Neer 2019). However, these approaches are not immune from Slatkin’s skew as they are based on the same standard MSC model as the  $D$ -statistic.

Using the true gene trees from the 6-taxon simulation as input, I inferred phylogenetic networks using SNaQ with the true species tree as the guide. I performed 10 runs per model, each with the maximum number of hybrid nodes ranging from 0–3. Following recommendations from Solís-Lemus et al. (2017), I chose the best number of hybrid nodes based on a slope heuristic (**Fig. 4.2b**). For the clustered model, incorporating even a single reticulation improved the pseudolikelihood from 3.392177 to 0.1021881 (**Fig. 4.2b**). The improvement for the unclustered model was far less dramatic: no hybrid nodes ( $\text{loglik} = 0.0788371$ ) compared to 1 hybrid node ( $\text{loglik} = 0.0332104$ ). The unclustered runs never inferred more than 1 hybrid node, regardless of how high the max number was set.

When will Slatkin’s skew matter? In short, when spatial autocorrelation persists between speciation intervals. We can prime our expectations by considering a “separation of timescales” approach to the coalescent (Wakeley, 1999; Wilkins, 2004). The first timescale – the *scattering phase* – is characterized by spatial autocorrelation. Rates of coalescence in this phase are greater than the expected  $1 / N_e$  since lineages nearer one another share an ancestor more recently than expected given the size of the population. The spread of ancestry across a range is diffusive and occurs at a rate of  $\sigma\sqrt{n}$ , where  $n$  is the generation and  $\sigma$  is the dispersal distance (Bradburd and Ralph, 2019; Kelleher et al. 2016). Once a lineage has ancestors that have diffused across the entire landscape backward in time, the coalescent enters the *collecting phase* (Wakeley 1999).

This phase can be thought of as behaving according to the standard coalescent model, in which the average time to a common ancestor for  $k = 2$  is  $2N_e$  and for  $k = i$  is  $4N_e / [i(i - 1)]$ . Since the rate of diffusion is dictated by  $\sigma$ , the timescale separating these two phases are proportional to the dispersal potential and the length of the range. If the length of the speciation interval is short relative to the timescale separating the scattering and collecting phases, then the spatial distribution of the ancestral population will dictate gene tree stoichiometry and Slatkin's skew will be in play. However, for speciation intervals  $\geq 2N_e$  (or  $\sim 1.0$  coalescent unit), even when dispersal is very low, gene tree proportions are indistinguishable from expectations under the MSC.

### **Wicked Forests: Species Tree Inference and Gene Tree Variance**

At a fundamental level, phylogenetic inference seeks to infer the relationships between species (topology) and the distance between them (branch-lengths). The currency of phylogenetics are gene trees, which are themselves biased subsets of the true population genealogy (Kelleher et al. 2016). From the distribution of gene trees, the multispecies coalescent model can be used to infer the species tree – the “true” history of the population – by conditioning a proposed species tree on a set of gene trees (Maddison 1997; Degnan and Rosenberg 2009; Chifman and Kubatko 2014). It’s important to note that while gene trees are real things – whether or not we can observe them directly – species trees are not. They are merely representations of the “true” history of a species in respect to how it has shaped the underlying gene tree distributions.

The genetic distance between any two species is on the order of  $T_D + 2N_e$  generations, where  $T_D$  is the time of divergence (the “true” history). This is because two lineages cannot coalesce until at least time  $T_D$ , and once they are in the same population it on average takes  $2N_e$  generations for them to coalesce (Wakeley 2000). Since we cannot observe gene trees directly and must rely on mutation to supply us with the information to infer them, it’s useful to note that, assuming panmixia, the pairwise sequence divergence between species 1 and 2 is  $\pi_{12} = 2\mu(T_D + 2N_e)$ , where  $\mu$  is the per generation mutation rate (Wakeley 2000).

Two important sources of bias from this expectation emerges when the  $z$ -axis is considered. First, population structure acts to inflate  $N_e$  relative to  $T_D$ , which leads to a greater contribution to  $\pi_{12}$  from  $2N_e$  (Slatkin 1991; Wakeley 2000; Hancock and Blackmon 2020). Second, spatial structure increases the variance in the  $T_{MRCA}$  as geographically close lineages coalesce rapidly whereas more distant lineages take much longer (Wall 2000). This second component also impacts the topological concordance of a given gene tree and is biased towards discordant topologies. This is because there is a hard lower-bound to the node height – it cannot be lower than  $T_D$  – whereas there is no such upper-bound. Therefore, spatial structure is expected to on average inflate branch-lengths and increase topological discordance among gene trees.

In general, we are interested in extracting some estimate of time from node heights. Under neutrality, if we have a good estimate of the per generation mutation rate,  $\mu$ , then we can estimate time directly as  $\pi_{12} / 2\mu$ . This is because the rates of substitution and mutation are equivalent under neutrality, and, importantly, independent of  $N_e$  (Kimura 1987). This finding is generally robust, even in the presence of population structure, as long as the rate of coalescence is still  $1 / 2N_e$  (Lehman 2014). However, when  $N_e$  varies across a landscape, as is often the case in spatially structured populations (Wilkins and Wakeley 2002; Bradburd and Ralph 2019), the

rate fluctuates across the range causing the local  $N_e$  to impact substitution rates (Allen et al. 2015). This has implications for many widely used molecular clocks (e.g., Knowlton and Weigt 1998; Weir and Schluter 2008; Herman et al. 2018) because the calibration source is unlikely to adequately represent the spatial population dynamics of a different species.

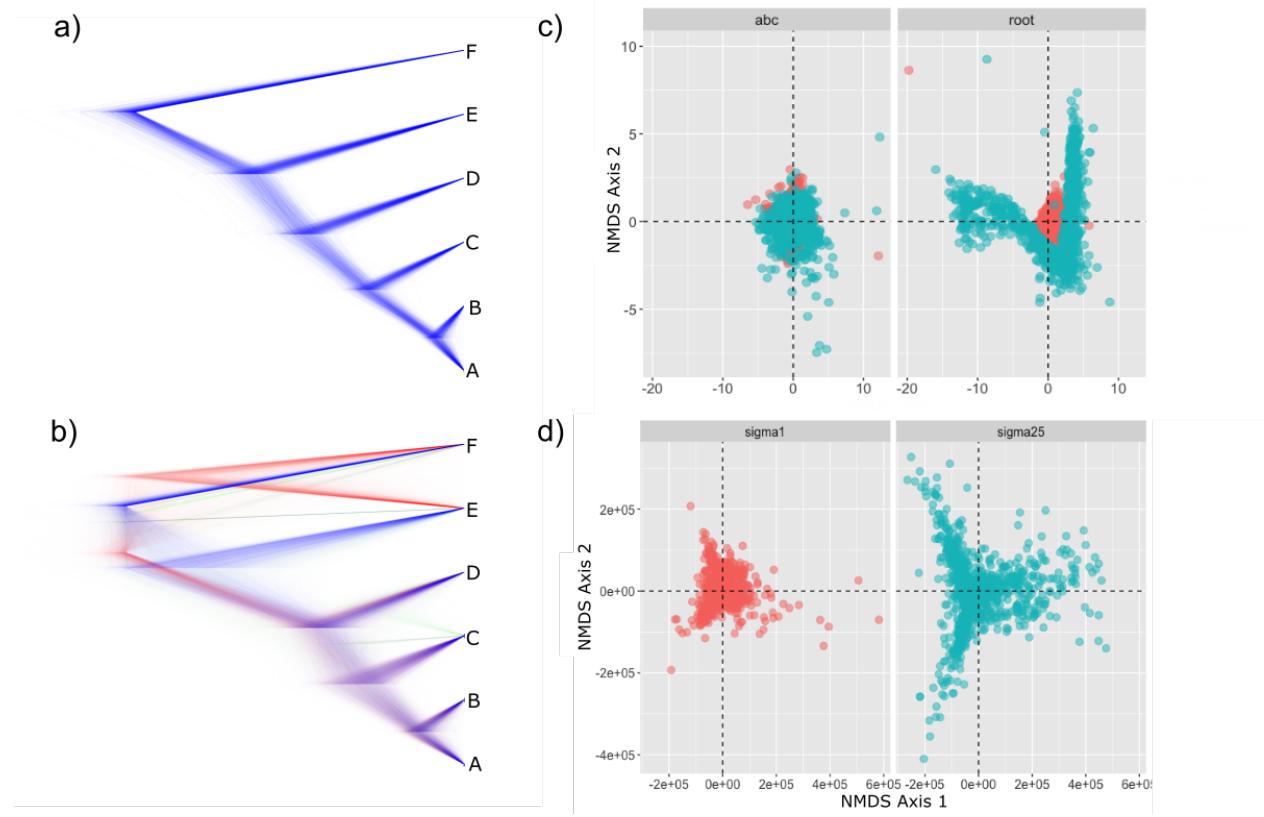
The rate of substitution can also be calibrated by an outside source, such as a fossil or a known biogeographic break. These calibrations are often in the form of node constraints (Bromham et al. 2018). However, the placement of the constraint can have an impact on the estimated age and rate of evolution (Sauquet et al 2012; Duchêne et al 2014). Duchêne et al (2014) recommended including as many fossil calibrations as possible, with the deeper ones being preferred. However, Brown and Smith (2018) showed that more doesn't necessarily equal better, as the interactions between the tree prior and node priors can cause deviations from the user-specified ages.

To illustrate the impact of spatial structure on tree variance and branch-length estimation, I randomly selected one genome from each species of the 6-taxon species tree vicariance simulations. These models have a uniform  $T_D$  of 10,000 generations except for the most recent divergence between A and B, which occurred 5,000 generations in the past. After applying neutral mutations in *msprime*, I output all SNPs in FASTA format. To estimate the species tree, I used two inference programs, \*BEAST2 (Ogilvie et al 2017) and SNAPP (Bryant et al 2012). For \*BEAST2, I first created pseudo-genes by concatenating SNPs in intervals of 500 for the first 5000 in the genome for a total of 10 pseudo-genes. These are “pseudo” in the sense that while their topology and branch-lengths tend to be correlated, they are not completely linked – what Springer and Gatesy (2016) call *concatenalescence*. While not ideal, this is standard practice for almost all phylogenomic datasets as the actual length of a “c-gene” may be as small

as 10 bases (Springer and Gatesy 2016). The dataset for SNAPP consisted of 10,000 randomly selected SNPs from across the genome. Unlike \*BEAST2, SNAPP does not estimate gene trees directly, instead iterates over all possible gene trees from biallelic SNPs to generate a species tree (Bryant et al 2012). In \*BEAST2, I performed two separate runs per dataset: 1) with the root-age calibrated; and 2) with the clade ((AB)C) calibrated – or the “midpoint” calibration. I was particularly interested in differences that arose between the clustered ( $\sigma = 0.25$ ) and unclustered ( $\sigma = 1.0$ ) models relative to estimated clock rates, tree heights, and topologies. I visualized the estimated clock rates and calibration priors using Tracer v1.4.3 (Rambaut et al. 2018), and incongruence between estimated ages and topologies using the package DensiTree (Bouckaert 2010). In addition, I performed non-metric multidimensional scaling (NMDS) in R on sets of 1000 randomly selected trees from the posterior distribution. These trees were compared using weighted Robinson-Foulds (RF) distances, which account for both topology and branch length (Robinson and Foulds 1979). Finally, I compared the variance in estimated trees with the true gene tree variance by randomly sampling 1000 true gene trees, generating a weighted RF distance matrix, and plotting them using NMDS.

The results from these analyses demonstrate the stark differences that emerge when spatial structure is strong (**Fig. 4.3**). I found that spatial structure greatly increased the variance in both the true gene trees (**Fig. 4.3d**) and the trees sampled from the posterior distribution when calibrated at the root (**Fig. 4.3c**). However, for the \*BEAST2 analyses, constraining the lineages (ABC) to be monophyletic reduced the inferred species tree variation as shown in **Fig. 4.3c**; the clustered and unclustered distributions largely overlap each other in tree-space. In addition, the species tree inferred from the midpoint calibration was incorrect with topology (((AB)C)D)(EF)). While the rooted calibration model generated the correct topology, node

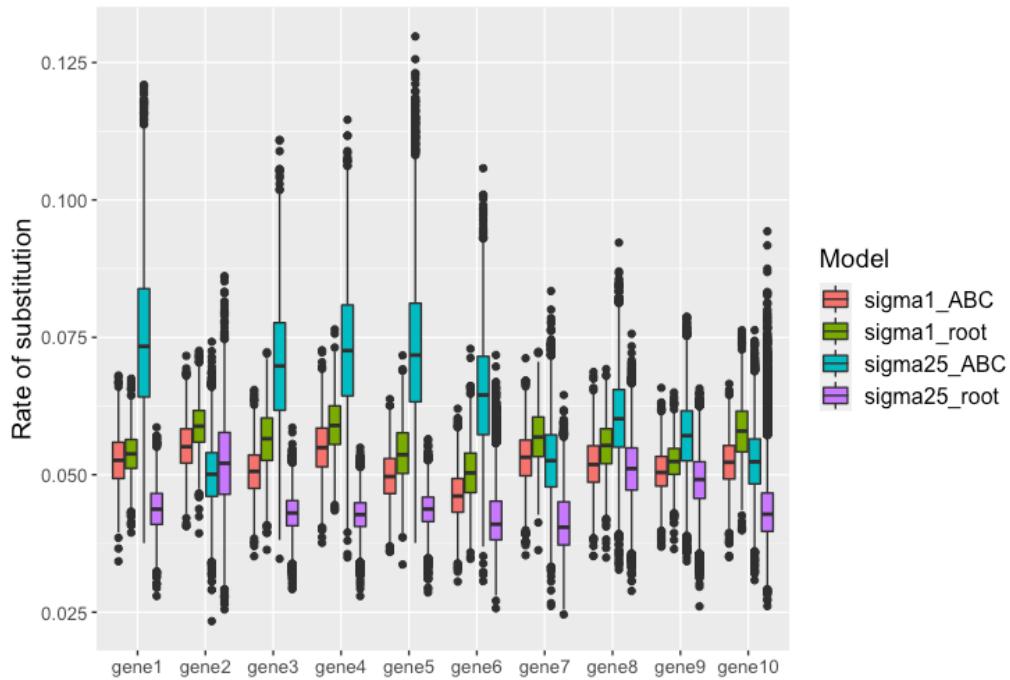
support was much lower than the unclustered models, which all had posterior probabilities of 1.0 (**Fig. S4.4**). The placement of the calibration impacted the estimated clock rate in both the clustered and unclustered models, but the rate differences between root and midpoint placements for the clustered model were far more extreme (**Fig. 4.4**). Irrespective of calibration placement, the 95% highest posterior density (HPD) of the node ages all included the true age for the unclustered models. For the root-calibrated clustered model, the true ages were obtained for all nodes, but the variance in the age estimate was much higher than in the unclustered model (**Fig. 4.4**). However, when the calibration was placed at the (ABC) node, the clustered model failed to identify the true ages for several nodes, and actually underestimated the root age by 22–60%.



**Figure 4.3.** Gene tree variance. A) reconstructed gene trees from the unclustered model (sigma1); B) trees from the clustered model (sigma25); C) variance in topologies and branch lengths based on weighted RF distances, where “abc” is midpoint calibration and “root” is root calibration; D) true gene tree variance based on weighted RF distances.

The SNAPP analysis, despite being designed to handle biallelic SNPs, performed much poorer. Both the clustered and unclustered analyses produced aberrant topologies with high posterior probability, indicating that any violation of the panmixia assumption will bias species tree inference in SNAPP. However, estimates of  $\theta$  generally followed expectations, inferring higher values for the ancestral branches of the clustered versus the unclustered (**Fig. S4.5**). In addition, the estimated tree height was greater in the clustered compared to the unclustered model (**Fig. S4.5**).

Degnan and Rosenberg (2006) and Degnan and Rhodes (2015) coined the term “wicked forests” to refer to the curious case where a set of species trees are one another’s most likely anomalous tree. In this section, I use the term more colloquially to refer to the expanded tree-space (or “forest”) that grows as spatial structure increases. These forests are “wicked” because they reduce our confidence in substitution rate, divergence-time, and topology estimates in species tree inference. Importantly, as was shown previously by Duchêne et al. (2014), the placement of the node calibration can rival the impact of spatial structure, and when placed at the midpoint of the tree actually masks the true gene tree variance. This masking also appears to reduce the ability of \*BEAST2 to accurately infer the species tree. Since node calibration relies on constraining monophyly, the MCMC is limited in the space of trees it searches and therefore does not have a holistic sampling of the extent of ILS.



**Figure 4.4.** Substitution rate estimates from \*BEAST2 for 10 pseudo-loci.

### Species Delimitation Delimits Structure

Identifying species boundaries is at the heart of systematics. The collection of techniques that fall under the category of “species delimitation” all aim, fundamentally, to distinguish between population structure and true species boundaries. These methods include four broad categories: 1) traditional morphological taxonomy; 2) distance-based methods; 3) tree-based methods; and 4) coalescent methods. It has been recommended to adopt a kind of holistic, consensus approach by considering results from each in deciding how robust the delimitation is (Carstens et al. 2013). The rise of the latter three methods coincides with increased sequencing effort for non-model organisms and the coincident “taxonomic impediment” (Lipscomb et al. 2003; Scotland et al. 2003). Each of these methods rely on a phylogenetic species concept (de

Queiroz 2007) since they do not explicitly examine reproductive barriers – instead, these are inferred from an absence of gene flow between groups of samples.

Sukumaran and Knowles (2017), by simulating speciation under the protracted speciation model (PSM), were the first to point out that coalescent methods conflate lineage splitting with speciation. These results were replicated with an empirical system by Chambers and Hillis (2020), who showed that coalescent methods over-split species in the case of geographically widespread taxa. Leaché et al. (2019) respond by noting that the use of the PSM by Sukumaran and Knowles (2017) assumes instantaneous speciation, and that due to this feature no method would be capable of distinguishing population splits from speciation.

Both coalescent and tree-based methods were recently reviewed by Luo et al. (2018) under a series of speciation scenarios. They found that BPP (Rannala and Yang 2003) – a common coalescent software – generally performed better than the two tested tree-based methods, the General mixed Yule coalescent (GMYC; Pons et al. 2006) and Poisson tree processes (PTP; Zhang et al. 2013). However, their models of speciation did not incorporate population structure, only gene flow following speciation. Talavera et al. (2013) investigated the performance of the GMYC on a large-scale butterfly dataset and found that it often overestimated the number of species relative to the recognized morphospecies. Papadopoulou et al. (2008) found that the GMYC was only impacted by population structure under an island model of migration when the product of the population size and migration rate,  $Nm$ , was  $< 10^{-5}$ . However, Lohse (2009) showed that by decreasing the number of demes sampled the GMYC would over-split even at higher migration rates.

In the previous section, I showed that low dispersal increased the variance in the  $T_{MRCA}$  between species. In the context of species delimitation, where you may not know whether you've

sampled from different species or merely different sites across the range of a single species, the variance in  $T_{MRCA}$  within a single range becomes important. This variance is poorly captured by discrete population models but is an emergent property of continuous landscapes with finite edges. Samples taken from the center of the range on average share an ancestor deeper in the past than those from the periphery (Wilkins 2002; Wilkins and Wakeley 2004). This is the result of an increased population density in the center relative to the range edges. How does this variance in  $T_{MRCA}$  across sampling locations impact species delimitation? Species delimitation methods, especially tree-based methods, rely on recognizing a shift in the coalescent dynamics of samples that are within versus between species. Spatial structure can increase the probability of monophyly between sampling locations, which makes them resemble discrete population clusters. The height of the node separating these clusters will dictate whether species delimitation methods identify them as separate species or merely clusters of samples within a single species. When samples are from the periphery, the distance between the clusters will be greater because each cluster, on average, shares a most recent common ancestor more recently than samples taken from the center of the range (Fig. S4.1–2). I contend that this feature of the coalescent in continuous populations demonstrates the importance of considering not merely the distance between two samples or the number of sampled sites, but whether those samples come from the core or the periphery of the range.

While these shortcomings have been identified for coalescent and tree-based methods, I contend that *all* species delimitation methods – including morphology-based – can be fooled by spatial structure. Unlike Sukumaran and Knowles (2017) and Leaché et al. (2019), I focus here solely on the ability of tree-based and coalescent-based models to distinguish population splits (i.e., I make no assumption that speciation has occurred).

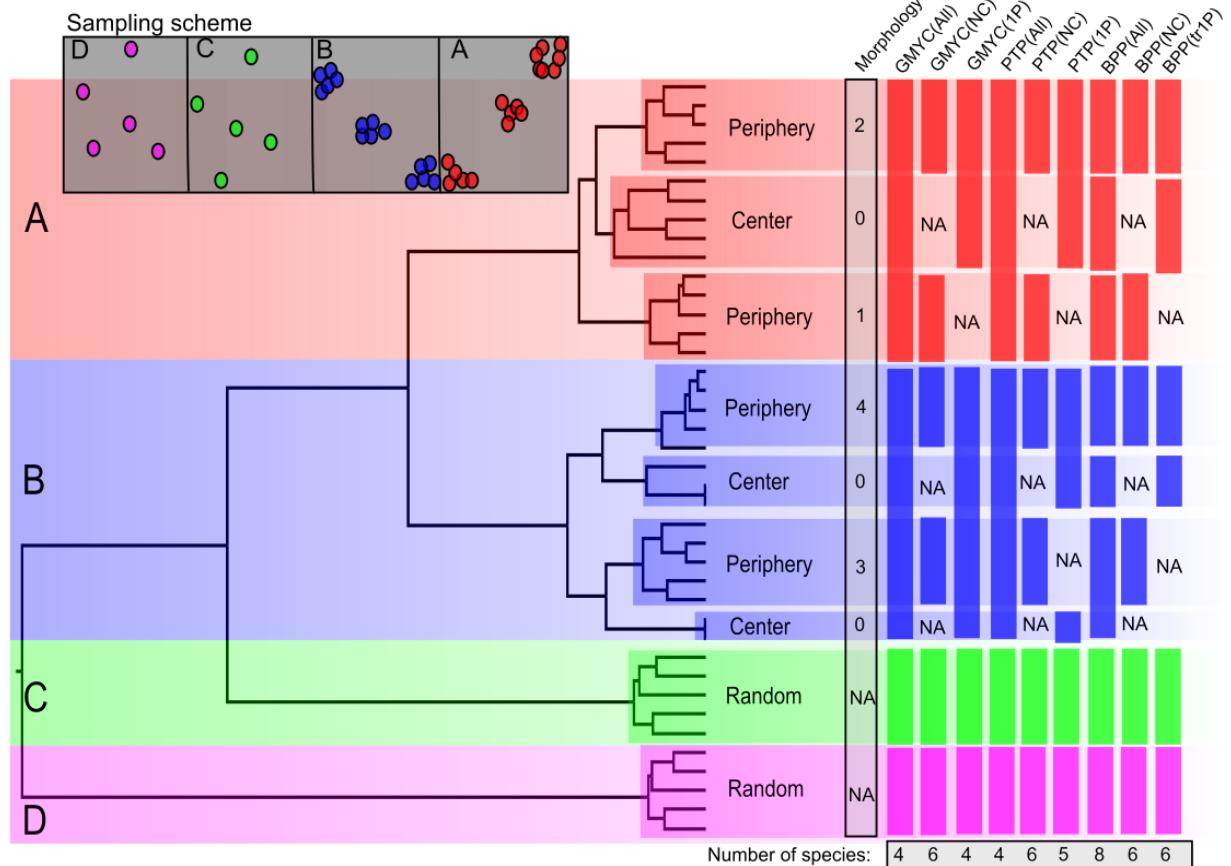
To investigate the performance of species delimitation methods on sequence data that arises from continuously distributed populations, I first sampled across the range of my simulated 4-taxon species tree vicariance models for both the clustered and unclustered settings. For species A and B, I sampled 5 individuals each from three locations, including two edges and the center (**Fig. 4.5**). For species C and D, I randomly sampled 5 individuals across the range. I then randomly sampled 10,000 SNPs from the genome alignment and constructed an ultrametric tree using BEAST2 with a birth-death tree prior. I generated consensus trees using TreeAnnotator and imported these trees into the R platform. For the GMYC method, I used the package *splits* (Fujisawa and Barraclough 2013) to calculate the ML estimate for the transition between a Yule and coalescent model. For the PTP method, I used the web server (<https://species.h-its.org/ptp/>) and both a ML and Bayesian estimation of the number of splits in the tree. For the coalescent-method, I used BPP, which can jointly infer the species tree and perform species delimitation using a reversible-jump MCMC algorithm (Rannala and Yang 2013). For BPP runs, I generated 5 pseudo-loci (see section *Wicked Forests: Species Tree Inference and Gene Tree Variance*) of 500 SNPs. The MCMC was run for 100,000 generations with a 20,000 generation burn-in. To illustrate how morphological delimitation can be deceived in the presence of population structure the same as tree- and coalescent-based methods, I simulated 20 binary traits onto the phylogeny produced by BEAST2 with *phytools* (Revell 2012) in R under a Brownian-motion model. The  $Q$ -matrix was defined such that there was an equal transition probability between trait 0 and 1. I then counted traits that clustered by sampling location as opposed to by species.

Delimitation studies aimed at identifying species, especially in the case of cryptic species complexes, generally rely on a consensus approach to interpret their delimitations. In **Fig. 4.5**,

we see that *all* delimitation methods can be misled by spatial structure when sampling is biased towards low diverse areas of the range (such as the periphery). Furthermore, if we were to search for specific morphological traits to support our delimitations, we have a good chance of finding at least 1 trait (**Fig. 4.5**). The specific sampling scheme also has a clear impact, but not necessarily because some samples were random whereas others were clustered. Instead, what matters is where the samples are collected in respect to the local population density. When samples are drawn from the periphery, they will have shorter coalescent times and therefore longer branches to other samples from the same population. However, drawing samples from the center of the range, where the local  $N_e$  is the highest, largely ensures that deeper node heights will be included. These deeper heights reduce the clustering pattern and the probability of monophyly. For example, the B species samples from the center break the monophyly between the peripheral populations (**Fig. 4.5**). In the case where all three samples are monophyletic, as in species A, the deep node heights of the center samples are sufficient to allow the tree-based methods to accurately identify the transition between Yule and coalescent models. Removing these central samples accentuates the clustering pattern, and consistently fools both tree-based and coalescent-based methods. The geographic locations were not monophyletic in the unclustered model and the node heights on the inferred tree were as deep as the random samples. Therefore, I did not perform the tree-based methods as no empirical study would use these methods in the absence of any geographic clustering.

While the tree-based methods performed better when all samples were included, BPP consistently overestimated the number of species. The species number with the highest posterior probability in each run matched the number of geographic samples, regardless if the center samples were included or absent (**Fig. 4.5**). In the unclustered model, BPP identified 5 species

(PP = 0.598) when the guide species tree included all 8 geographic locations. This was much closer to the true 4 species total than in the clustered model.



**Figure 4.5.** Species delimitation in continuous populations. Top-left boxes represent the sampling scheme, and the tree showed includes all samples, with clades colored by the species they belong to. Bars on the right are the inferred number of species from each method. “All” designates all samples included; “NC” is no center; “IP” includes a single periphery sample with the center. For the Morphology section, numbers beside clades are the number of simulated binary traits on the complete tree that support that clade as distinct from the rest of its conspecifics.

Unlike previous examinations of species delimitation methods in which independent lineages may be designated as “lineages” or as *bona fide* “species” (Sukumaran and Knowles 2017; Leaché et al. 2019), I only assess the ability of these methods to identify actual population

splits irrespective of if reproductive isolation had occurred. In addition, by using a continuous landscape I am better able to assess sampling schemes and how variation in local population density impacts delimitation methods. In contrast to Luo et al. (2018), I find that the tree-based methods outperform BPP when samples from the center of the range are included, and that of those the GMYC consistently converged on the true number of species (**Fig. 4.5**).

### **Out of Space: Continuity for Discontinuous Data**

In a review of Battey et al. (2020) for the *Molecular Ecologist*, Jeremy Yoder described continuous space as the “final frontier” for population genetics (Yoder 2020). In this chapter, I extend Yoder’s proclamation to phylogenetics as well. As continuously distributed populations impact virtually all commonly used summary statistics in population genetics (Battey et al. 2020), I show above that it also affects broader macroevolutionary patterns. Each of these patterns have been discussed in the literature previously (e.g., Edwards and Beerli 2000; Slatkin and Pollack 2009; Sukumaran and Knowles 2017; Hancock and Blackmon 2020) but have not explicitly been tied to continuous populations or with a focus on the mode of speciation. Below, I review the preceding sections by answering three questions that empiricists should consider when conducting phylogenetic inference.

*When does space matter for phylogenetics?*

Answer: When spatial autocorrelation of ancestry persists between speciation intervals.

This is most likely to occur when there is heterogeneity in population density across the range, low dispersal or local retention, and speciation intervals are short relative to ancestral  $N_e$ .

As discussed in *Slatkin's Skew and Gene Tree Asymmetry* above, spatial autocorrelation of ancestry decays with time at a rate dictated by the size of the range and the rate of dispersal. Given enough time, the ancestors of a sample of individuals will appear as a random sample across the range and the rate of coalescence will collapse to Kingman's coalescent (Kingman 1982; Wakeley 2009). To provide a simple illustration of this, imagine a Markov process with transition matrix

$$\mathbf{P}^t = \begin{bmatrix} a_a & a_{a \rightarrow b} & a_{a \rightarrow c} \\ b_{b \rightarrow a} & b_b & b_{b \rightarrow c} \\ c_{c \rightarrow a} & c_{c \rightarrow b} & c_c \end{bmatrix}^t,$$

where  $a$ ,  $b$ , and  $c$  represent transition probabilities to either different states (above and below the diagonal) or remaining within the currently occupied state (along the diagonal). As a concrete example where  $t = 1$ :

$$\mathbf{P}^1 = \begin{bmatrix} 0.99 & 0.01 & 0.0 \\ 0.01 & 0.98 & 0.01 \\ 0.0 & 0.01 & 0.99 \end{bmatrix}.$$

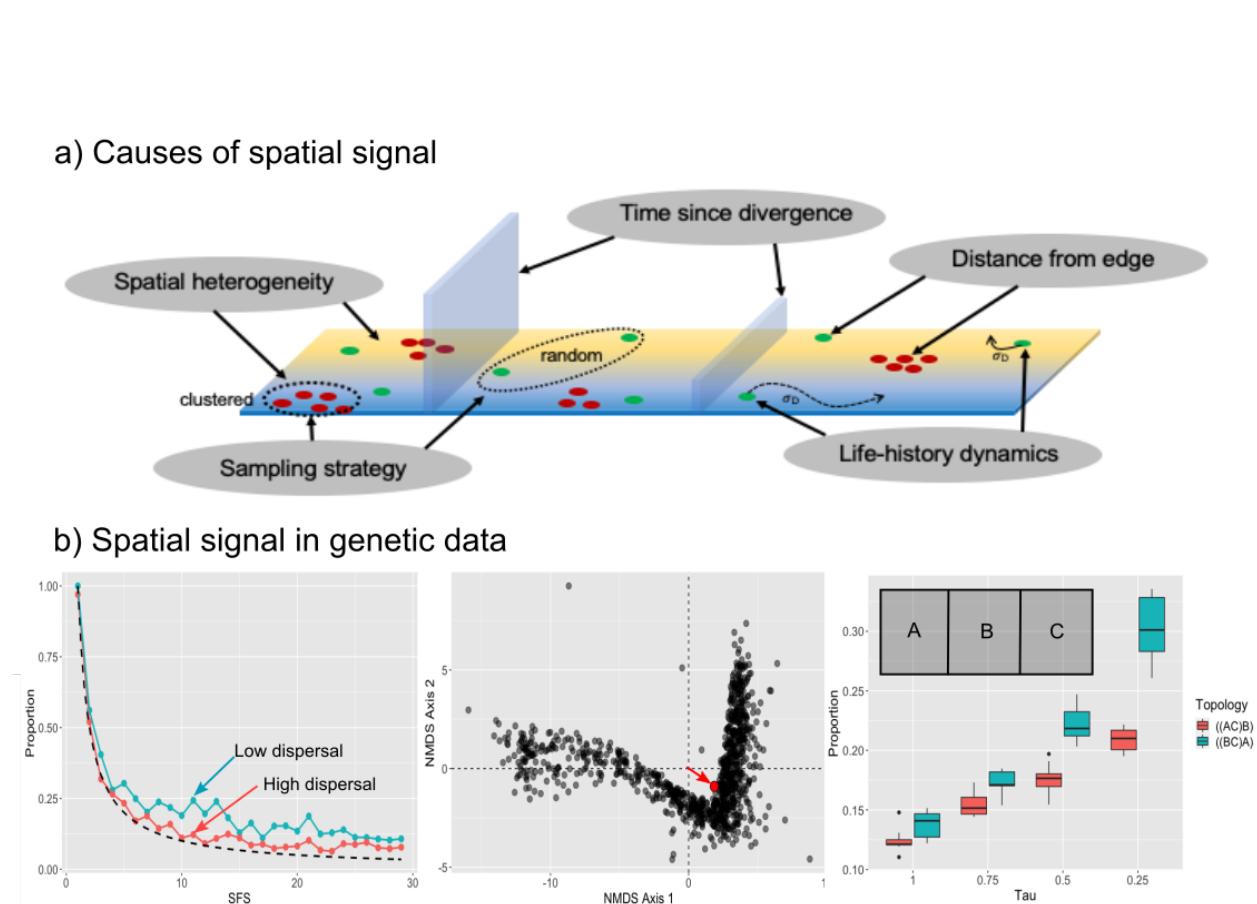
In this example, there are three demes represented by each row in the transition matrix. The transition probabilities can be thought of as dispersal rates between demes. Clearly, when  $t$  is small the ancestors of any deme will be biased towards the deme that they were originally sampled from. For example, when  $t = 5$ :

$$\mathbf{P}^5 = \begin{bmatrix} 0.99 & 0.01 & 0.0 \\ 0.01 & 0.98 & 0.01 \\ 0.0 & 0.01 & 0.99 \end{bmatrix}^5 = \begin{bmatrix} 0.952 & 0.047 & 0.001 \\ 0.047 & 0.906 & 0.047 \\ 0.001 & 0.047 & 0.952 \end{bmatrix}.$$

While spatial autocorrelation is still strong at  $t = 5$ , we see that there is now a non-zero probability of a lineage originating in deme 1 having an ancestor in deme 3, and vice-versa. This toy example also shows that central demes spread out to the periphery much faster than peripheral populations reach one another, which indicates that if the sample originates from the center of the population the transition to the scattering phase may be faster than for samples from the periphery. How much faster? In my example, by  $t = 45$  (or 45 generations before the present), samples from the center have a 50:50 chance that their ancestor was also in the center deme. Alternatively, the peripheral demes each have  $\sim 70\%$  probability that their ancestors at generation 45 are still found in the sampled deme.

The transition to the collecting phase occurs as the probability that a sampled individual has an ancestor in any deme is equal. In fact, Wakeley (1999) defined the end of the scattering phase as the  $t$  when all ancestral lineages are in separate demes (i.e., they are scattered evenly across space). Continuing with our example, this would occur when the transition probabilities are all  $\sim 33\%$ , which does not occur until  $t \approx 400$ . In a phylogenetic context we might imagine one of the peripheral demes has a complete barrier to gene flow with the others until  $T_D$ . At this time, bias in tree topologies and branch lengths will occur if ancestors are still correlated with their sampled deme. In this case, if the true tree is ((Center, Periphery1), Periphery2), we would expect a skew in topologies towards ((Center, Periphery2), Periphery1) over ((Periphery1, Periphery2), Center), where the MSC would predict they should occur with equal probability.

Furthermore, trees that do show the latter topology will be biased towards older node heights since it on average takes longer for the edges to find ancestors with one another.



**Figure 4.6.** How space impacts phylogenetic datasets. A) a schematic of the different causes of spatial signal; B) specific signals that can be found in phylogenetic data. From left to right: skews in the SFS at lower dispersal; pull of the mean tree away from the origin in NMDS space; skews in gene tree topologies that reflect a vicariant history.

Clearly, the rate of transition between the scattering and collecting phases is a function of the number of demes and the rate of migration between them. While this is much more complex in continuous space (there are more dimensions to transition through, for example), this simple example can serve to prime our expectation of when space will matter in macroevolutionary studies. Also, important to note is that the above example is independent of  $N_e$ , which is the parameter that actually determines the rate and average times to coalescence in real populations.

*What sort of organisms will be the most affected?*

Answer: Those with large ranges relative to their dispersal potential with a history of vicariance. A famous example from the biogeography literature is the vicariance event that impacted coastal marine organisms separated by the Florida peninsula (Bert 1988; Bowen and Avise 1990; Avise 1992; Flannagan et al. 2016). While several hypotheses have been purported to explain the cause of vicariance in these marine taxa (see Portnoy and Gold 2012 and Hancock et al. 2019 for recent reviews), sister species are otherwise widespread with varying degrees of population structure. Species that show low structure within populations, such as oysters (*Crassostrea virginica*) with pelagic dispersal, may approximate our “unclustered” models above, which show little to no bias on phylogenetic inference. Alternatively, species such as brooding amphipods (Hancock et al. 2019), which have poor dispersal potential and strong population structure, are more akin to our “clustered” models and are expected to have high gene tree discordance, increased variance in node heights across the posterior sample, and likely upwardly biased estimates of species diversity (all of which were documented in Hancock et al. 2019). Another example is the Gulf pipefish (*Sygnathus scovelli*), which displays population structure within species as well as discrete barriers between those isolated by the Mississippi River and the Florida peninsula (Flannagan et al. 2016).

*How can we detect if spatial structure is influencing our phylogenies?*

Answer: Fortunately, our data contain a myriad of clues into a history of spatial structure. At the phylogeographic level, in which researchers are often examining species complexes or closely related taxa with broad geographic sampling, signals of vicariance and continuous structure can be quite obvious. Hancock and Blackmon (2020) evaluated the impact of isolation-by-distance on phylogenetic inference in a species complex of Brazilian endemic lizards. Despite the ability of species delimitation methods to identify discrete clusters (Domingos et al. 2017), when samples were drawn from the range edges the estimated ages were upwardly biased compared to neighboring localities (Hancock and Blackmon 2020). This pattern can only emerge if the range was once widespread with ancestral IBD.

The impacts of space will also lead to high variance in reconstructed topologies and low posterior probabilities, especially at deeper nodes (**Fig. 4.4**). A model of ILS alone is sufficient to predict the topological incongruence in **Fig. 4.4** – i.e., we expect some of the most common trees that are incongruent with the species tree to be (((AB)C)D)(EF) and (((AB)(CD))E)F – but it does not explain the high variance in reconstructed node heights. Furthermore, when the calibration is placed at the root, I find that the increased variance in both topologies and node heights in the posterior emerge as a result of the same variance in the underlying true gene trees. While there are many factors in practice that can increase variance, such as low information content, sequencing error, etc., the most fundamental is the true gene tree variance. Assuming that other sources of variance can be isolated and accounted for, recognizing this underlying gene tree structure could be an indirect way to peak at the demographic processes responsible for patterns in the data. A possible test of spatial structure would then be to examine the tree-space of reconstructed topologies and branch-lengths from the posterior. A peculiar pattern that emerges when examining the spread of trees in multidimensional space is that strong spatial

structure pulls the mean tree away from the origin in both the true trees (**Fig. 4.4d**) and the reconstructed trees (**Fig. 4.4c**).

Another promising avenue may be to compare the impact of node calibration placement on gene tree variance and rate estimation. While both the unclustered and clustered models showed some impact of calibration placement, the clustered model resulted in extreme differences between rates estimated with different placements (**Fig. 4.5**). In addition, like the topologies and branch-lengths mentioned above, the variance in these estimates were much higher than in the unclustered models. Thus, a side-effect of strong spatial structure is reduced confidence in rate estimation, topological inference, and branch-length estimation, even when incorporating node calibrations that include the true age. I recommend researchers not jointly infer substitution rates and topology; instead, first reconstruct the species tree and then infer rates on the fixed tree.

Gene tree stoichiometries are important sources of information about the underlying demographic histories of populations but can be misleading if not considered within an explicitly spatial context. In the case of a 4-taxon tree, one could investigate the possible impact of space by first summing the proportion of trees that include D (or the outgroup) as the ingroup (as shown in **Fig. 4.2c**). Next, one could regress the physical distance between D and its sister species for a given tree against the total proportion of gene trees. Significant negative relationships between topological proportions and distance indicate that gene tree topologies may be driven by historic spatial relationships. This does not preclude hybridization, but it does demonstrate that a more robust criteria must be adopted to distinguish spatial structure from introgression.

Theunert and Slatkin (2017) provided a method for distinguishing between ancestral population structure and admixture using the unfolded site frequency spectrum (SFS). They showed that structure causes a skew from the theoretical expectation of  $\frac{\theta}{i}$  towards intermediate frequency alleles, whereas admixture did not alter the SFS. I recapitulate their results in **Fig. 4.6b** and show that the unclustered model leads to much less skew in the SFS than the clustered. In a phylogenetic context, the SFS could be used in conjunction with methods that reconstruct ancestral areas, such as RASP (Yu et al. 2015), to distinguish ancient hybridization from continuous spatial structure. This could be done via model comparison in a Bayesian framework by conditioning a model of hybridization versus spatial structure on the shape of the SFS and the probability of hybridization given reconstructed ancestral areas.

In many respects, ancestral ranges actually confound our ability to distinguish structure from hybridization as individuals need to be in geographic proximity to hybridize in the first place, but it is this very proximity that makes structure a likely explanation for geographically skewed topologies. Ideally, we would weight the posterior probability of a model of hybridization versus structure by incorporating the SFS as a prior. But what about when hybridization *and* structure characterize the history of our samples? In this case, we expect a skew in the SFS towards intermediate frequency alleles in conjunction with overlapping ancestral areas, exactly as predicted by continuous structure alone. However, we could leverage simulations over various histories of admixture proportions in conjunction with spatial structure to determine which combinations of models are identifiable. For example, for periods of strong admixture in the distant past we might expect to see skews in gene tree topologies that exceed that predicted by spatial structure alone. Indeed, if hybridization was prevalent enough and

occurred  $> 2N_e$  generations in the past, the influence of spatial structure may be largely washed-out and most, if not all, of the skew should be caused by hybridization (**Fig. 4.2**).

The subfield that may be most improved by considering spatial structure is species delimitation. There exist several methods in spatial population genetics whose underlying structure could be co-opted to improve existing techniques. For example, Bradburd et al. (2018) introduced an R package, *conStruct*, which seeks to identify both discrete and continuous population boundaries by modelling an expected decay in allelic covariance with distance. Like STRUCTURE (Pritchard et al. 2000), *conStruct* accepts a user specified number of layers,  $K$ , but unlike STRUCTURE allelic covariance is allowed to decay with distance within a single layer. To incorporate geographic distance and the covariance within layer  $k$  of allele  $i$  and  $j$ , they use

$$G_{ij}^{(k)} = \alpha_0^{(k)} \left( \exp \left( - \left( \alpha_D^{(k)} D_{ij} \right)^{\alpha_2^{(k)}} \right) \right) + \phi^{(k)},$$

where  $\alpha^{(k)}$  parameters dictate the decay of covariance and  $D_{ij}$  is the geographic distance between sample  $i$  and  $j$ , and  $\phi^{(k)}$  is the background covariance within the entire layer (Bradburd et al. 2018). By setting  $\alpha_0^{(k)} = 0$ , we recover the nonspatial model (i.e.,  $k$  is governed by a single shared frequency parameter,  $\phi^{(k)}$ ). Bradburd et al. (2018) use a cross-validation approach to compare the spatial and nonspatial models. In this way, one can imagine the  $K$  with the highest support representing the maximum number of species present in a sample while accounting for the expected decay of relatedness with distance. This model could be extended to incorporate time separating the  $K$  discrete layers as allelic covariance is expected to decay across both space and time. Thus, for more distantly related species this decay will be less influenced by space than by time, and higher numbers of  $K$  should be preferred.

## *A Future in Space*

The field of phylogenetics has taken the form of an hourglass. At one bulge is phylogeography, in which multiple geographic locations and individuals within a species are sampled. Phylogeographers straddle the boundaries between population genetics and phylogenetics, plucking methods from both to query the recent past of relatively closely related individuals. At the opposite bulge is deep-time phylogenetics that is investigating the origins of families, orders, or even phyla, where the ancestors of sampled individuals lived on a landscape dramatically different from the one today. These phylogenies may have hundreds of tips, but each species is often represented by a single individual, and tips are not related to one another in space. Both ends of the hourglass have seen tremendous developments in the past decades thanks to improved sequencing technology and more computationally tractable methods for dealing with large numbers of tips. However, both tend to use phylogenies as only informative about time.

I contend that the next important steps in phylogenetics will not merely be in time, but in space. By incorporating the *z*-axis into our models, we can better identify the mode of speciation and characterize the demographic factors influencing our reconstructed trees. Furthermore, by adopting a spatial framework conceptually provides us with an intuition about gene tree variance and can caution our interpretations of widespread inferences of hybridization. In addition, considering space encourages us to visualize tree-space, which can lead to more rigorous scrutiny of consensus methods that often fail in the presence of population structure (DeGiorgio and Rosenberg 2016). Finally, I believe that spatial models of IBD can dramatically improve

species delimitation methods by conditioning the inferred number of species on an expected decay of allelic covariance with distance.

## BENEDICTION

“So long and thanks for all the fish.”  
—Douglas Adams, *The Hitchhiker’s Guide to the Galaxy*

*“I called the doctor on the telephone  
Said doctor, doctor, please  
I got this feelin’  
Rockin’ and a’reelin’  
Tell me, what can it be?  
Is it some new disease?”*  
—The Sylvers, “Boogie Fever”

“Well, Smith, I’ve just had me last shit.”  
—J.B.S. Haldane to J. Maynard Smith before surgery for rectal carcinoma  
(he survived and would go on to have more shits)

In the months before the first outbreak of the novel coronavirus in a wet market in Wuhan, Faith and I drove ~6,000 miles across the entire Gulf of Mexico and Atlantic coastlines collecting the amphipods that were the subject of Chapter II. We drove my mother’s Altima, backseat loaded with a giant portable freezer, collecting gear, ice chest of beer and food, a tent, and a tray with plastic bottles where live samples would be kept. We collected our first amphipods on the small island of Grand Isle in southern Louisiana – a foul<sup>4</sup> little fishing community bound for the drink in a mere 20 years or so. The puddles of brackish water that lined the road were festering with mosquito larvae and the air was thick with swarms of the bloodsuckers. We drove on to New Orleans that night and stayed with Liz Marchio, my old roommate Dan’s ex-wife. We wandered through the French Quarter as the sun descended the hazy horizon, crowds gathering outside the bars that lined the narrow streets. The city smelled of piss and beer, homeless people in rags stumbled through the alleys like zombies, swindlers hovered on every corner waiting to pounce on unsuspecting tourists. We perused the shops,

---

<sup>4</sup> A committee-member disagreed with this characterization of Grand Isle, calling it “quaint”. I point this out only in fairness; I prefer my adjective.

spending some half an hour in an establishment that bore the name *Marie Laveau's House of Voodoo*. The walls were decorated with odd artifacts, dolls skewered with pins, horrific masks with mangy goat hair, bizarre idols lining shelves and, in the back, an altar to the old dead witch herself, complete with pieces of bone, aged tarot cards, and smoke rings rising from the trays of burning incense. We ate at a little diner tucked in the basement of a motel, drank local beer and shared bowls of jambalaya and boudin. We slept on the floor of Liz's small house listening to the sounds of the city, echoing voices and the low hum of car engines, dogs barking in the distance.



“Strange Gods, Strange Altars”. Photo by Zach Hancock.

The Gulf adopts a dramatically different persona on the other side of the Mississippi. Along the Texas and Louisiana coastlines, where meandering rivers create extensive deltas of silt and mud, the ocean is turbid and hazy. The surf is gentle and deposits the dark soil of the

Mississippi across the barrier islands; a gloomy, filthy dress of organic matter and driftwood. In the eastern Gulf, the ocean becomes clear and blue, the sand rich with quartz and white as snow. The surf is strong and creates a sharp berm on the beach-face. Waves break against the shore and sampling is rough – almost impossible – at high tide.

As you come into Pass Christian, Mississippi off Interstate 90 after crossing the Bay of St. Louis, you find yourself in a kind of daze as if passing into the Twilight Zone. Behind you is the gloomy swamps and poverty of Louisiana – before you, lining the interstate overlooking the calm blue ocean, are rows of mansions three- and four-stories high, neatly manicured yards with expensive sports cars parked in the cobblestone driveways. We were struck by the abruptness of this transition. Yet as we drove, we came upon a crowd of people gathered on the side of the street. Glancing up, we could see the largest mansion on the block – a towering monstrosity complete with narrow spires crowned with weathervanes and an immense balcony that extended over an impressive garden below. Massive flames bellowed out of its top story windows, blackening the once white wood finish. The fire department had not yet arrived and so the flames swept freely to the neighboring houses. Black smoke rose in furious columns into the clear blue sky. People had their phones out, snapping photographs. They chatted back and forth, pointing at the burning palace. Unchecked, the hungry flames would spread along the entire coastline and lay waste to the idyllic little town.

In August, the Gulf can easily exceed 100° F and we spent most of the time drenched in sweat, stooped over metal sieve plates and digging through coarse, shelly sand. Haustoriid population densities are much lower in the eastern Gulf, and this meant that each site took longer to find the numbers we needed. The going was slow. But given our proximity to the ocean at all times, at night we could relax by taking dips in the cool waters. We cooked our meals on a single

electric burner, often in the dark, and we'd sit up at night and chat about science, politics, and the politics of science. But the going was slow. And hot. The lowest point was in Melbourne, Florida, where the storm that had been following us most of the trip finally caught us, dumping heavy sheets of rain upon our tent. The campsite quickly flooded, which led to my getting electrocuted by our little box fan. Faith still finds this humorous. Me, less so.

I was actually working on the edits from the typesetter of my manuscript (Chapter I) to *Journal of Biogeography* during the trip, and the morning after we camped at Melbourne, we indulged ourselves with a hearty breakfast at a local diner with free wifi. After almost a week of sleeping in tents in the August heat and warming up meals with a burner by cell phone light, the food was heavenly. I had chicken and biscuits smothered in creamy sausage gravy. Three or four cups of coffee. We stayed until I finished my edits.

Over the next few days we drove up through Jacksonville and stopped for a few nights in Hilton Head, South Carolina, where we crashed with a fellow grad student in his parent's retirement home. We swam at the crowded beach and ate lunch at the Frosty Frog, known for their giant daiquiris. Hilton Head is a beach town that is ashamed of itself – where most embrace the ocean culture, including the tackily colored homes and scam souvenir shops – Hilton Head enforces a strict “color policy”. All the homes and businesses were earthy tones, dark greens and browns. It made the town appear drab, despite its clear affluence.

We continued north through North Carolina and into Virginia. The ocean became cooler and the nights more tolerable. The accents changed and by the time we entered New York City the sweltering heat of the Gulf was largely forgotten. By this time, we had taken to buying dry ice to keep our portable freezer running, as the power it sucked had blown the fuse in the Altima. On Staten Island, we bought a pound of dry ice from a small shop called *Piazza's*, where a man

with no eyebrows and burn scars on his arms told us that dry ice was quite safe. We collected haustoriids on Long Beach – the locals crowded around us, asking questions about every step of the sampling process and assuring us that they'd seen the amphipods before, no doubt, crawling around in the sand. We then drove up through Brooklyn (*Leaving Brooklyn? Fuggedaboutit!*) and the length of Manhattan, crossed through the Bronx and then were finally out of the city, up the Sound on our way to Cape Cod.

The ocean in New England is icy cold, even in August, and we watched seals wading in the surf not 20 meters off the beach. A French-Canadian couple filmed them as we dug through the rocky sand for amphipods. We found the largest ‘pods I’d ever seen there, some half the length of your palm, and we found them by the dozens. We did not swim here, though, but not because of the cold. Signs along the beach warned of great white sharks and their favorite snacks – wading seals. I didn’t go more than ankle deep.

In Maine, we camped on a lake and braved the frigid water in the evening after driving most of the day. Maine bleeds seamlessly into Canada – towering evergreens that stretch as far as you can see, a harsh gray landscape interrupted by the occasional bog. Signs warned of “Moose Crossings” and Tim Horton’s began to appear at every intersection. Prince Edward Island was the last sampling location. The beaches here were crimson from the dark clay that washed down from the eroded hillsides. The water was warm and there were no amphipods. But there were fields of wildflowers – rolling hills splashed with shades of yellow, red, and pink. The beauty of the island erased our feelings of failure at our last site; instead, the island was itself the destination, the end of the pilgrimage. I would not trade its red sands and sea of wildflowers, its quaint architecture and unpaved granite roads, for all the amphipods in the world.



*Left:* “Elysium”. Yellow fields of Prince Edward Island. *Right:* “A House on the Hill”. Photos by Faith Hardin.

By March, the World Health Organization had designated what is taxonomically known as SARS-CoV-2 as a global pandemic. As of this writing, 22 million people around the world have been infected and almost 800,000 have died. The Spring and Summer semesters of my 4<sup>th</sup> year were times of intense stress punctuated by seemingly endless periods of boredom. The whole world has gone into lockdown. The future is a black box wrapped in anxiety; the only constant is quarantine, the only hope a vaccine that can restore some sense of normalcy to our lives. If anything, this pandemic has laid bare the tragedy of the United States. Thinking ourselves untouchable – a mansion on a hill – we instead are bearing the brunt of the deaths, the infections, all the while insisting that we have it under control. Forcing normalcy at the cost of human life. The incompetency of our state and federal governments is rivaled only by university administrators that place their bottom-line above the health and safety of the community.

For four years I attended public defenses and seminars. I permitted myself to imagine what my own defense would be like – how my friends and family would be there, how I’d layout for them the work that I’d done the past 5 years in some sweeping story. I thought of how I would thank all of my great friends with some emotional speech on a slide with all of their faces,

and I'd see them in the audience, smiling and nodding their heads. I imagined walking across the stage at graduation, receiving my hood from Dr. Wicksten and the flurry of photographs that my family would take afterward, flashes blinding me, my mother's tears of joy, the celebratory dinner to follow.

None of that will ever happen. I will never know what it's like to give a normal public defense, to hug your friends and family after the closing slide, to physically stand in front of your committee and justify your methods. The taste of a post-defense beer with friends at the bar, all buying you shots, calling you *Doctor*. The experience of being hooded in front of your peers and family, all gathered to cheer-on your hard-fought accomplishment.

But I will always have the wildflowers of Prince Edward Island. The virus can take away my public graduation, but it cannot rob me of the nights on the roof of the field station in Calnali. It cannot claim the feeling of the warm sun on my back as we rowed through the Laguna de Términos over seagrass beds peppered with bright purple sponges. Nor the refreshingly cool ponds of Chahuaco Falls, the mist of the waterfall descending upon me in gentle clouds.

It cannot take this away from me.

*This is mine.*



“EXIT HERE.” Photo by Zach Hancock.

## REFERENCES

### CHAPTER I

- Adamkewicz, S.L. & Harasewych, M.G. (1996). Systematics and biogeography of the genus *Donax* (Bivalvia: Donacidae) in eastern North America. *American Malacological Bulletin*, 12(1/2), 97–103.
- Aharon, P. (2003). Meltwater flooding events in the Gulf of Mexico revisited: implications for rapid climate changes during the last deglaciation. *Paleoceanography*, 18, 1–14.
- Anderson, J.D. (2007). Systematics of the North American menhadens: molecular evolutionary restructures in the genus *Brevoortia* (Clupeiformes: Clupeidae). *Fishery Bulletin*, 205, 368–378.
- Avise, J.C. (1992). Molecular population structure and the biogeographic history of a regional fauna: A case history with lessons for conservation biology. *Oikos*, 62(1), 62–76.
- Bert, T.M. (1986). Speciation in western Atlantic stone crabs (genus *Menippe*): the role of geological processes and climatic events in the formation and distribution of species. *Marine Biology*, 93, 157–170.
- Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.H., Xie, D., Suchard, M.A., Rambaut, A., Drummond, A.J. (2014). BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS ONE Computational Biology*, 10(4), e1003537.
- Bousfield, E.L. (1962). New haustoriid amphipods from the Canadian Atlantic region. *Bulletin of the National Museum of Canada*, 183, 63–75.
- Bousfield, E.L. (1970). Adaptive radiation in sand-burrowing amphipod crustaceans. *Chesapeake Science*, 11(3), 143–154.
- Cahill, A.E., de Jode, A., Dubois, S., Bouzaza, Z., Aurelle, D., Boissin, E., Chabrol, O., David, R., Egea, E., Ledoux J.B., Mérigot, B., Weber A.A., Chenail, A. (2017). A multispecies approach reveals hot spots and cold spots of diversity and connectivity in invertebrate species with contrasting dispersal modes. *Molecular Ecology*, 26, 6563–6577.
- Conlan, K.E. (1991). Precopulatory mating behavior and sexual dimorphism in the amphipod Crustacea. In Watling, L. (Ed.) *VIIth International Colloquium on Amphipoda. Developments in Hydrobiology* (Vol. 4, pp. 255–282). Springer, Dordrecht.
- Copilaş-Ciocianu, D. & Petrušek, A. (2017). Phylogeography of a freshwater crustacean species complex reflects a long-gone archipelago. *Journal of Biogeography*, 44, 421–432.

- Croker, R.A. (1967). Niche diversity in five sympatric species of intertidal amphipods (Crustacea: Haustoriidae). *Ecological Monographs*, 37(3), 173–200.
- Darriba, D., Taboada, G.L., Doallo, R., Posada, D. (2012). jModelTest2: more models, new heuristics and parallel computing. *Nature Methods*, 9, 772.
- Drovetski, S.V., Semenov, G., Red'kin, Y.A., Sotnikov, V.N., Fadeev, I.V., Koblik, E.A. (2015). Effects of asymmetric nuclear introgression, introgressive mitochondrial sweep, and purifying selection on phylogenetic reconstruction and divergence estimates in the Pacific clade of *Locustella* warblers. *PLoS ONE*, 10(4), e0122590.
- Drumm, D.T. & Kreiser, B. (2012) Population genetic structure and phylogeography of *Mesokalliaipseudes mcsweenyi* (Crustacea: Tanaidacea) in the northwestern Atlantic and Gulf of Mexico. *Journal of Experimental Marine Biology and Ecology*, 412, 58–65.
- Edwards, S.V. & Beerli, P. (2000). Gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution*, 54(6), 1839–1854.
- Excoffier, L. & Lischer, H.E.L. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetic analyses under Linux and Windows. *Molecular Ecology Resources*, 10, 564–567.
- Fişer, C., Robinson, C.T., Malard, F. (2017) Cryptic species as a window into the paradigm shift of the species concept. *Molecular Ecology*, 27, 613–635.
- Fişer, Z., Altermatt, F., Zakšek, V., Knapič, T., Fişer, C. (2015) Morphologically cryptic amphipod species are “ecological clones” at regional but not at local scale: A case study of four *Niphargus* species. *PLoS ONE*, 10(7), e0134384.
- Flanagan, S.P., Rose, E., Jones, A.G. (2016). Population genomics reveals multiple drivers of population differentiation in a sex-role-reversed pipefish. *Molecular Ecology*, 25, 5043–5072.
- Foster, J.M. & LeCroy, S.E. (1991). *Haustorius jayneae*, a new species of haustoriid amphipod from the northern Gulf of Mexico, with notes on its ecology (Haustoriidae: Haustoriinae). *Gulf Research Reports*, 8(2), 189–197.
- Fu, Y.X. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, 147(2), 915–925.
- Fujisawa, T., Aswad, A., Barraclough, T.G. (2016). A rapid and scalable method for multilocus species delimitation using Bayesian model comparisons and rooted triplets. *Systematic Biology*, 65(5), 759–771.

Galloway, W.E., Whiteaker, T.L., Ganey-Curry, P. (2011). History of Cenozoic North American drainage basin evolution, sediment yield, and accumulation in the Gulf of Mexico basin. *Geosphere*, 4, 938–973.

Goudet, J. (1995). FSTAT (Version 1.2): A computer program to calculate F-statistics. *Journal of Heredity*, 86(6), 485–486.

Goulding, T.C. & Cohen, C.S. (2014). Phylogeography of a marine acanthocephalan: lack of cryptic diversity in a cosmopolitan parasite of mole crabs. *Journal of Biogeography*, 41, 965–976.

Grant, J. (1981) Dynamics of competition among estuarine sand-burrowing amphipods. *Journal of Experimental Marine Biology and Ecology*, 49, 255–265.

Hafner, M.S., Sudman, P.D., Villablanca, F.X., Spradling, T.A., Demastes, J.W., Nadler, S.A. (1994). Disparate rates of molecular evolution in cospeciating hosts and parasites. *Science*, 265(5175), 1087–1090.

Hancock, Z.B. & Wicksten, M.K. (2018). Two new species of sand-burrowing amphipods of the genus *Haustorius* Müller, 1775 (Amphipoda: Haustoriidae) from the northwestern Gulf of Mexico. *Zootaxa*, 4459(1), 101–127.

Harpending, H.C. (1994). Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Human Biology*, 66(4), 591–600.

Heled, J. & Drummond, A.J. (2010). Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution*, 27(3), 570–580.

Heller, R., Chikhi, L., Siegmund, H.R. (2013). The confounding effect of population structure on Bayesian skyline plot inferences on demographic history. *PLoS ONE*, 8(5), e62992.

Herke, S.W. & Foltz, D.W. (2002). Phylogeography of two squid (*Loligo pealei* and *L. plei*) in the Gulf of Mexico and northwestern Atlantic Ocean. *Marine Biology*, 140, 103–115.

Hou, Z., Fu, J., Li, S. (2007). A molecular phylogeny of the genus *Gammarus* (Crustacea: Amphipoda) based on mitochondrial and nuclear gene sequences. *Molecular Phylogenetics and Evolution*, 45(2), 596–611.

Huelsenbeck, J.P. & Rannala, B. (2004). Frequentist properties of Bayesian posterior probabilities of phylogenetic trees under simple and complex substitution models. *Systematic Biology*, 53, 904–913.

Huelsenbeck, J.P. & Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics Applications Note*, 17(8), 754–755.

Hurtado, L.A., Mateos, M., Santamaria, C.A. (2010). Phylogeography of supralittoral rocky intertidal *Ligia* isopods in the Pacific region from Central California to Central Mexico. *PLoS ONE*, 5(7), e11633.

Katoh, K. & Standley, D.M. (2013). MAFFT Multiple Sequence Alignment Software Version 7: Improvements in performance and U.S.A.ability. *Molecular Biology and Evolution*, 30(4), 772–780.

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, 16(2), 111–120.

Kingman, J.F.C. (1982). The coalescent. *Stochastic Processes and their Applications*, 13, 235–248.

Kinlan, B.P., Gaines, S.D., Lester, S.E. (2005). Propagule dispersal and the scales of marine community process. *Diversity and Distributions*, 11, 139–148.

Knowlton, N. & Weigt, L.A. (1998). New dates and new rates for divergence across the Isthmus of Panama. *Proceedings of the Royal Society B*, 265(1412), 2257–2263.

Kumar, S., Stecher, G., Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for bigger datasets. *Molecular Biology and Evolution*, 33(7), 1870–1874.

Lagrule, C., Wattier, R., Galipaud, M., Gauthey, Z., Rullmann, J.P., Dubreuil, C., Rigaud, T., Bollache, L. (2014). Confrontation of cryptic diversity and mate discrimination within *Gammarus pulex* and *Gammarus fossarum* species complexes. *Freshwater Biology*, 59, 2555–2570.

Lanfear, R., Frandsen, P.B., Wright, A.M., Senfeld, T., Calcott, B. (2017). PartitionFinder2: New methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Molecular Biology and Evolution*, 34(4), 772–773.

Laurindo, L., Mariano, A., Lumpkin, R. (2017). An improved near-surface velocity climatology for the global ocean drifter observations. *Deep-Sea Research I*, 124, 73–92.

Lecocq, T., Vereecken, N.J., Michez, D., Dellicour, S., Lhomme, P., Valterová, I., Rasplus, J.Y., Rasmont, P. (2013). Patterns of genetic and reproductive traits differentiation in Mainland vs. Corsican populations of bumblebees. *PLoS ONE*, 8(6), e65642.

Lefébure, T., Douady, C.J., Gouy, M., Gibert, J. (2006). Relationship between morphological taxonomy and molecular divergence within Crustacea: Proposal of a molecular threshold to help species delimitation. *Molecular Phylogenetics and Evolution*, 40, 435–447.

Leigh, J.W. & Bryant, D. (2015). PopART: Full-feature software for haplotype network construction. *Methods in Ecology and Evolution*, 6(9), 1110–1116.

Lessios, H.A., Kane, J., Robertson, D.R. (2003) Phylogeography of the pantropical sea urchin *Tripneustes*: Contrasting patterns of population structure between oceans. *Evolution*, 57(9), 2026–2036.

Lumpkin, R. & Johnson, G.C. (2013). Global ocean surface velocities from drifters: Mean, variance, El Niño-Southern Oscillation response, and seasonal cycle. *Journal of Geophysical Research: Oceans*, 118, 2992–3006.

Macdonald, K.S., Yampolsky, L., Duffy, J.E. (2005). Molecular and morphological evolution of the amphipod radiation of Lake Baikal. *Molecular Phylogenetics and Evolution*, 35(2), 323–343.

Maddison, W.P. & Maddison, D.R. (2018). Mesquite: a modular system for evolutionary analysis. Version 3.51. <http://www.mesquiteproject.org>.

Matthews, L.M., Schubart, C.D., Neigel, J.E., Felder, D.L. (2002). Genetic, ecological, and behavioural divergence between two sibling snapping shrimp species (Crustacea: Decapoda: *Alpheus*). *Molecular Ecology*, 11, 1427–1437.

McClure, M.R. & Greenbaum, I.F. (1999). Allozymic variation and biogeography of snapping shrimp (*Alpheus*) from the Gulf of Mexico and northwestern Atlantic coasts. *The Southwestern Naturalist*, 44(4), 462–469.

McClure, M.R. & McEachran, J.D. (1992). Hybridization between *Prionotus alatus* and *P. paralatus* in the northern Gulf of Mexico (Pisces: Triglidae). *Copeia*, 1992(4), 1039–1046.

Mendelssohn, I.A., Byrnes, M.R., Kneib, R.T., Vittor, B.A. (2017) Coastal habitats of the Gulf of Mexico. In C. Herb Ward (Ed.), *Habitats and Biota of the Gulf of Mexico: Before the Deepwater Horizon Oil Spill* (Vol. 1, pp. 359–582). Houston, TX: Springer.

Miller, M.A., Pfeiffer, W., Schwartz, T. (2010). Creating the CIPRES Science Gateway for inference of large phylogenetic trees. *Proceedings of the Gateway Computing Environments Workshop*, 1–8.

Portnoy, D.S. & Gold, J.R. (2012). Evidence of multiple vicariance in a marine suture-zone in the Gulf of Mexico. *Journal of Biogeography*, 39, 1499–1507.

Puillandre, N., Lambert, A., Brouillet, S., Achaz, G. (2012). ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Molecular Ecology*, 21: 1864–1877.

Rambaut, A. (2011). FigTree, version 1.4.3. Available at: <https://github.com/rambaut/figtree/releases>.

Rambaut, A., Drummond, A.J., Xie, D., Baele, G., Suchard, M.A. (2018). Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*, 67(5), 901–904.

Revsbech, N.P., Madsen, B., Jorgensen, B.B. (1980). Oxygen in the sea bottom measured with a microelectrode. *Science*, 207, 1355–1356.

Roberts, H.H. (1997). Dynamics changes of the Holocene Mississippi River Delta plain: The delta cycle. *Journal of Coastal Research*, 13(3), 605–627.

Roberston, P.B. & Shelton, C.R. (1980). *Lepidactylus triarticulatus* n. sp., a new haustoriid amphipod from the northern Gulf of Mexico. *Gulf Research Reports*, 6(4), 415–420.

Rosas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J.C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S.E., Sánchez-Gracia, A. (2017). DnaSP 6: DNA Sequence Polymorphism Analysis of Large Datasets. *Molecular Biology and Evolution*, 34: 3299–3302.

Rousset, F. (1997). Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. *Genetics*, 145(4), 1219–1228.

Sameoto, D.D. (1969). Some aspects of the ecology and life cycle of three species of subtidal sand-burrowing amphipods (Crustacea: Haustoriidae). *Journal Fisheries Research Board of Canada*, 26 (5), 1321–1345.

Saunders, I.W., Tavaré, S., Watterson, G.A. (1984). On the genealogy of nested subsamples from a haploid population. *Advances in Applied Probability*, 16(3), 471–491.

Schneider-Broussard, R., Felder, D.L., Chlan, C.A., Neigel, J.E. (1998). Tests of phylogeographic models with nuclear and mitochondrial DNA sequence variation in the stone crabs, *Menippe adina* and *Menippe mercenaria*. *Evolution*, 52(6), 1671–1678.

Shelton, C.R. & Robertson, P.B. (1981). Community structure of intertidal macrofauna on two surf-exposed Texas sandy beaches. *Bulletin of Marine Science*, 31(4), 833–842.

Simpson, C.T. (1900). On the evidence of the Unionidae regarding the former courses of the Tennessee and other southern rivers. *Science*, 12(291), 133–136.

Slatkin, M. (1991). Inbreeding coefficients and coalescence times. *Genetics Research*, 58, 167–175.

Slatkin, M. (1995). A measure of population subdivision based on microsatellite allele frequencies. *Genetics*, 139, 457–462.

Stamatakis, A. (2014). RAxML Version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30(9), 1312–1313.

Sutherland, D.L., Hogg, I.D., Waas, J.R. (2009). Phylogeography and species discrimination in the *Paracalliope fluvialis* species complex (Crustacea: Amphipoda): can morphologically

similar heterospecifics identify compatible mates? *Biological Journal of the Linnean Society*, 99, 196–205.

Sukumaran J. & Holder, M.T. (2010). DendroPy: a Python library for phylogenetic computing. *Bioinformatics*, 26, 1569–1571.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, 12, 585–595.

Takada, Y., Sakuma, K., Fujii, T., Kojima, S. (2018). Phylogeography of the sandy beach amphipod *Haustorioides japonicus* along the Sea of Japan: Paleogeographical signatures of cryptic divergences. *Estuarine, Coastal and Shelf Science*, 200, 19–30.

Tamura, K. & Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution*, 10(3), 512–526.

Vergilino, R., Dionne, K., Nozais, C., Dufresne, F., Belzile, C. (2012). Genome size differences in *Hyalella* cryptic species. *Genome*, 55, 134–139.

Villesen, P. (2007). FaBox: an online toolbox for FASTA sequences. *Molecular Ecology Notes*, 7, 965–968.

Wellborn, G.A. & Cothran, R.D. (2003) Phenotypic similarity and differentiation among sympatric cryptic species in a freshwater amphipod species complex. *Freshwater Biology*, 49(1), 1–13.

Witmer, A.D. & Roelke, D.L. (2014). Human interference prevents recovery of infaunal beach communities from hurricane disturbance. *Ocean & Coastal Management*, 87, 52–60.

Witt, J.D.S. & Herbert P.D.N. (2000). Cryptic species diversity and evolution in the amphipod genus *Hyalella* within central glaciated North America: a molecular phylogenetic approach. *Canadian Journal of Fisheries & Aquatic Sciences*, 57, 687–698.

Witt, J.D.S., Blinn, D.W., Hebert, P.D.N. (2003) The recent evolutionary origin of the phenotypically novel amphipod *Hyallela montezuma* offers an ecological explanation for morphological stasis in a closely allied species complex. *Molecular Ecology*, 12(2), 405–413.

Zavala-Hidalgo, J., Morey, S.L., O'Brien, J.J. (2003) Seasonal circulation on the western shelf of the Gulf of Mexico using a high-resolution numerical model. *Journal of Geophysical Research*, 108(C12), 3389.

Zhang, J., Kapli, P., Pavlidis, P., Stamatakis, A. (2013). A general species delimitation method with applications to phylogenetic placements. *Bioinformatics*, 29(22), 2869–2876.

## CHAPTER II

- Alfsnes, K, Leinaas, HP, Hessen, DO. 2017 Genome size in arthropods; different roles of phylogeny, habitat and life history in insects and crustaceans. *Ecology and Evolution*, 7: 5939–5947.
- Arnqvist, G, Sayadi, A, Immonen, E, Hotzy, C, Rankin, D, Tuda, M, Hjelmen, CE, Johnston, JS. 2015 Genome size correlates with reproductive fitness in seed beetles. *Proceedings of the Royal Society B: Biological Sciences*, 282(1815): 20151421.
- Atkinson, D & Sibly, RM. 1997 Why are organisms usually bigger in colder environments? Making sense of a life history puzzle. *Trends in Ecology & Evolution*, 12(6): 235–239.
- Barrón, MG, Fiston-Lavier, AS, Petrov, DA, González, J. 2014 Population genomics of transposable elements in *Drosophila*. *Annual Review of Genetics*, 48, 561-581.
- Bert, TM 1986. Speciation in western Atlantic stone crabs (genus *Menippe*): The role of geological processes and climatic events in the formation and distribution of species. *Marine Biology*, 93, 157–170.
- Bouckaert, R, Vaughan, TG, Barido-Sottani, J, Duchene, S, Fourment, M, Gavryushkina, A, et al. 2019 BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, 15(4): e1006650.
- Bousfield, EL 1970. Adaptive radiation in sand-burrowing amphipod crustaceans. *Chesapeake Science*, 11(3), 143–154.
- Cavalier-Smith, T. 1978. Nuclear volume control by nucleoskeletal DNA, selection for cell volume and cell growth rate, and the solution of the DNA C-value paradox. *Journal of Cell Science*, 34:247–278.
- Cheng, J & Chang, ES. 1994 Determinants of postmolt size in the American lobster (*Homarus americanus*). II. Folding of premolt cuticle. *Canadian Journal of Fisheries and Aquatic Sciences*, 51(8): 1774–1779.
- Davidowitz, G & Nijhout, HF. 2004 The physiological basis of reaction norms: The interaction among growth rate, the duration of growth, and body size. *Integrative & Comparative Biology*, 44(6): 443–449.
- Folmer, O, Black, M, Hoeh, W, Lutz, R, Vrijenhoek, R. 1994 DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, 3(5): 294–299.
- Gregory, TR 2005 *The Evolution of the Genome*. Academic Press, 768 pp.
- Gregory, TR. 2020 The animal genome size database. <http://www.genomesize.com>.

Gregory, TR & Johnston, JS. 2008 Genome size diversity in the family Drosophilidae. *Heredity*, 101: 228–238.

Hancock, ZB & Wicksten, MK. 2018 Two new species of sand-burrowing amphipods of the genus *Haustorius* Müller, 1775 (Haustoriidae: Amphipoda) from the northwestern Gulf of Mexico. *Zootaxa*, 4459(1), 101–127.

Hancock, ZB, Hardin, FO, Light, JE. 2019 Phylogeography of sand-burrowing amphipods (Haustoriidae) supports an ancient suture zone in the Gulf of Mexico. *Journal of Biogeography*, 46(11): 2532–2547.

Hare, E & Johnston, JS. 2011 Genome size determination using flow cytometry of propidium iodide-stained nuclei. In: Orgogozo, V., Rockman, M. (eds) *Molecular Methods for Evolutionary Genetics: Methods in Molecular Biology (Methods and Protocols)*, vol. 772. Humana Press, pp. 3–12.

Hidalgo O, Pellicer J, Christenhusz M, Schneider H, Leitch AR, Leitch, IJ. 2017 Is there an upper limit to genome size? *Trends in Plant Science*, 22(7): 567–573.

Hloušková, P, Mandáková, T, Pouch, M, Trávníček, P, Lysak MA. 2019 The large genome size variation in the *Hesperis* clade was shaped by the prevalent proliferation of DNA repeats and rarer genome downsizing. *Annals of Botany*, 124(1): 103–120.

Holmquist, JG, Powell, GVN, Sogard, SM. 1989 Decapod and stomatopod assemblages on a system of seagrass-covered mud banks in Florida Bay. *Marine Biology*, 473–483.

Hudson, RR, Boos, DD, Kaplan, NL. 1992 A statistical test for detecting geographic subdivision. *Molecular Biology and Evolution*, 9(1): 138–151.

Hultgren, KM, Jeffery, NW, Moran, A, Gregory, TR. 2018 Latitudinal variation in genome size in crustaceans. *Biology Journal of the Linnean Society*, 123: 348–359.

Jeffery, NW, Yampolsky, L, Gregory, TR. 2016 Nuclear DNA content correlates with depth, body size, and diversification rate in amphipod crustaceans from ancient Lake Baikal, Russia. *Genome*, 60: 303–309.

Johnston, JS, Bernardini, A, Hjelmen, CE. 2019. Genome size estimation and quantitative cytogenetics in insects. In *Insect Genomics* (pp. 15–26). Humana Press, New York, NY.

Kelley, JL, Peyton, JT, Fiston-Lavier, AS, Teets, NM, Yee, MC, Johnston, JS, Bustamante, CD, Lee, RE, Denlinger, DL. 2014 Compact genome of the Antarctic midge is likely an adaptation to an extreme environment. *Nature Communications*, 5: 4611.

Knowlton, N, & Weigt, LA 1998. New dates and new rates for divergence across the Isthmus of Panama. *Proceedings of the Royal Society B*, 265(1412), 2257–2263.

Kozlowski, J, Konarzewski, M, Gawelczyk, AT. 2003 Cell size as a link between noncoding DNA and metabolic rate scaling. *PNAS*, 100(24): 14080–14085.

Lanfear, R, Frandsen, PB, Wright, AM, Senfeld, T, Calcott, B. 2017 PartitionFinder2: New methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Molecular Biology and Evolution*, 34(4), 772–773.

LeCroy, SE 2002. *An illustrated identification guide to the nearshore marine and estuarine gammaridean amphipoda of Florida* (Vol. 2, p. 224). Tallahassee, FL: State of Florida Department of Environmental Protection.

Lefébure, T, Morvan, C, Mallard, F, François, C, Konecny-Dupré, L, Guéguen, L, Weiss-Gayet, M, Seguin-Orlando, A, Ermini, L, Sarkissian, CD, Charrier, NP, Eme, D, Mermilliod-Blondin, F, Duret, L, Vieira, C, Orlando, L, Douady, CJ. 2017 Less effective selection leads to larger genomes. *Genome Research*, 27: 1016–1028.

Lower, SS, Johnston, JS, Stanger-Hall, KF, Hjelman, CE, Hanrahan, SJ, Korunes, K, Hall, D. 2017 Genome size in North American fireflies: Substantial variation likely driven by neutral processes. *Genome Biology & Evolution*, 9(6): 1499–1512.

Lynch, M. 2007 *Origins of Genome Architecture*. Sinauer Associates Inc., 340 pp.

Lynch, M. 2011 Statistical inference on the mechanisms of genome evolution. *PLoS Genetics*, 7(6): e1001389.

Lynch, M & Connery, JS. 2003 The origins of genome complexity. *Science*, 302(5649): 1401–1404.

Lyu, H, He, Z, Wu, CI, Shi, S. 2018 Convergent adaptive evolution in marginal environments: unloading transposable elements as a common strategy among mangrove genomes. *New Phytologist*, 217(1): 428–438.

Maddison, WP, & Maddison, DR. 2018. Mesquite: A modular system for evolutionary analysis. Version 3.51. <http://www.mesquiteproject.org>.

McClure, MR, & Greenbaum, IF. 1999 Allozymic variation and bio- geography of snapping shrimp (*Alpheus*) from the Gulf of Mexico and northwestern Atlantic coasts. *The Southwestern Naturalist*, 44(4): 462–469.

Misof, B. 2019 Diversity and evolution of the transposable element repertoire in arthropods with particular reference to insects. *BMC Evolutionary Biology*, 19(1): 11.

Novák, P, Neumann, P, Macas, J. 2010 Graph-based clustering and characterization of repetitive sequences in next-generation sequencing data. *BMC Bioinformatics*, 11: 378.

- Pagel, M 1999. Inferring the historical patterns of biological evolution. *Nature*, 401: 877–884.
- Paradis, E, Schleip, K. 2019 ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, 35: 526–528.
- Orme, D, Freckleton, R, Thomas, G, Petzoldt, T, Fritz, S, Isaac, N, Pearse, W. 2018 caper: Comparative analyses of phylogenetics and evolution in R. <https://cran.r-project.org/web/packages/caper/caper.pdf>.
- Rambaut, A. 2011 FigTree, version 1.4.3. <https://github.com/rambaut/figtree/releases>.
- Rambaut, A, Drummond, AJ, Xie, D, Baele, G, & Suchard, MA. 2018 Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*, 67(5), 901–904.
- Revell, LJ. 2012 phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, 3: 217–223.
- Revsbech, NP, Madsen, B, Jorgensen, BB. 1980 Oxygen in the sea bottom measured with a microelectrode. *Science*, 207, 1355–1356.
- Roebuck, K. 2017 Nuclear genome size diversity of marine invertebrate taxa using flow cytometric analysis. *Lopez Laboratory: Genomics & Microbial Data*. 2. [https://nsuworks.nova.edu/lopez\\_lab/2](https://nsuworks.nova.edu/lopez_lab/2)
- Rosas, J, Ferrer-Mata, A, Sánchez-DelBarrio, JC, Guirao-Rico, S, Librado, P, Ramos-Onsins, SE, & Sánchez-Gracia, A. 2017 DnaSP 6: DNA sequence polymorphism analysis of large datasets. *Molecular Biology and Evolution*, 34, 3299–3302.
- Ross, E & Behringer, D. 2019 Changes in temperature, pH, and salinity affect the sheltering responses of Caribbean spiny lobsters to chemosensory cues. *Scientific Reports*, 9: 4375.
- Sessegolo, C, Burlet, N, Haudry, A. 2016 Strong phylogenetic inertia on genome size and transposable element content among 26 species of flies. *Biology Letters*, 12(8): 20160407.
- Takada, Y, Sakuma, K, Fujii, T, & Kojima, S. 2018 Phylogeography of the sandy beach amphipod *Haustorioides japonicus* along the Sea of Japan: Paleogeographical signatures of cryptic divergences. *Estuarine, Coastal and Shelf Science*, 200, 19–30.
- Takolander, A, Leskinen, E, Cabeza, M. 2017 Synergistic effects of extreme temperature and low salinity on foundational macroalga *Fucus vesiculosus* in the northern Baltic Sea. *Journal of Experimental Marine Biology and Ecology*, 495: 110–118.
- Tiersch, TR & Wachtel, SS. 1991 On the evolution of genome size in birds. *Journal of Heredity*, 82(5): 363–368.

Twombly, S & Tisch, N. 2000 Body size regulation in copepod crustaceans. *Oecologia*, 122: 318–326.

Vergilino, R, Dionne, K, Nozais, C, Dufresne, F, Belzile, C. 2011 Genome size differences in *Hyallela* cryptic species. *Genome*, 55: 134–139.

Walters, RJ & Hassall, M. 2005 The temperature-size rule in ectotherms: May a general explanation exist after all? *The American Naturalist*, 167(4): 510–523.

Whitney, KD & Garland, T. 2010 Did genetic drift drive increases in genome complexity? *PLoS Genetics*, 6(8): e1001080.

### CHAPTER III

Aguillon, S.M., Fitzpatrick, J.W., Bowman, R., Schoech, S.J., Clark, A.G., Coop, G., and N. Chen. 2017. Deconstructing isolation-by-distance: The genomic consequences of limited dispersal. *PLoS Genet.* 13(8): e1006911.

Angelis, K. and M. Dos Reis. 2015. The impact of ancestral population size and incomplete lineage sorting on Bayesian estimation of species divergence times. *Current Zoology*. 61(5): 874–885.

Arbogast, B.S., Edwards, S.V., Wakeley, J., Beerli, P., Slowinski, J.B. 2002. Estimating divergence times from molecular data on phylogenetic and population genetic timescales. *Ann. Rev. Ecol. Syst.* 33: 707–740.

Aris-Brosou, S., and L. Excoffier. 1996. The impact of population expansion and mutation rate heterogeneity on DNA sequence polymorphism. *Mol. Biol. Evo.* 13(3): 494–504.

Bocxlaer, I.V., Roelants, K., Biju, S.D., Nagaraju, J., Bossuyt, F. 2006. Late Cretaceous vicariance in Gondwanan amphibians. *PloS ONE*, 1(1): e74.

Bouckaert, R. 2010. DensiTree: making sense of sets of phylogenetic trees. *Bioinformatics*, 26(10): 1372–1373.

Bouckaert, R., and Bryant, D. 2015. A rough guide to SNAPP. <https://www.beast2.org/snapp/>

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C-H., Xie, D., Suchard, M.A., Rambaut, A., and Drummond, A.J. (2014). BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLoS Comp. Bio.* 10(4): e1003537.

Broquet, T., Ray, N., Petit, E., Fryxell, J.M., and F. Burel. 2006. Genetic isolation by distance and landscape connectivity in the American marten (*Martes americana*). *Land. Eco.* 21: 877–889.

- Brown, J.W., and S.A. Smith. 2018. The past sure is tense: On interpreting phylogenetic divergence time estimates. *Syst. Biol.* 67(2): 340–353.
- Bryant, D., Bouckaert, R., Felsenstein, J., Rosenberg, N.A. and A. RoyChoudhury. 2012. Inferring species trees directly from biallelic genetic markers: bypassing gene trees in a full coalescent analysis. *Mol. Bio. Evo.* 29(8): 1917–1932.
- Ceballos, G., Ehrlich, P.R., and R. Dirzo. 2017. Biological annihilation via the ongoing sixth mass extinction signaled by vertebrate population losses and declines. *PNAS.* 114(30): E6089–E6096.
- Coyne, J. and H.A. Orr. 2004. *Speciation*. Sinauer Associates, Inc. Sunderland, Massachusetts, 545 pp.
- Dobzhansky, T. 1973. Nothing in biology makes sense except in the light of evolution. *American Biology Teacher.* 35: 125–129.
- Domingos, F.M.C.B, Colli, G.R., Lemmon, A., Lemmon, E.M., Beheregaray, L.B. 2017. In the shadows: Phylogenomics and coalescent species delimitation unveil cryptic diversity in a Cerrado endemic lizard (Squamata: *Tropidurus*). *Mol. Phylo. Evo.*, 107: 455–465.
- Douzery, E.J.P., Snell, E.A., Bapteste, E., Delsuc, F., and H. Philippe. 2004. The timing of eukaryotic evolution: Does a relaxed molecular clock reconcile proteins and fossils? *PNAS.* 101(43): 15386–15391.
- Dowle, E.J., Bracewell, R.R., Pfrender, M.E., Mock, K.E., Bentz, B.J., Ragland, G.J. 2017. Reproductive isolation and environmental adaptation shape the phylogeography of mountain pine beetle (*Dendroctonus ponderosae*). *Mol. Ecol.* 26(21): 6071–6084.
- Edwards, S.V., and P. Beerli. 2000. Gene divergence, population divergence, and the variance in coalescent time in phylogeographic studies. *Evolution.* 54(6): 1839–1854.
- Excoffier, L., Dupanloup, I., Huerta-Sánchez, E., Sousa, V.C., and M. Foll. 2013. Robust demographic inference from genomic and SNP data. *PLoS Genet.* 9(10): e1003905.
- Fang, B., Merilä, J., Matschiner, M., Momigliano, P. 2020. Estimating uncertainty in divergence times among three-spined stickleback clades using the multispecies coalescent. *Mol. Phylogen. Evol.* 142: 106646.
- Gaggiotti, O.E., and L. Excoffier. 2000. A simple method for removing the effect of a bottleneck and unequal population sizes on pairwise genetic distances. *Proc. B.* 267(1438): 81–87.
- Haller, B.C., and P.W. Messer. 2019. SLiM 3: Forward genetic simulations beyond the Wright-Fisher Model. *Mol. Bio. Evo.* 36(3): 632–637.

- Haller, B.C., Galloway, J., Kelleher, J., Messer, P.W., and P.L. Ralph. 2019. Tree-sequence recording in SLiM opens new horizons for forward-time simulation of whole genomes. Mol. Eco. Res. 19(2): 552–566.
- Hancock, Z.B., Hardin, F.O., Light, J.E. 2019. Phylogeography of sand-burrowing amphipods (Haustoriidae) supports an ancient suture zone in the Gulf of Mexico. J. Biogeogr. 46(11): 2532–2547.
- Hare, M.P., Nunney, L., Schwartz, M.K., Ruzzante, D.E., Burford, M., Waples, R.S., Ruegg, K., and F. Palstra. 2011. Understanding and Estimating Effective Population Size for Practical Application in Marine Species Management. 25(3): 438–449.
- Heath, T.A., Hueslenbeck, J.P., and T. Stadler. 2014. The fossilized birth-death process for coherent calibration of divergence-time estimates. PNAS. E2957-E2966.
- Herman, A., Brandvain, Y., Weagley, J., Jeffery, W.R., Keene, A.C., Kono, T.J.Y., Bilandžija, H., Borowsky, R., Espinasa, L., O'Quin, K., Ornelas-García, C.P., Yoshizawa, M., Carlson, B., Maldonado, E., Gross, J.B., Cartwright, R.A., Rohner, N., Warren, W.C., and S.E. McGaugh. 2018. The role of gene flow in rapid and repeated evolution of cave-related traits in Mexican tetra, *Astyanax mexicanus*. Mol. Eco. 27(22): 4397–4416.
- Hey, J. 1991. A multi-dimensional coalescent process applied to multi-allelic selection models and migration models. Theo. Pop. Bio. 39: 30–48.
- Kelleher, J., Etheridge, A.M., and G. McVean. 2016. Efficient Coalescent Simulation and Genealogical Analysis for Large Sample Sizes. PLoS Comput. Biol. 12(5): e1004842.
- Kimura, M. 1953. “Stepping Stone” model of population. Ann. Rept. Nat. Inst. Genetics, Japan. 3: 62–63.
- Kimura, M. 1968. Evolutionary rate at the molecular level. Nature. 217: 624–626.
- Kimura, M. 1983. The Neutral Theory of Molecular Evolution. Cambridge Univ. Pres., Cambridge, UK.
- Kimura, M. and G. Weiss. 1964. The stepping stone model of population structure and the decrease of genetic correlation with distance. Genetics. 49: 561–576.
- King, J.L., and T.H. Jukes. 1969. Non-Darwinian evolution. Science. 164: 788–798.
- Knowlton, N. and L.A. Weigt. 1998. New dates and new rates for divergence across the Isthmus of Panama. Proc. B. 265(1412): 2257–2263.
- Kubatko, L.S., and J.H. Degnan. 2007. Inconsistency of Phylogenetic Estimates from Concatenated Data under Coalescence. Sys. Bio. 56(1): 17–24.

- Leaché, A., and Bouckaert, R. 2018. Species Trees Estimation with SNAPP: A Tutorial and Example. Workshop on Population and Speciation Genomics, Český Krumlov.
- Lepage, T., Bryant, D., Philippe, H., and N. Lartillot. 2007. A general comparison of relaxed molecular clock models. *Mol. Biol. Evo.* 24(12): 2669–2680.
- Leslie, M.S. and P.A. Morin. 2018. Structure and phylogeography of two tropical predators, spinner (*Stenella longirostris*) and pantropical spotted (*S. attenuata*) dolphins, from SNP data. *R. Soc. 5:171615.*
- Li, J., Huang, J-P., Sukumaran, J., and L.L. Knowles. 2018. Microevolutionary processes impact macroevolutionary patterns. *BMC Evo. Bio.* 18: 123.
- Lynch, M. 1990. The rate of morphological evolution in mammals from the standpoint of the neutral expectation. *Am. Nat.* 136(6): 727–741.
- Lynch, M. 2007. *The Origins of Genome Architecture*. Sinauer Associates, Inc. Sunderland, MA.
- Magallón, S. 2010. Using Fossils to Break Long Branches in Molecular Dating: A Comparison of Relaxed Clocks Applied to the Origin of Angiosperms. *Sys. Bio.* 59(4): 384–399.
- Malécot, G. 1968. *The Mathematics of Heredity*. Translated from the French edition (Paris, 1948). Ed. Yermanos, D.M. Freeman, San Francisco, 1969.
- Manthey, J.D., Klicka, J., Spellman, G.M. 2015. Chromosomal patterns of diversity and differentiation in creepers: a next-gen phylogeographic investigation of *Certhia americana*. *Heredity.* 115: 165–172.
- Manthey, J.D., Moyle, R.G., Gawin, D.F., Rahman, M.A., Ramji, M.F.S., Sheldon, F.H. 2017. Genomic phylogeography of the endemic Mountain Black-eye of Borneo (*Chlorocharis emiliae*): montane and lowland populations differ in patterns of Pleistocene diversification. *J. Bio.* 44(10): 2272–2283.
- Maruyama, T. 1970a. Effective number of alleles in a subdivided population. *Theo. Pop. Bio.* 1: 273–306.
- Maruyama, T. 1970b. The rate of decrease of heterozygosity in a population occupying a circular or linear habitat. *Genetics.* 67: 437–454.
- Maruyama, T. 1971. Analysis of population structure: II. Two-dimensional stepping stone models of finite length and other geographically structured populations. *Ann. Hum. Gen., Lon.* 35: 179–196.
- Meirmans, P.G. 2012. The trouble with isolation by distance. *Mol. Eco.* 21(12): 2839–2846.

Miller, M.A., Pfeiffer, W., and T. Schwartz. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. Proceedings of the Gateway Computing Environments Workshop (GCE). New Orleans, LA.

Nei, M., and W.H. Li. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. PNAS. 76(10): 5269-5273.

Nei, M., and N. Takahata. 1993. Effective population size, genetic diversity, and coalescence time in subdivided populations. J. Mol. Evo. 37: 240–244.

Peterson, M.A., and R.F. Denno. 1998. The influence of dispersal and diet breadth on patterns of genetic isolation by distance in phytophagous insects. Am. Nat. 152(3): 428–446.

Pond, S.K., and S.V. Muse. 2005. Site-to-site variation of synonymous substitution rates. Mol. Biol. Evo. 22(12): 2375–2385.

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: <https://www.R-project.org/>.

Riddle, B.R., Hafner, D.J., Alexander, L.F., Jaeger, J.R. 2000. Cryptic vicariance in the historical assembly of a Baja California Peninsular Desert biota. PNAS, 97(26): 14438–14443.

Rieman, B.E., and F. W. Allendorf. 2011. Effective Population Size and Genetic Conservation Criteria for Bull Trout. North American Journal of Fisheries Management. 21(4): 756–764.

Rosenberg, N.A., and M.W. Feldman. 2002. The relationship between coalescent times and population divergence times. In: Modern Developments in Theoretical Population Genetics. Eds. Slatkin, M. and M. Veuille. Oxford Univ. Pres. New York, NY.

Slatkin, M. 1991. Inbreeding coefficients and coalescence times. Genet. Res., Camb. 58: 167–175.

Slatkin, M. 2004. Seeing ghosts: the effect of unsampled populations on migration rates estimated for sampled populations. Mol. Eco. 14(1): 67–73.

Spalink, D., MacKay, R., Sytsma, K.J. 2019. Phylogeography, population genetics and distribution modelling reveal vulnerability of *Scirpus longii* (Cyperaceae) and the Atlantic Coastal Plain Flora to climate change. Mol. Ecol. 28(8): 2046–2061.

Stange, M., Sánchez-Villagra, M.R., Salzburger, W., and M. Matschiner. 2018. Bayesian divergence-time estimation with genome-wide single-nucleotide polymorphism data of sea catfish (Ariidae) supports Miocene closure of the Panamanian Isthmus. Sys. Bio. 67(4): 681–699.

- Turner, T.F., Wares, J.P., and J.R. Gold. 2002. Genetic Effective Size Is Three Orders of Magnitude Smaller Than Adult Census Size in an Abundant, Estuarine-Dependent Marine Fish (*Sciaenops ocellatus*). *Genetics*. 162(3): 1329–1339.
- Wakeley, J. 1998. Segregating sites in Wright's island model. *Theoret. Pop. Biol.* 53:166–175.
- Wakeley, J. 1999. Non-equilibrium migration in human history. *Genetics*. 153:1863–1871.
- Wakeley, J. 2000. The effects of subdivision on the genetic divergence of populations and species. *Evolution*. 54(4): 1092–1101.
- Weir, J.T., and D. Schluter. 2008. Calibrating the avian molecular clock. *Mol. Eco.* 17: 2321–2328.
- Whitlock, M.C. and D.E. McCauley. 1999. Indirect measures of gene flow and migration: FST not equal to  $1/(4Nm + 1)$ . *Heredity*. 82: 117–125.
- Wilkins, J.F., and J. Wakeley. 2002. The coalescent in a continuous, finite, linear population. *Genetics*. 161: 873–888.
- Wright, S. 1931. Evolution in Mendelian populations. *Genetics*. 16(2): 97–159.
- Wright, S. 1943. Isolation by distance. *Genetics*. 28:114–138.
- Yu, G., Smith, D.K., Zhu, H., Guan, Y., and T.T.Y. Lam. 2017. ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Meth. Eco. Evo.* 8(1): 28–36.
- ## CHAPTER IV
- Allen B, Sample C, Dementieva Y, Medeiros RC, Paoletti C, Nowak MA (2015) The molecular clock of neutral evolution can be accelerated or slowed by asymmetric spatial structure. *PLoS Computational Biology*, 11(2): e1004108.
- Árnason E, Halldórsdóttir K (2019) Codweb: Whole-genome sequencing uncovers extensive reticulations fueling adaptation among Atlantic, Arctic, and Pacific gadids. *Science Advances*, 5(3): eaat8788.
- Avise JC (1992) Molecular population structure and the biogeographic history of a regional fauna: a case history with lessons for conservation biology. *OIKOS*, 63: 62–76.
- Battey CJ, Ralph PL, Kern AD (2020) Space is the Place: Effects of continuous spatial structure on analysis of population genetic data. *Genetics*, 215(1): 193–214.

Bert TM, Harrison RG (1988) Hybridization in western Atlantic stone crabs (genus *Meippe*): Evolutionary history and ecological context influences species interactions. *Evolution*, 42(3): 528–544.

Blackmon H, Adams RH (2015) evobiR: Comparative and population genetic analyses. <http://coleoguy.github.io/#resources>.

Bowen BW, Avise JC (1990) Genetic structure of Atlantic and Gulf of Mexico populations of sea bass, menhaden, and sturgeon: Influence of zoogeographic factors and life-history patterns. *Marine Biology*, 107: 371–381.

Bouckaert RR (2010) DensiTree: making sense of sets of phylogenetic trees. *Bioinformatics*, 26(10): 1372–1373.

Bradburd GS, Ralph PL (2019) Spatial population genetics: It's about time. *Annual Review of Ecology, Evolution, and Systematics*, 50: 427–449.

Bradburd GS, Coop GM, Ralph PL (2018) Inferring continuous and discrete population genetic structure across space. *Genetics*, 210(1): 33–52.

Bromham L, Duchêne S, Hua X, Ritchie AM, Duchêne DA, Ho SYW (2017) Bayesian molecular dating: opening up the black box. *Biological Reviews*, 93(2): 1165–1191.

Brown JW, Smith SA (2018) The past sure is tense: On interpreting phylogenetic divergence time estimates. *Systematic Biology*, 67(2): 340–353.

Bryant D, Bouckaert R, Felsenstein J, Rosenberg NA, RoyChoudhury A (2012) Inferring species trees directly from biallelic genetic markers: bypassing gene trees in a full coalescent analysis. *Molecular Biology and Evolution*, 29(8): 1917–1932.

Carstens BC, Pelletier TA, Reid NM, Satler JD (2013) How to fail at species delimitation. *Molecular Ecology*, 22(17): 4369–4383.

Castellanos-Morales G, Gámez N, Casillo-Gámez RA, Eguiarte LE (2016) Peripatric speciation of an endemic species driven by Pleistocene climate change: The case of the Mexican prairie dog (*Cynomys mexicanus*). *Molecular Phylogenetics and Evolution*, 94: 171–181.

Chambers AE, Hillis DM (2020) The multispecies coalescent over-splits species in the case of geographically widespread taxa. *Systematic Biology*, 69(1): 184–193.

Chifman J, Kubatko L (2014) Quartet inference from SNP data under the coalescent model. *Bioinformatics*, 30(23): 3317–3324.

Coyne JA, Orr HA (2004) *Speciation*. Sinauer Associates: Sunderland, MA.

Cui R, Schumer M, Kruesi K, Walter R, Andolfatto P, Rosenthal GG (2013) Phylogenomics reveals extensive reticulate evolution in *Xiphophorus* fishes. *Evolution*, 67(8): 2166–2179.

Dawkins & Wong (2016) *The Ancestor's Tale: The Dawn of Evolution*. Mariner Books: New York, NY.

de Queiroz K (2007) Species concepts and species delimitation. *Systematic Biology*, 56(6): 879–886.

DeGiorgio M, Rosenberg NA (2016) Consistency and inconsistency of consensus methods for inferring species trees from gene trees in the presence of ancestral population structure. *Theoretical Population Biology*, 110: 12–24.

Degnan JH (2018) Modeling hybridization under the network multispecies coalescent. *Systematic Biology*, 67(5): 687–799.

Degnan JH, Rhodes JA (2015) There are no caterpillars in a wicked forest. *Theoretical Population Biology*, 105: 17–23.

Degnan JH, Rosenberg NA (2006) Discordance of species trees with their most likely gene trees. *PLoS Genetics*, 2(5): e68.

Degnan JH, Rosenberg NA (2009) Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology & Evolution*, 24(6): 332–340.

Domingos FMCB, Colli GR, Lemmon A, Lemmon EM, Beheregaray LB (2017) In the shadows: Phylogenomics and coalescent species delimitation unveil cryptic diversity in a Cerrado endemic lizard (Squamata: *Tropidurus*). *Molecular Phylogenetics and Evolution*, 107: 455–465.

Donald PF, Greenwood JJD (2008) Spatial patterns of range contraction in British breeding birds. *Ibis*, 143(3): 593–601.

Duchêne S, Lanfear R, Ho SYW (2014) The impact of calibration and clock-model choice on molecular estimates of divergence times. *Molecular Phylogenetics and Evolution*, 78: 277–289.

Durand EY, Patterson N, Reich D, Slatkin M (2011) Testing for ancient admixture between closely related populations. *Molecular Biology and Evolution*, 28(8): 2239–2252.

Edwards SV, Beerli P (2000) Gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution*, 54(6): 1839–1854.

Eriksson A, Manica A (2012) Effect of ancient population structure on the degree of polymorphism shared between modern human populations and ancient hominids. *PNAS*, 109(35): 13956–13960.

- Etienne RS, Morlon H, Lambert A (2014) Estimating the duration of speciation from phylogenies. *Evolution*, 68(8): 2430–2440.
- Etienne RS, Rosindell J (2012) Prolonging the past counteracts the pull of the present: Protracted speciation can explain observed slowdowns in diversification. *Systematic Biology*, 61(2): 204.
- Fisher RA (1923) XXI. – On the Dominance Ratio. *Proceedings of the Royal Society of Edinburgh*, 42: 321–341.
- Flannagan SP, Rose E, Jones AG (2016) Population genomics reveals multiple drivers of population differentiation in a sex-role-reversed pipefish. *Molecular Ecology*, 5043–5072.
- Fujisawa T, Aswad A, Barraclough TG (2016) A rapid and scalable method for multilocus species delimitation using Bayesian model comparison and rooted triplets. *Systematic Biology*, 65(5): 759–771.
- Fujisawa T, Barraclough TG (2013) Delimiting species using single-locus data and the generalized mixed Yule coalescent approach: A revised method and evaluation on simulated data sets. *Systematic Biology*, 62(5): 707–724.
- Green RE et al. (2010) A draft sequence of the Neandertal genome. *Science*, 328(5979): 710–722.
- Haller BC, Galloway J, Kelleher J, Messer PW, Ralph PL (2019) Tree-sequence recording in SLiM opens new horizons for forward-time simulation of whole genomes. *Molecular Ecology Resources*, 19(2): 552–566.
- Haller BC, Messer PW (2019) SLiM 3: Forward genetic simulations beyond the Wright-Fisher model. *Molecular Biology and Evolution*, 36(3): 632–637.
- Hancock ZB, Blackmon H (2020) Ghosts of a structured past: Impacts of ancestral patterns of isolation-by-distance on divergence-time estimation. *Journal of Heredity*, in press.
- Hancock ZB, Hardin FO, Light JE (2019) Phylogeography of sand-burrowing amphipods (Haustoriidae) supports an ancient suture zone in the Gulf of Mexico. *Journal of Biogeography*, 46(11): 2532–2547.
- Hengeveld R, Haeck J (1982) The distribution of abundance. I. Measurements. *Journal of Biogeography*, 9: 303–316.
- Herman A, et al. (2018) The role of gene flow in rapid and repeated evolution of cave-related traits in Mexican tetra, *Astyanax mexicanus*. *Molecular Ecology*, 27(22): 4397–4416.
- Huson DH (1998) SplitsTree: analyzing and visualizing evolutionary data. *Bioinformatics*, 14(1): 68–73.

Huynh S, Marcussen T, Felber F, Parisod C (2019) Hybridization preceded radiation in diploid wheats. *Molecular Phylogenetics and Evolution*, 139: 106554.

Jasper M, Schmidt TL, Ahmad NW, Sinkins SP, Hoffmann AA (2019) A genomic approach to inferring kinship reveals limited intergenerational dispersal in the yellow fever mosquito. *Molecular Ecology Resources*, 19(5): 1254–1264.

Kelleher J, Etheridge AM, McVean G (2016) Efficient coalescent simulation and genealogical analysis for large sample sizes. *PLoS Computational Biology*, 12(5): e1004842.

Kelleher J, Thornton KR, Ashander J, Ralph PL (2018) Efficient pedigree recording for fast population genetics. *PLoS Computation Biology*, 14(11): e1006581.

Kimura M (1987) Molecular evolutionary clock and the neutral theory. *Journal of Molecular Evolution*, 26: 24–33.

Kimura M, Weiss GH (1964) The stepping stone model of population structure and the decrease of genetic correlation with distance. *Genetics*, 49(4): 561–576.

Kingman JFC (1982) The coalescent. *Stochastic Processes and their Applications*, 13(3): 235–248.

Knowlton N, Weigt LA (1998) New dates and new rates for divergence across the Isthmus of Panama. *Proceedings of the Royal Society B*, 265: –2263.

Krings M, Stone A, Schmitz RW, Krainitzki H, Stoneking M, Pääbo S (1997) Neandertal DNA sequences and the origin of modern humans. *Cell*, 90(1): 19–30.

Leaché AD, Zhu T, Rannala B, Yang Z (2019) The spectre of too many species. *Systematic Biology*, 68(1): 168–181.

Leaché AD, Rannala B (2011) The accuracy of species tree estimation under simulation: A comparison of methods. *Systematic Biology*, 60(2): 126–137.

Lehman L (2014) Stochastic demography and the neutral substitution rate in class-structured populations. *Genetics*, 197(1): 351–360.

Li J, Huang J, Sukumaran J, Knowles LL (2017) Microevolutionary processes impact macroevolutionary patterns. *BMC Evolutionary Biology*, 18: 123.

Lipscomb D, Platnick N, Wheeler Q (2003) The intellectual content of taxonomy: a comment on DNA taxonomy. *Trends in Ecology & Evolution*, 18(2): 65–66.

Lohse K (2009) Can mtDNA barcodes be used to delimit species? A response to Pons et al. (2006). *Systematic Biology*, 58(4): 439–442.

- Lomilino MV, Channell R (1995) Splendid isolation: Patterns of geographic range collapse in endangered mammals. *Journal of Mammalogy*, 76(2): 335–347.
- Luo A, Ling C, Ho SYW, Zhu C (2018) Comparison of methods for molecular species delimitation across a range of speciation scenarios. *Systematic Biology*, 67(5): 830–846.
- MacGuigan DJ, Neer TJ (2019) Phylogenomic signatures of ancient introgression in a rogue lineage of darters (Teleostei: Percidae). *Systematic Biology*, 68(2): 329–346.
- Maddison WP (1997) Gene trees in species trees. *Systematic Biology*, 46(3): 523–536.
- Maddison WP, Knowles LL (2006) Inferring phylogeny despite incomplete lineage sorting. *Systematic Biology*, 55(1): 21–30.
- Malay MCMD, Paulay G (2010) Peripatric speciation drives diversification and distributional pattern of reef hermit crabs (Decapoda: Diogenidae: *Calcinus*). *Evolution*, 64(3): 634–662.
- Mayr E (1954) Change of genetic environment and evolution. In: Huxley J, Hardy AC, Ford EB (eds), *Evolution as a Process*. Unwin Brothers, London, pp. 157–180.
- Moran PAP (1958) Random processes in genetics. *Mathematical Proceedings of the Cambridge Philosophical Society*, 54(1): 60–71.
- Nordborg M (1998) On the probability of Neanderthal ancestry. *American Journal of Human Genetics*, 63(4): 1237–1240.
- Papadopoulou A, Bergsten J, Fujisawa T, Monaghan MT, Barraclough TG, Volger AP (2008) Speciation and DNA barcodes: testing the effects of dispersal on the formation of discrete sequence clusters. *Philosophical Transactions of the Royal Society B*, 363(1506): 2987–2996.
- Pearse DE, Arndt AD, Valenzuela N, Miller BA, Cantarelli V, Sites JW (2006) Estimating population structure under nonequilibrium conditions in a conservation context: continent-wide population genetics of the giant Amazon river turtle, *Podocnemis expansa* (Chelonia; Podocnemididae). *Molecular Ecology*, 15(4): 985–1006.
- Poissant J, Knight TW, Ferguson MM (2005) Nonequilibrium conditions following landscape rearrangement: the relative contribution of past and current hydrological landscapes on the genetic structure of a stream-dwelling fish. *Molecular Ecology*, 14(5): 1321–1331.
- Pons J, Barraclough TG, Gomez-Zuria J, Cardoso A, Duran DP, Hazell S, Kamoun S, Sumlin WD, Volger AP (2006) Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology*, 55(4): 595–609.
- Portnoy DS, Gold JR (2012) Evidence of multiple vicariance in a marine suture-zone in the Gulf of Mexico. *Journal of Biogeography*, 39(8): 1499–1507.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, 155(2): 945–959.

Ogilvie HA, Bouckaert RR, Drummond AJ (2017) StarBEAST2 brings faster species tree inference and accurate estimates of substitution rates. *Molecular Biology and Evolution*, 34(8): 2101–2114.

Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA (2018) Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*, syy032.

Rannala B, Yang Z (2003) Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. *Genetics*, 164(4): 1645–1656.

Revell LJ (2012) phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, 3: 217–223.

Robinson DF, Foulds LR (1979) Comparison of weighted labelled trees. In: Horadam AF, Wallis WD (eds) *Combinatorial Mathematics VI. Lecture Notes in Mathematics*, Vol. 748. Springer, Berlin, Heidelberg.

Rosenberg NA (2002) The probability of topological concordance of gene trees and species trees. *Theoretical Population Biology*, 61(2): 225–247.

Sauquet H, et al. (2012) Testing the impact of calibration on molecular divergence times using a fossil-rich group: The case of *Nothofagus* (Fagales). *Systematic Biology*, 61(2): 289–313.

Schumer M, Xu C, Powell DL, Durvasula A, Skov L, Holland C, Blazier JC, Sankararaman S, Andolfatto P, Rosenthal GG, Przeworski M (2018) Natural selection interacts with recombination to shape the evolution of hybrid genomes. *Science*, 360(6389): 656–660.

Scotland R, Hughes C, Bailey D, Wortley A (2003) The *big machine* and the much-maligned taxonomist DNA taxonomy and the web. *Systematics and Biodiversity*, 1(2): 139–143.

Slatkin M (1991) Inbreeding coefficients and coalescence times. *Genetics Research*, 58(2): 167–175.

Slatkin M, Pollack JL (2009) Subdivision in an ancestral species creates asymmetry in gene trees. *Molecular Biology and Evolution*, 25(10): 2241–2246.

Springer MS, Gatesy J (2016) The gene tree delusion. *Molecular Phylogenetics and Evolution*, 94: 1–33.

Solís-Lemus C, Ané C (2016) Inferring phylogenetic networks with maximum pseudolikelihood under incomplete lineage sorting. *PLoS Genetics*, 12(3): e1005896.

Solís-Lemus C, Bastide P, Ané C (2017) PhyloNetworks: A package for phylogenetic networks. *Molecular Biology and Evolution*, 34(12): 3292–3298.

Sukumaran J, Knowles LL (2017) Multispecies coalescent delimits structure, not species. *PNAS*, 114(7): 1607–1612.

Takada Y, Sakuma K, Fujii T, Kojima S (2018) Phylogeography of the sandy beach amphipod *Haustoriodes japonicus* along the Sea of Japan: Paleogeographical signatures of cryptic regional divergences. *Estuarine, Coastal, and Shelf Science*, 200(5): 19–30.

Talavera G, Lukhtanov VA, Rieppel L, Pierce NE, Vila R (2013) In the shadow of phylogenetic uncertainty: The recent diversification of *Lysandra* butterflies through chromosomal changes. *Molecular Phylogenetics and Evolution*, 69(3): 469–478.

Theunert C, Slatkin M (2017) Distinguishing recent admixture from ancestral population structure. *Genome Biology and Evolution*, 9(3): 427–437.

Wakeley J (1999) Nonequilibrium migration in human history. *Genetics*, 153(4): 1863–1871.

Wakeley J (2000) The effects of subdivision on the genetic divergence of populations and species. *Evolution*, 54(4): 1092–1101.

Wakeley J (2009) *Coalescent Theory*. Roberts and Company Publishers, Greenwood Village, CO.

Waldrop E, Hobbs JA, Randall JE, DiBattista JD, Rocha LA, Kosaki RK, Berumen ML, Bowen BW (2016) Phylogeography, population structure and evolution of coral-eating butterflyfishes (Family Chaetodontidae, genus *Chaetodon*, subgenus *Corallochaetodon*). *Journal of Biogeography*, 43(6): 116–129.

Wall JD (2000) Detecting ancient admixture in humans using sequence polymorphism data. *Genetics*, 154(3): 1271–1279.

Weir JT, Schlüter D (2008) Calibrating the avian molecular clock. *Molecular Ecology*, 17(10): 2321–2328.

Wilkins JF (2004) A separation-of-timescales approach to the coalescent in a continuous population. *Genetics*, 168(4): 2227–2244.

Wilkins JF, Wakeley J (2002) The coalescent in a continuous, finite, linear population. *Genetics*, 161(2): 873–888.

Wright S (1931) Evolution in Mendelian populations. *Genetics*, 16(2): 97–159.

Wright S (1943) Isolation by distance. *Genetics*, 28(2): 114–138.

Yang MA, Malaspinas A, Durand EY, Slatkin M (2012) Ancient structure in Africa unlikely to explain Neanderthal and non-African genetic similarity. *Molecular Biology and Evolution*, 29(10): 2987–2995.

Yoder J (2020) For population genetics, continuous space might be the final frontier. *The Molecular Ecologist*, <https://www.molecularecologist.com/2020/03/for-population-genetics-continuous-space-might-be-the-final-frontier/>.

Yu Y, Harris AJ, Blair C, He X (2015) RASP (Reconstruct Ancestral State in Phylogenies): A tool for historical biogeography. *Molecular Phylogenetics and Evolution*, 87: 46–49.

Zhang J, Kapli P, Pavlidis P, Stamatakis A (2013) A general species delimitation method with applications to phylogenetic placements. *Bioinformatics*, 29(22): 2869–2876.

## APPENDIX A

### CHAPTER I

Maps throughout this article were created using ArcGIS® software by Esri. ArcGIS® and ArcMap™ are the intellectual property of Esri and are used herein under license. Copyright © Esri. All rights reserved. For more information about Esri® software, please visit [www.esri.com](http://www.esri.com).

Data has been stored in the Dryad repository; doi:10.5061/dryad.pd0ks46.

### CHAPTER II

Code from this chapter is available at <https://github.com/hancockzb/GenomeSize>.

### CHAPTER III

Code from this chapter is available at <https://github.com/hancockzb/ancestralIBD>.

### CHAPTER IV

#### Methods

##### *Model details and discussion*

The continuous-space model is an extension of Rogan et al. (*in prep*), which itself was a modification of Bradburd & Ralph (2019) and Battey et al. (2020). Unlike in these previous applications, the model is constrained to non-overlapping generations by enforcing a pseudo-age structure, which sets fitness to 0 at age 1. Beyond this, the model is unlike the classical WF model in that population size is not constant but is an emergent property of the population carrying-capacity,  $K$ . Furthermore, while the mean and variance of the number of offspring should converge to a Poisson process with  $\lambda = 1$  as  $N \rightarrow \infty$  (Wakeley 2009), in my model the population is only sustained when  $\lambda = 2$ . Therefore, the mean and variance in offspring is slightly higher than under the WF model.

The landscape is defined by a matrix,  $M$ , of height  $h$  and length  $ls$ , where

$$\begin{cases} \text{length}(M) = ls, & \text{when } s = 1 \\ \text{length}(M) = ls + 4s, & \text{when } s = 1 \end{cases}$$

The inclusion of  $4s$  provides a distance of 4 cells between each species,  $s$ , following the split while ensuring that the species-specific  $M = h \times l$ . This distance exceeds the maximum dispersal distance in the unclustered model. The number of species simulated were 3, 4, and 6, each with different applications (see the different subsections below). I used interpolation to allow the fitness values of cells to be a continuous gradient instead of discrete steps between cells. The cell-specific fitnesses,  $f_c$ , were equal to 1.0 (i.e., there were no cell-specific fitness declines).

Fitness was also determined by the local population density with a maximum strength of  $1 / 2\pi\sigma_I^2$  and max distance of  $3\sigma_I$ , which was drawn from a Gaussian distribution. To reduce clumping along the edges, I enforce fitness decline relative to distance from the edge as

$$f_e = \min \left( 1, \sqrt{\frac{x_i}{\sigma_D}} \right) \min \left( 1, \sqrt{\frac{y_i}{\sigma_D}} \right) \min \left( 1, \sqrt{\frac{(M-x_i)}{\sigma_D}} \right) \min \left( 1, \sqrt{\frac{(M-y_i)}{\sigma_D}} \right) F,$$

where  $F$  is the cumulative fitness accounting for both  $f_c$  and  $\sigma_I$ , and  $x_i$  and  $y_i$  are the spatial positions of individual  $i$  in the  $x$  and  $y$  coordinates, respectively (Battey et al. 2020). This causes fitness declines towards the edges to be more extreme when  $\sigma_D < 1.0$ , which promotes heterogeneity in population density across the range.

Individuals are hermaphroditic but incapable of selfing. Mate-choice strength is drawn from a Gaussian distribution with a maximum of  $1 / 2\pi\sigma_M^2$  within a max distance of  $3\sigma_M$ . Immediately after offspring production, individuals disperse across the range by taking two draws from a random uniform distribution with a minimum of  $-3\sigma_D$  and maximum of  $3\sigma_D$ . The two life-history models we refer to as “clustered” and “unclustered”, which characterize the behavior of the simulated individuals. The clustered model sets  $\sigma_D$  and  $\sigma_M$  each = 0.25, and  $\sigma_I$  = 1.0; the unclustered model sets  $\sigma_D$  and  $\sigma_M$  = 1.0, and  $\sigma_I$  = 0.25.

For the vicariant model of speciation, the initial range of size  $M$  is populated with  $M*K$  individuals, where  $K$  is the local carrying-capacity, with initial positions drawn from a random uniform distribution. For the 3 and 4 species models, a period of 10,000 generations of burn-in proceeds the initial split. For the 6 species models, this period is extended to 50,000 generations. The range then proceeds through a series of fragmentations that carve the range into  $s$  equal final ranges of size 20×20. For the peripatric speciation model, the fitness of the initial range is set to 0.0 except for the far-left corner of size 20×20. The population then proceeds through an expansion period of 100 generations where the habitable range size has extended to a length of  $ls + 4s$ . After 100 generations, the range is split by setting the fitness values of 4 columns of the landscape matrix to 0.0. This occurs  $s$  times.

Simulated individuals are diploid with genome sizes of 1000 Mb and a recombination rate of  $10^{-9}$ . No mutations are simulated during the runs as these can be overlaid after to increase computational efficiency. Tree-sequence recording is enabled to track the true local ancestry of all individuals; however, throughout the run internal simplification occurs that discards nodes that do not contribute offspring to the final generation.

Several important theoretical points about the model should be noted. First, both the clustered and unclustered models have a period of spatial autocorrelation and therefore neither represent true random mating. The unclustered model is merely meant to *approximate* random mating. Along these lines, the neighborhood sizes of each model are < 1000 and therefore will not generate random mating expectations (Wright 1943). Furthermore, in the clustered model the neighborhood size may be low enough that it *never* converges to the  $n$ -coalescent (Kingman 1982; Wilkins 2004; Wakeley 2009). Another way to state this is that there is no transition between the scattering and collecting phases – the population is locked in the scattering phase. However, the process of recapitation in *msprime* (see *Trees in Space: Models and Datasets*) ensures that the collecting phase occurs as it simulates coalescence among all multiple roots under the  $n$ -coalescent irrespective of geography, life-history, etc. I justify this forced transition by pointing out that any uncoalesced nodes present at the beginning of the simulation either 1) belong to the same species, usually the last  $k = 2$  samples to coalesce or 2) are from different

species, and therefore represent a case of deep coalescence. The former case will add high frequency SNPs on the last long branches leading to the most recent common ancestor. These high frequency SNPs lead to deviations from the expected unfolded site frequency spectrum,  $E[\xi_i]$ , which under neutrality should be  $\frac{\theta}{i}$ ; we see an inflated proportion of high frequency SNPs in the site frequency spectrum (**Fig. 4.6**). Importantly, this occurs in both the clustered and unclustered models. However, as seen in **Fig. 4.6**, the clustered model leads to greater deviations in  $E[\xi_i]$  at intermediate frequencies. The impact of (2) on topological inference will be conservative as it will reduce the  $T_{\text{MRCA}}$  by forcing random mating.

#### *Identifiability of 4-taxon trees*

In the section *Slatkin's Skew and Gene Tree Asymmetry*, I provide the results of both 3- and 4-taxon simulations. There's an important point that should be noted regarding  $t$  in each of these models. Since  $t$  is in units of  $2N_e$ , to accurately reflect theoretical expectations of divergence times we need to know  $N_e$ . For the 3-taxon tree,  $t$  can be calculated directly. However, for the 4-taxon tree there are an infinite number of combinations of  $t_1$  and  $t_2$  that produce a given proportion of gene trees. Therefore, the actual  $t_1$  and  $t_2$  values – those that reflect the “true”  $N_e$  – cannot be identified. The lack of identifiability is twofold: 1) the infinite combinations issue already pointed out; and 2) that reducing  $T_D$  also reduces the amount of time a population has to equilibrate following the split. Therefore, estimating  $N_e$  from  $\pi_{12}$  would be incorrect because the relationship  $\pi_{12} = 4N_e\mu$  only holds if the population is in equilibrium. This is further exacerbated in the clustered model as it takes much longer to return to equilibrium following a split than the clustered model does. Due to these issues, for the 4-taxon tree we examine gene tree proportions in units of  $2N$ , where  $N$  is the census population size. This means that the coalescent units in **Fig. 4.2c** between “sigma1” and “sigma25” are not the same. However, since we focus only on the topologies that are expected to be equal at all values of  $T_D$ , I do not think this disparity impacts the results presented.

## APPENDIX B

### CHAPTER I

**Table S1.1.** Sampling information, including morphological identification, OTU identified using tree-based species delimitation methods, number of individuals sequenced for *COI*, *16S*, *28S*, and *18S*, and the sample site location. Localities match those presented in Fig. 1. The numbers following LT-Pass Christian refer to the two unique genetic lineages of *Lepidactylus triarticulatus* found at this locality.

Identification	# Individuals sequenced for <i>COI</i>	<i>16S</i>	<i>28S</i>	<i>18S</i>	Location (Fig. 1 locality #)	Voucher #
<i>Haustorius galvezi</i>	9	-	1	1	Galveston, Texas (7)	-
<i>Haustorius galvezi</i>	10	-	1	-	Jamaica Beach, Texas (6)	USNM 1492390 (paratype)
<i>Haustorius galvezi</i>	1	-	-	-	Matagorda, Texas (5)	USNM 1492388 (paratype)
<i>Haustorius galvezi</i>	4	-	-	1	South Padre, TX (4)	-
<i>Haustorius galvezi</i>	10	3	2	-	Tamiahua, MX-VE (3)	-
<i>Haustorius galvezi</i>	7	6	-	-	Antón Lizardo, MX-VE (2)	BRTC 24863
<i>Haustorius galvezi</i>	3	7	-	-	Tabasco-Campeche border, MX (1)	-
<i>Haustorius jayneae</i>	4	5	2	-	Dauphin Island, AL (12)	-
<i>Haustorius jayneae</i>	5	1	1	1	Grayton Beach, FL (14)	-
<i>Haustorius jayneae</i>	5	4	-	-	St. Andrews State Park, FL (15)	-

**Table S1.1.** Continued.

Identification	# Individuals sequenced for <i>COI</i>	16S	28S	18S	Location (Fig. 1 locality #)	Voucher #
<i>Haustorius jayneae</i>	5	4	-	-	Carrabelle Beach, FL (16)	BRTC 24858
<i>Haustorius allardi</i>	3	8	1	1	Holly Beach, LA (9)	USNM 1492392 (paratype)
<i>Haustorius canadensis</i>	2	-	2	2	Cape May, NJ	BRTC 24864
<i>Haustorius arenarius</i>	2	-	1	2	Belgian Institute	RBINS INV.138073
<i>Lepidactylus triarticulatus</i>	5	1	2	1	Grand Isle, LA (10)	BRTC 24861
<i>Lepidactylus triarticulatus</i>	5	4	1	1	Sea Rim State Park, TX (8)	BRTC 24862
<i>Lepidactylus triarticulatus</i>	4	3	1	1	Pass Christian, MS (11)	BRTC 24860
<i>Lepidactylus triarticulatus</i>	6	1	1	1	Pass Christian, MS (11)	-
<i>Lepidactylus triarticulatus</i>	5	5	1	1	Pensacola Bay, FL (13)	BRTC 24859
<b>Outgroups</b>						
<i>Gammarus balcanicus</i>	-	-	1	-	Accession #: JF966175	-
<i>Neogammarus nudus</i>	-	-	1	-	Accession #: KF478630	-
<i>Accubogamm arus sp.</i>	-	-	-	1	Accession #: KF478647	-
<i>Echinogamm arus anisocheirus</i>	-	-	-	1	Accession #: KF478629	-
<i>Gammarellus angulosus</i>	-	1	-	-	Accession #: AY926717	-

**Table S1.1.** Continued.

Identification	# Individuals sequenced for <i>COI</i>	<i>16S</i>	<i>28S</i>	<i>18S</i>	Location (Fig. 1 locality #)	Voucher #
<i>Parhyale plumicornis</i>	-	1	-	-	Accession #: KU565878	-
<i>Niphargus timavi</i>	1	-	-	-	Accession #: KR858492	-
<i>Niphargus krameri</i>	1	-	-	-	Accession #: KR858481	-

**Table S1.2.** Locus, primer sets, PCR conditions, and references for molecular markers used in this study.

Locus	Primer Set	PCR conditions	References
<i>COI</i>	L6625 — 5'-CCGGATCCTYTGRT YTYGGNCAYCC-3' H7005 — 5'-CCGGATCCACANCRT ARTANGTRTCRTG-3'	5 mins at 94°C, 45 cycles of 94°C for 1 min, 1 min at 51°C, 1 min 30 sec at 72°C, final annealing at 72°C for 10 mins	Hafner et al., 1994
<i>16S</i>	16STf — 5'GGTAWHYTRACYG TGCTAAG-3' 16Sbr — 5'-CCGGTTGAACTCAG ATCATGT-3'	4 mins at 95°C, 40 cycles of 95°C for 1 min, 1 min at 45°C, 2 min 30 sec at 72°C, final annealing at 72°C for 7 mins	Macdonald et al., 2005
<i>18S</i>	18SGF — 5'-GGATAACTGTGGTAA TTCCAGAGCT-3' 18SGR — 5'TAGTAGCGACGGGC GGTGTGTA-3'	1 min at 95°C, 35 cycles of 20 sec at 95°C, 20 sec at 63°C, 45 sec at 72°C, final annealing at 72°C for 7 mins	Hou et al., 2007
<i>28S</i>	28F — 5'-TTAGTAGGGGCGACC GAACAGGGAT-3' 28R — 5'-GTCTTCGCCCTATG CCCAACGA-3'	1 min at 95°C, 35 cycles of 20 sec at 95°C, 20 sec at 63°C, 45 sec at 72°C, final annealing at 72°C for 7 mins	Hou et al., 2007

**Table S1.3.** DNA substitution models supported by PartitionFinder2 and jModeltest2.

Gene	Model	Replaced as
<i>COI</i>	1st position — GTR + G	
	2nd position — TRNEF + G	GTR + G
	3rd position — F81 + I	
<i>18S</i>	TIM3 + G	GTR + G
<i>28S</i>	GTR + G	
<i>16S</i>	HKY + G	

**Table S1.4.** Pairwise Distance Matrix among OTUs. Below the diagonal are distances calculated using a K2P substitution model. Numbers in blue above the diagonal are standard errors from 500 bootstrap replicates in MEGA7. OTUs were identified using tree-based species delimitation methods and definitions of the abbreviations are presented in Table S1; LT = *Lepidactylus triarticulatus*.

	<i>H. galvezi</i> (MX)	<i>H. galvezi</i> (TX)	<i>H. jayneae</i>	<i>H. allardi</i>	LT Grand Isle	LT Sea Rim	LT Pass Christian[1]	LT Pensacola Bay	LT Pass Christian[2]	<i>H. canadensis</i>
<i>H. galvezi</i> (MX)		0.0087	0.0186	0.0245	0.0261	0.0283	0.0305	0.0274	0.0268	0.0275
<i>H. galvezi</i> (TX)	0.0385		0.0196	0.0243	0.0261	0.0273	0.0317	0.0270	0.0255	0.0271
<i>H. jayneae</i>	0.1152	0.1200		0.0274	0.0302	0.0311	0.0321	0.0301	0.0298	0.0283
<i>H. allardi</i>	0.1843	0.1823	0.2006		0.0285	0.0291	0.0325	0.0311	0.0304	0.0317
LT Grand Isle	0.1989	0.1897	0.2377	0.2369		0.0096	0.0264	0.0274	0.0266	0.0286
LT Sea Rim	0.2232	0.2040	0.2517	0.2406	0.0351		0.0244	0.0275	0.0256	0.0290

**Table S1.4.** Continued.

	<i>H. galvezi</i> (MX)	<i>H. galvezi</i> (TX)	<i>H. jayneae</i>	<i>H. allardi</i>	LT Grand Isle	LT Sea Rim	LT Pass Christi an[1]	LT Pensac ola Bay	LT Pass Christi an[2]	<i>H. canade nsis</i>
LT Pass Christi an[1]	0.2412	0.2478	0.2485	0.2563	0.1726	0.1547		<b>0.0292</b>	<b>0.0269</b>	<b>0.0313</b>
LT Pensac ola Bay	0.2353	0.2256	0.2479	0.2535	0.1913	0.1861	0.2111		<b>0.0085</b>	<b>0.0301</b>
LT Pass Christi an[2]	0.2250	0.2098	0.2355	0.2352	0.1806	0.1672	0.1858	0.0305		<b>0.0297</b>
<i>H. canade nsis</i>	0.2057	0.2048	0.2146	0.2615	0.2044	0.2087	0.2378	0.2405	0.2307	

**Table S1.5.** Pairwise  $F_{ST}$  matrix among localities shown in **Fig. 6**. Significant values are shown in bold; \* =  $P < 0.01$ , \*\* =  $P < 0.001$ , \*\*\* =  $P < 0.0001$ .

Locality	Anton Lizardo	South Tabasco	South Padre	Tamiah ua	Galvest on	Jamaica Beach	Dauphin Island	Grayton Beach	St Andrew s	Carrabel le Beach
Anton Lizardo	-									
South Tabasco	0.0215	-								
South Padre	<b>0.2409*</b>	<b>0.3645*</b>	-							
Tamiah ua	<b>0.3567**</b>	<b>0.3443**</b>	<b>0.2530**</b>	-						
Galvest on	<b>0.7080**</b>	<b>0.7168*</b>	<b>0.7103**</b>	<b>0.6540**</b>	-					
Jamaica Beach	<b>0.6987**</b>	<b>0.7038**</b>	<b>0.6896**</b>	<b>0.6425**</b>	0.0000	-				

**Table S1.5.** Continued.

Locality	Anton Lizardo	South Tabasco	South Padre	Tamiah ua	Galvest on	Jamaica Beach	Dauphin Island	Grayton Beach	St Andrew s	Carrabel le Beach
Dauphin Island	<b>0.9227*</b> **	<b>0.9609*</b>	<b>0.9623*</b>	<b>0.8809*</b> **	<b>0.9138*</b> *	<b>0.9069*</b> **	-			
Grayton Beach	<b>0.9251*</b> **	<b>0.9575*</b>	<b>0.9601*</b> **	<b>0.8852*</b> **	<b>0.9150*</b> **	<b>0.9095*</b> **	<b>0.6821*</b> *	-		
St Andrew s	<b>0.9189*</b> *	<b>0.9481*</b> **	<b>0.9506*</b> **	<b>0.8803*</b> **	<b>0.9109*</b> **	<b>0.9095*</b> **	0.0000	<b>0.4298*</b>	-	
Carrabel le Beach	<b>0.9225*</b> **	<b>0.9544*</b> **	<b>0.9565*</b> **	<b>0.8828*</b> *	<b>0.9137*</b> **	<b>0.9079*</b> **	0.3919	0.1129	0.0790	-

**Figure S1.1.** Topology comparison for *COI* inferred from each method (RAxML, MrBayes, BEAST2).

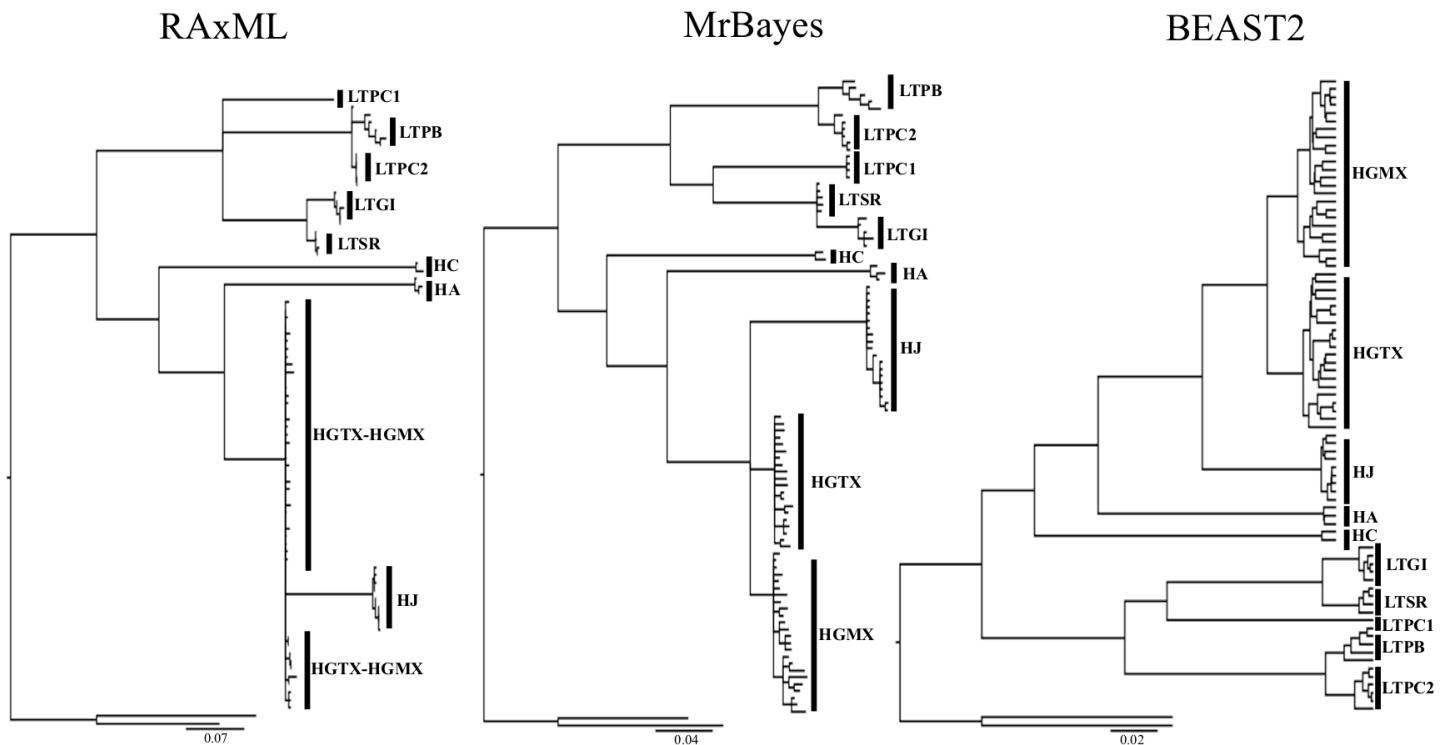
**Figure S1.2.** 28S gene tree produced in MrBayes. Node support: (posterior probability, bootstrap support from RAxML analyses).

**Figure S1.3.** 16S gene tree produced in MrBayes. Node support: (posterior probability, bootstrap support from RAxML analyses).

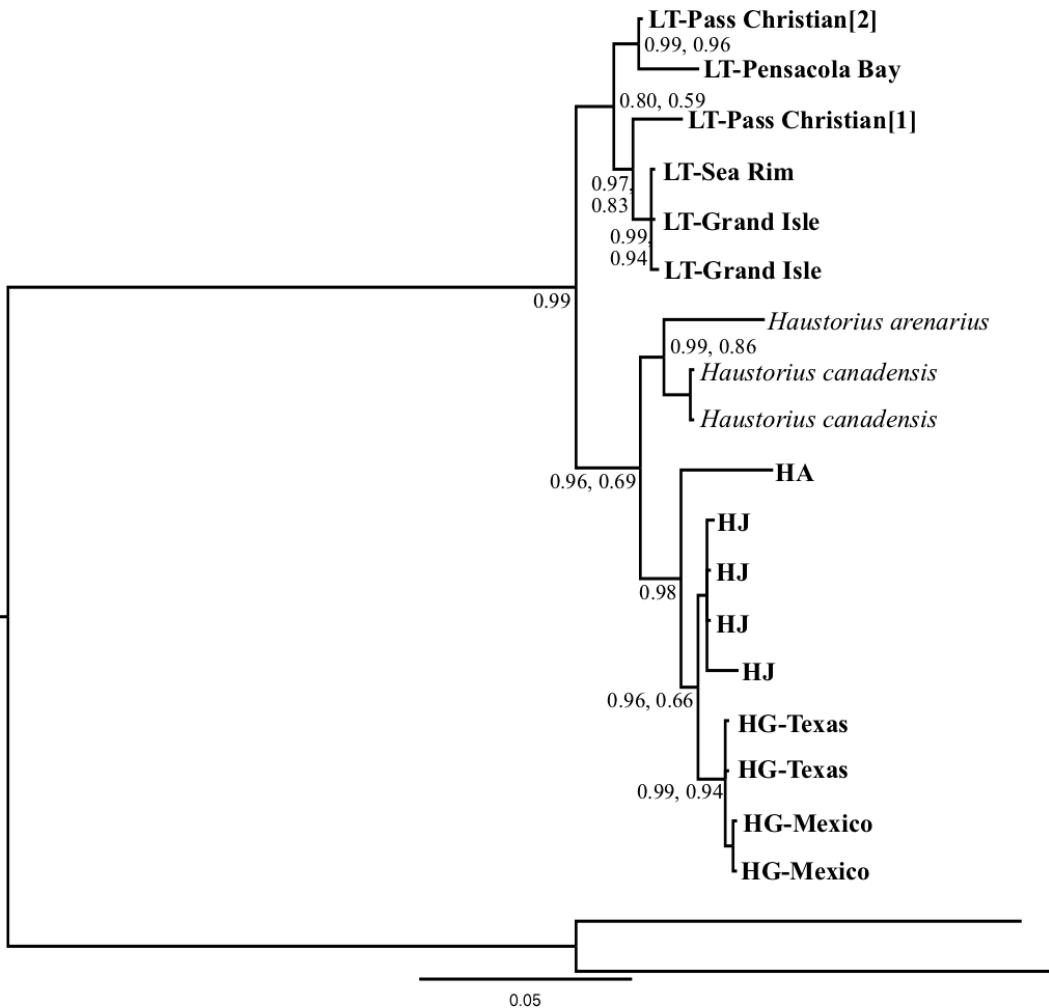
**Figure S1.4.** 18S gene tree produced in MrBayes. Node support: (posterior probability, bootstrap support from RAxML analyses).

**Figure S1.5.** TCS haplotype network for *COI* produced in PopART. Tick-marks are inferred mutations.

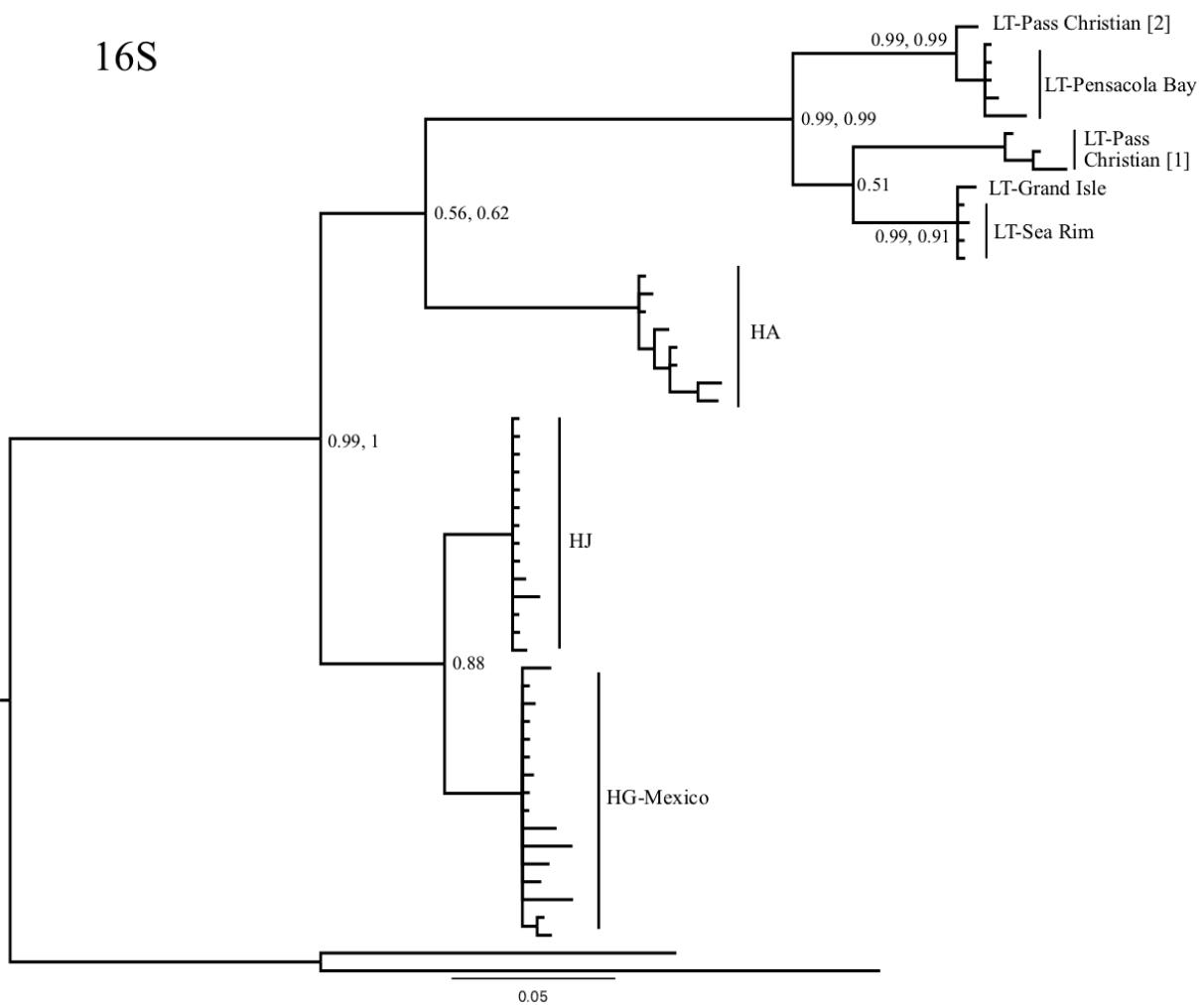
**Figure S1.1.**



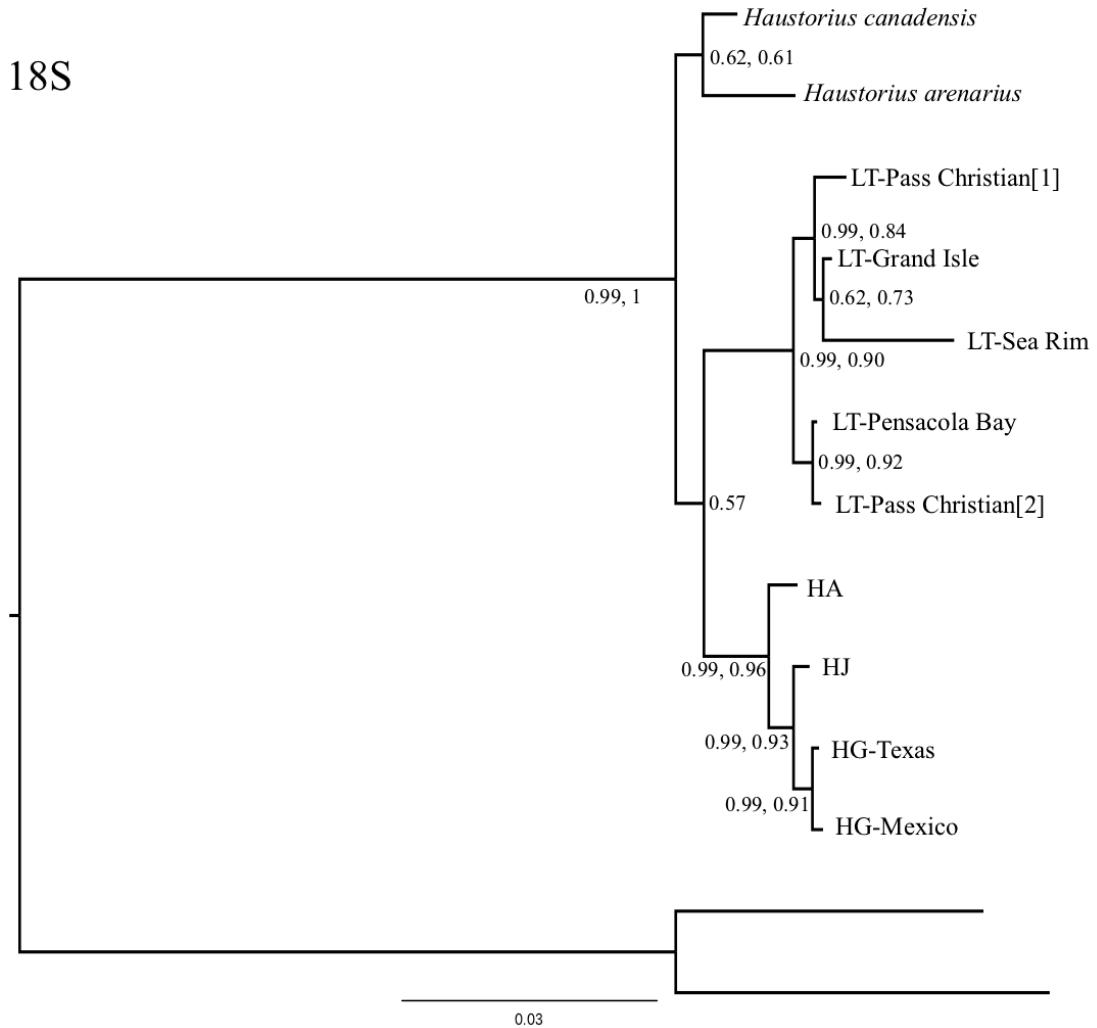
**Figure S1.2**



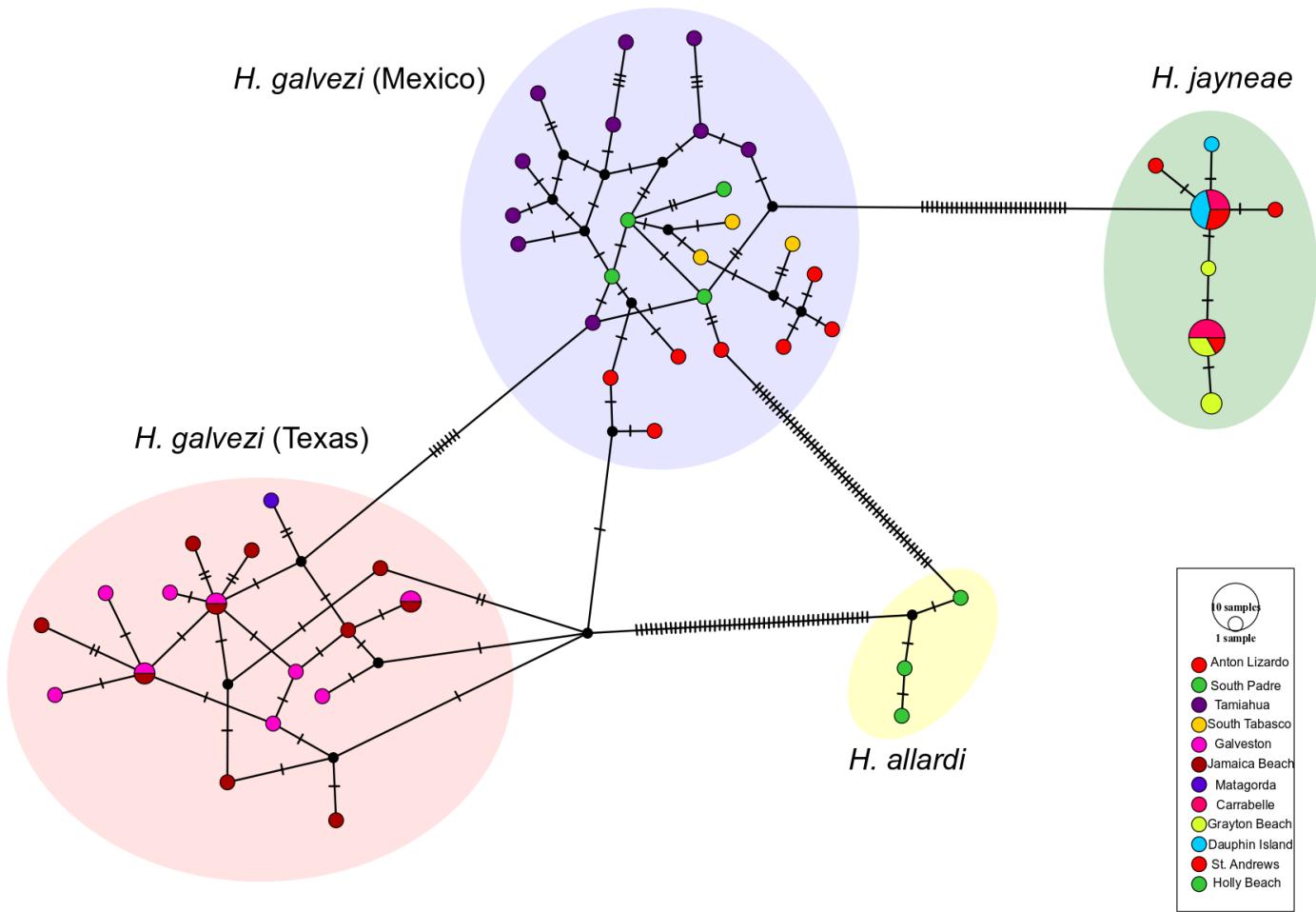
**Figure S1.3**



**Figure S1.4**



**Figure S1.5**



## CHAPTER II

### Tables

**Table S2.1.** Pairwise genetic differentiation measures.

Comparison	$H_{ST}$	$K_{ST}$	$K_{ST}^*$
<i>H. jayneae</i> (Dauphin Island) x <i>H. jayneae</i> (Grayton Beach)	<b>0.21032*</b>	<b>0.54839*</b>	<b>0.42426*</b>
<i>H. jayneae</i> (Dauphin Island) x <i>H. jayneae</i> (Carrabelle Beach)	0.12348	0.27273	0.20226
<i>H. jayneae</i> (Grayton Beach) x <i>H. jayneae</i> (Carrabelle Beach)	0.04545	0.06604	0.2850
<i>H. canadensis</i> (Cape Cod) x <i>H. canadensis</i> (Long Island)	<b>0.20687**</b>	0.15385	<b>0.15438*</b>
<i>H. canadensis</i> (Cape Cod) x <i>H. canadensis</i> (Melbourne)	<b>0.1340*</b>	<b>0.72846***</b>	<b>0.44586***</b>
<i>H. canadensis</i> (Cape Cod) x <i>H. canadensis</i> (Jacksonville)	<b>0.59241**</b>	<b>0.83879**</b>	<b>0.70653**</b>
<i>H. canadensis</i> (Cape Cod) x <i>H. canadensis</i> (Tybee)	<b>0.15020**</b>	<b>0.74960***</b>	<b>0.44608**</b>
<i>H. canadensis</i> (Long Island) x <i>H. canadensis</i> (Jacksonville)	<b>0.64770**</b>	<b>0.87097**</b>	<b>0.80899**</b>
<i>H. canadensis</i> (Long Island) x <i>H. canadensis</i> (Melbourne)	0.04698	<b>0.76483**</b>	<b>0.46874**</b>
<i>H. canadensis</i> (Long Island) x <i>H. canadensis</i> (Tybee)	0.04698	<b>0.74251**</b>	<b>0.41554**</b>
<i>H. canadensis</i> (Jacksonville) x <i>H. canadensis</i> (Melbourne)	<b>0.462628***</b>	<b>0.4662***</b>	<b>0.482366***</b>
<i>H. canadensis</i> (Melbourne) x <i>H. canadensis</i> (Tybee)	0.01296	<b>0.4400**</b>	<b>0.34241***</b>

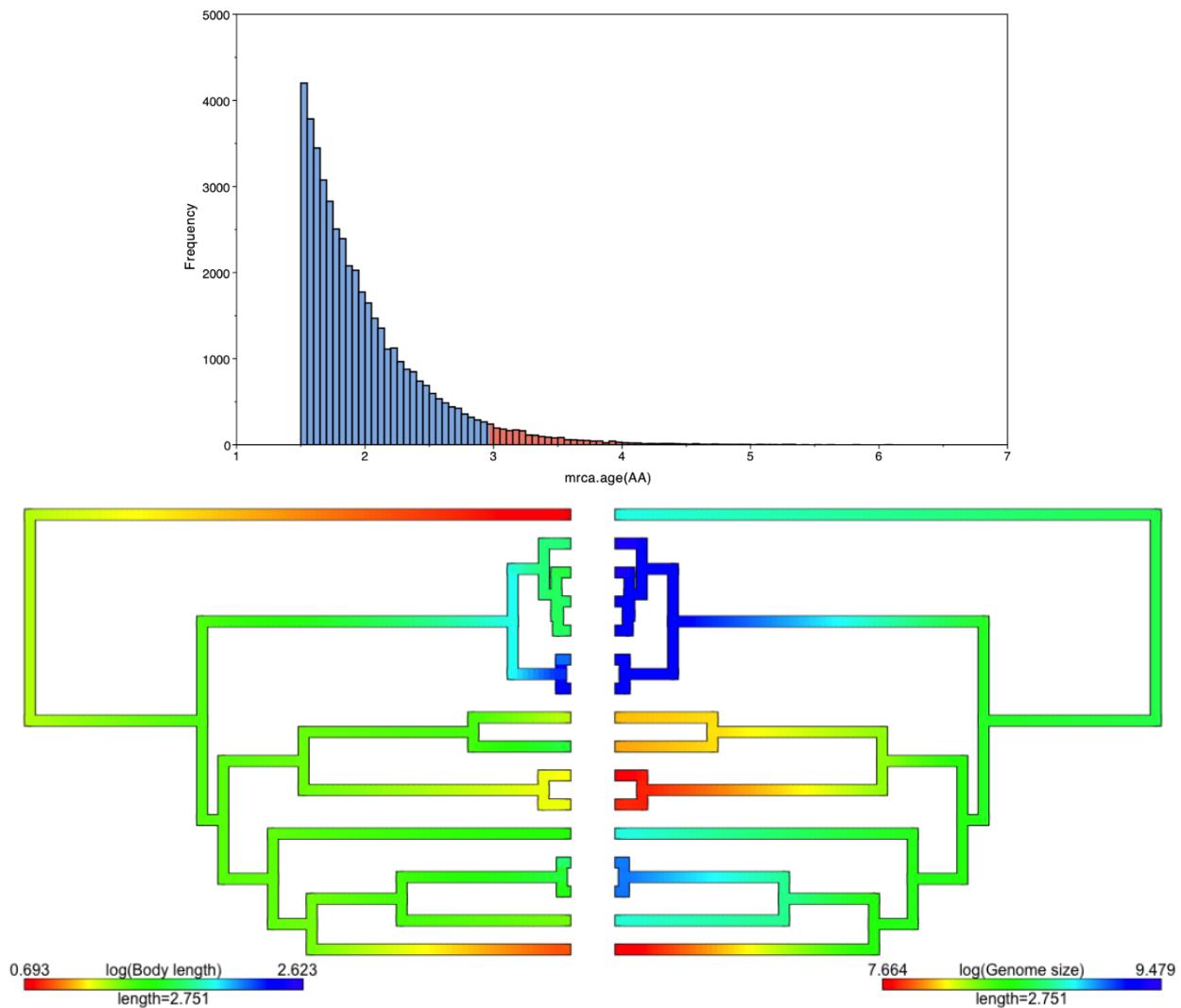
**Table S2.2.** Proportion of major clusters.

Species	Genome size (Mb)	Max reads	Coverage (%)	Total Repeats (%)	LIN E	DIR S	LT R	Penelope	Maverick	Satellite	Unknown	Bottom
<i>Nschmitzi</i>	7500	1655500	0.0221	71.08	7.47	7.09	18.86	0.1	0.6	2.78	8.78	9.83
<i>HcanadensisTy</i>	13080	2145387	0.0164	69.49	18.75	0	3.74	0.4	0.28	4.63	20.81	17.81
<i>HjayneaeG</i>	9330	2464505	0.0264	66.02	17.66	0.22	3.31	0.51	0.77	6.81	15.56	20.47
<i>HjayneaeDI</i>	9520	1420106	0.0149	70.97	29.31	0.01	2.09	0.09	0.61	5.79	14.49	17.38
<i>Hallardi</i>	2130	1128376	0.053	57	8.31	0.01	0.98	0.32	0.31	11.31	17.96	16.57
<i>HcanadensisCC</i>	11390	1845132	0.0162	71.78	21.15	1.52	4.44	0.21	0.06	6.39	19.3	15.58
<i>HcanadensisNC</i>	12830	2469903	0.0193	69.77	18.08	0.54	5.11	0.47	0.31	3.6	21.02	18.47
<i>HcanadensisLo</i>	11990	2585215	0.0216	69.46	16.45	0.57	5.51	0.39	0.33	4.93	20.27	18.92
<i>LdysticusNCBa</i>	7520	1857818	0.0247	72.46	18.22	3.14	17.65	0.49	1.47	2.02	14.84	14.28
<i>HcanadensisJa</i>	11690	2889172	0.0247	68.13	19.31	0	3.94	0.36	0.86	6.69	14.92	21.72
<i>HcanadensisMe</i>	11050	2354941	0.0213	69.28	20.55	0.38	4.42	0.43	1.13	1.47	18.51	22.41
<i>HgalveziMX</i>	7350	2501302	0.034	68.78	19.58	3.93	6.33	1.6	0.95	5.6	15.53	14.84
<i>LtriGI</i>	2380	644675	0.0271	47.67	6.88	0	0.94	0.1	0.13	15.28	10.98	12.89
<i>LtriS</i>	2200	785415	0.0357	48.46	8.43	0.04	1.44	0.07	0.16	12.56	11.92	13.84
<i>LtriPC1</i>	3000	316053	0.0105	54.17	4.01	0	4.1	0.04	0.57	20.8	10.85	13.58
<i>LtriPBAY</i>	2900	268152	0.0092	55.87	2.86	0	1.06	0.01	0.25	28.51	10.82	11.77

## Figures

**Figure S2.1.** Posterior distribution of the estimated age of the *Acanthohaustorius* clade calibration.

**Figure S2.2.** Ancestral reconstructions of log(body length) and log(genome size).



### CHAPTER III

**Table S3.1.** Pairwise Wilcoxon test results of  $T_{\text{MRCA}}$  for comparisons of  $T_D / ND$  of 10, 5, 2, 1, 0.5, and 0.1 from *fastsimcoal2* simulations. Significant ( $p < 0.05$ ) results are bolded.

Comparison	Ratio	Migration rate	p-value
Ends-Center	10	0.1	<b>&lt;2E-16</b>
Island-Center	10	0.1	0.26
Island-Ends	10	0.1	<b>&lt;2E-16</b>
Ends-Center	10	0.01	<b>&lt;2E-16</b>
Island-Center	10	0.01	<b>0.019</b>
Island-Ends	10	0.01	<b>&lt;2E-16</b>
Ends-Center	10	0.001	<b>&lt;2E-16</b>
Island-Center	10	0.001	<b>0.00000012</b>
Island-Ends	10	0.001	<b>&lt;2E-16</b>
Ends-Center	5	0.1	<b>&lt;2E-16</b>
Island-Center	5	0.1	0.22
Island-Ends	5	0.1	<b>&lt;2E-16</b>
Ends-Center	5	0.01	<b>&lt;2E-16</b>
Island-Center	5	0.01	0.18
Island-Ends	5	0.01	<b>&lt;2E-16</b>

**Table S3.1.** Continued.

Comparison	Ratio	Migration rate	p-value
Ends-Center	5	0.001	<b>&lt;2E-16</b>
Island-Center	5	0.001	<b>1.2E-08</b>
Island-Ends	5	0.001	<b>&lt;2E-16</b>
Ends-Center	2	0.1	<b>2.1E-11</b>
Island-Center	2	0.1	0.46
Island-Ends	2	0.1	<b>1.2E-10</b>
Ends-Center	2	0.01	<b>&lt;2E-16</b>
Island-Center	2	0.01	<b>0.029</b>
Island-Ends	2	0.01	<b>&lt;2E-16</b>
Ends-Center	2	0.001	<b>&lt;2E-16</b>
Island-Center	2	0.001	<b>0.000041</b>
Island-Ends	2	0.001	<b>&lt;2E-16</b>
Ends-Center	1	0.1	<b>0.0000038</b>
Island-Center	1	0.1	0.33893
Island-Ends	1	0.1	<b>0.00017</b>
Ends-Center	1	0.01	<b>&lt;2E-16</b>

**Table S3.1.** Continued.

Comparison	Ratio	Migration rate	p-value
Island-Center	1	0.01	0.53
Island-Ends	1	0.01	<b>&lt;2E-16</b>
Ends-Center	1	0.001	<b>&lt;2E-16</b>
Island-Center	1	0.001	<b>0.0000023</b>
Island-Ends	1	0.001	<b>&lt;2E-16</b>
Ends-Center	0.5	0.1	<b>0.027</b>
Island-Center	0.5	0.1	0.759
Island-Ends	0.5	0.1	<b>0.013</b>
Ends-Center	0.5	0.01	<b>&lt;2E-16</b>
Island-Center	0.5	0.01	<b>0.032</b>
Island-Ends	0.5	0.01	<b>&lt;2E-16</b>
Ends-Center	0.5	0.001	<b>&lt;2E-16</b>
Island-Center	0.5	0.001	0.25
Island-Ends	0.5	0.001	<b>&lt;2E-16</b>
Ends-Center	0.1	0.1	1
Island-Center	0.1	0.1	1

**Table S3.1.** Continued.

Comparison	Ratio	Migration rate	p-value
Island-Ends	0.1	0.1	1
Ends-Center	0.1	0.01	<b>0.0000068</b>
Island-Center	0.1	0.01	0.91
Island-Ends	0.1	0.01	<b>0.0000068</b>
Ends-Center	0.1	0.001	<b>&lt;2E-16</b>
Island-Center	0.1	0.001	0.077
Island-Ends	0.1	0.001	<b>&lt;2E-16</b>

**Table S3.2.** Kruskal-Wallis test comparing  $(T_{MRCA} - T_D) / T_D$  estimated from SNAPP of center and end species for different rates of migration and ratios of  $T_D / ND$ . Bolded p-values indicate  $p < 0.05$ .

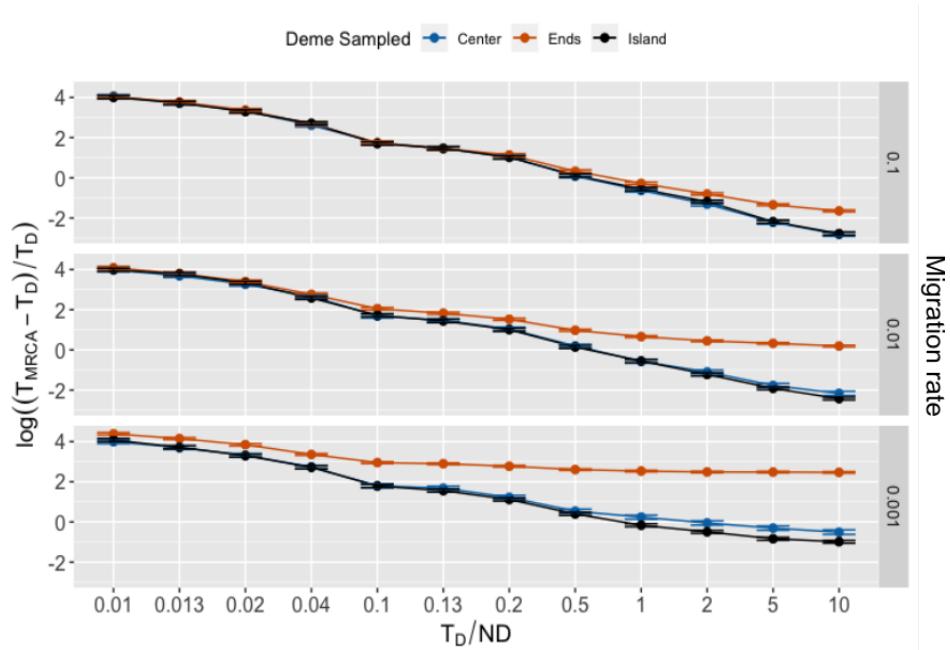
Ratio	Migration_rate	p-value
50	0.1	0.8724
50	0.01	0.497
50	0.001	<b>3.821e-09</b>
25	0.1	<b>0.001449</b>
25	0.01	0.5983
25	0.001	<b>&lt; 2.2e-16</b>
10	0.1	0.2231
10	0.01	<b>0.03289</b>
10	0.001	<b>&lt; 2.2e-16</b>
5	0.1	<b>7.891e-05</b>
5	0.01	<b>8.788e-06</b>
5	0.001	<b>&lt; 2.2e-16</b>
1	0.1	0.4134
1	0.01	0.07628
1	0.001	<b>&lt; 2.2e-16</b>

**Table S3.3.** Table of estimated divergence-times in SNAPP.

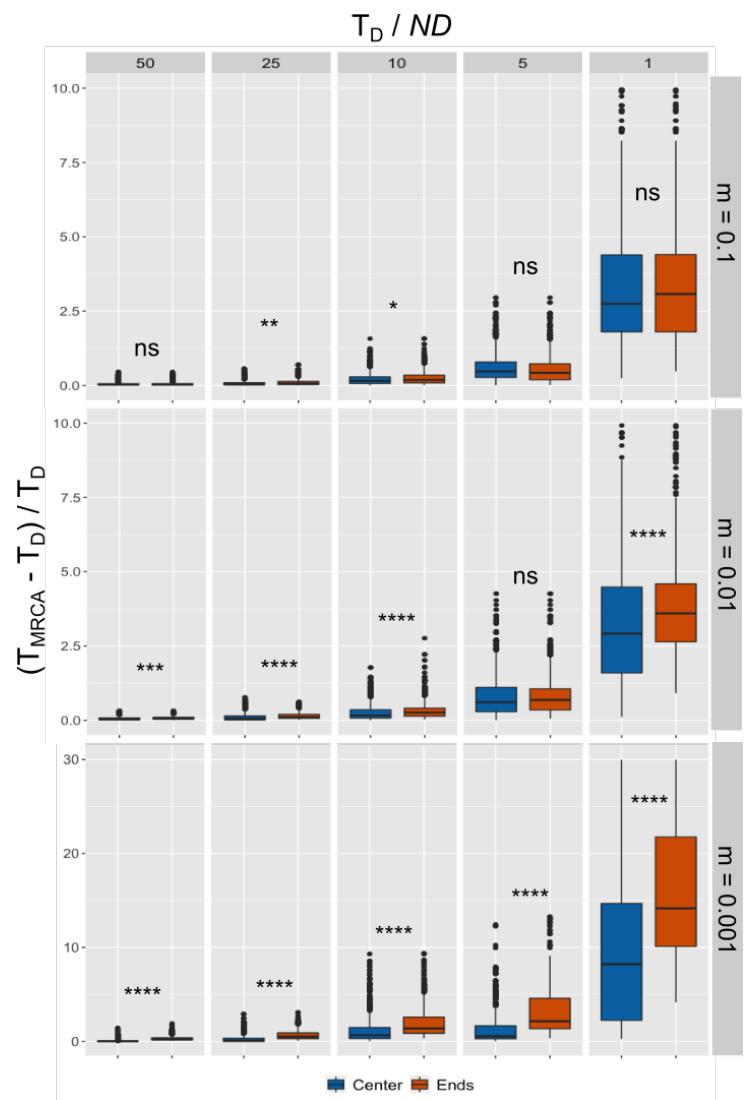
T <sub>D</sub> / ND	Deme sampled	Migration rate	True Age (T <sub>D</sub> )	Expected Estimate (eT <sub>D</sub> )	Actual Estimate (T <sub>MRCA</sub> )	T <sub>MRCA</sub> - eT <sub>D</sub>	(T <sub>MRCA</sub> - eT <sub>D</sub> ) / eT <sub>D</sub>
50	End	0.1	50000	52000	32929	-19071	-0.36675
50	End	0.01	50000	52000	36330	-15670	-0.301346154
50	End	0.001	50000	52000	40028	-11972	-0.230230769
50	Center	0.1	50000	52000	46270	-5730	-0.110192308
50	Center	0.01	50000	52000	54301	2301	0.04425
50	Center	0.001	50000	52000	49004	-2996	-0.057615385
25	End	0.1	25000	27000	25893	-1107	-0.041
25	End	0.01	25000	27000	26492	-508	-0.018814815
25	End	0.001	25000	27000	48430	21430	0.793703704
25	Center	0.1	25000	27000	22264	-4736	-0.175407407
25	Center	0.01	25000	27000	25488	-1512	-0.056
25	Center	0.001	25000	27000	26147	-853	-0.031592593
10	End	0.1	10000	12000	12439	439	0.036583333
10	End	0.01	10000	12000	14085	2085	0.17375

**Table 3.3.** Continued.

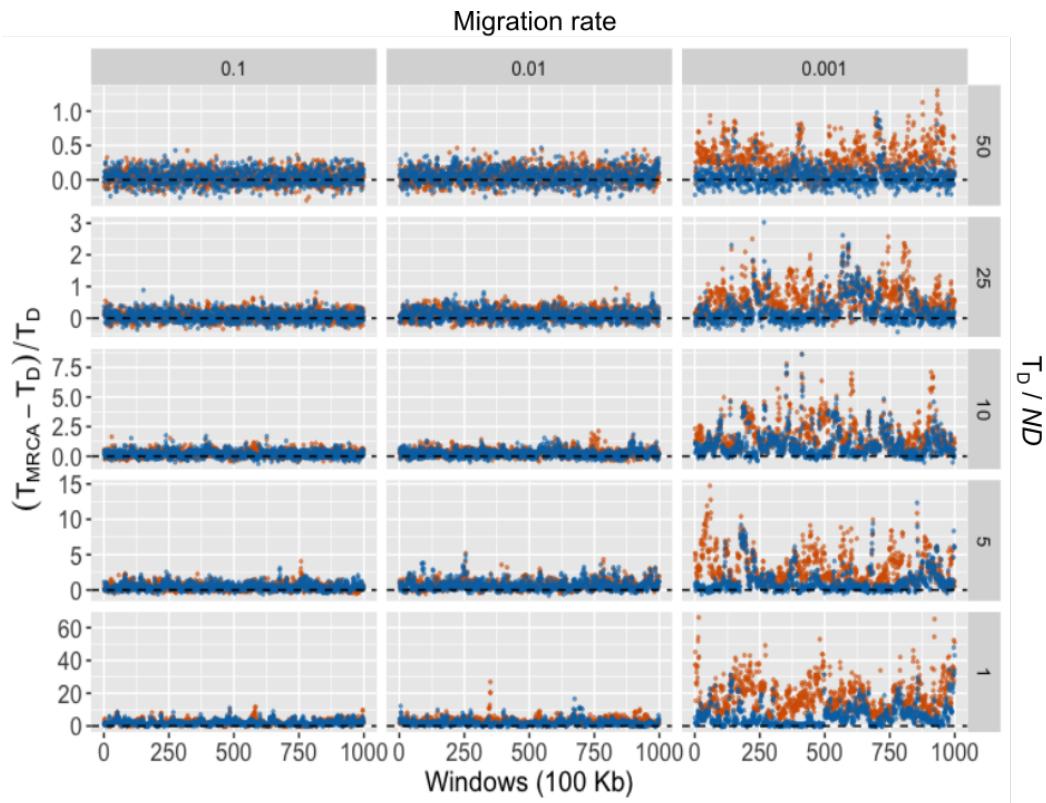
T <sub>D</sub> / ND	Deme sampled	Migration rate	True Age (T <sub>D</sub> )	Expected Estimate (eT <sub>D</sub> )	Actual Estimate (T <sub>MRCA</sub> )	T <sub>MRCA</sub> - eT <sub>D</sub>	(T <sub>MRCA</sub> - eT <sub>D</sub> ) / eT <sub>D</sub>
10	End	0.001	10000	12000	28629	16629	1.38575
10	Center	0.1	10000	12000	13234	1234	0.102833333
10	Center	0.01	10000	12000	13040	1040	0.086666667
10	Center	0.001	10000	12000	21782	9782	0.815166667
5	End	0.1	5000	7000	10618	3618	0.516857143
5	End	0.01	5000	7000	12730	5730	0.818571429
5	End	0.001	5000	7000	19936	12936	1.848
5	Center	0.1	5000	7000	11144	4144	0.592
5	Center	0.01	5000	7000	10595	3595	0.513571429
5	Center	0.001	5000	7000	11697	4697	0.671
1	End	0.1	1000	3000	1512.9	-1487.1	-0.4957
1	End	0.01	1000	3000	2700	-300	-0.1
1	End	0.001	1000	3000	23845	20845	6.948333333
1	Center	0.1	1000	3000	1513.2	-1486.8	-0.4956
1	Center	0.01	1000	3000	2649.1	-350.9	-0.116966667
1	Center	0.001	1000	3000	1647.2	-1352.8	-0.450933333



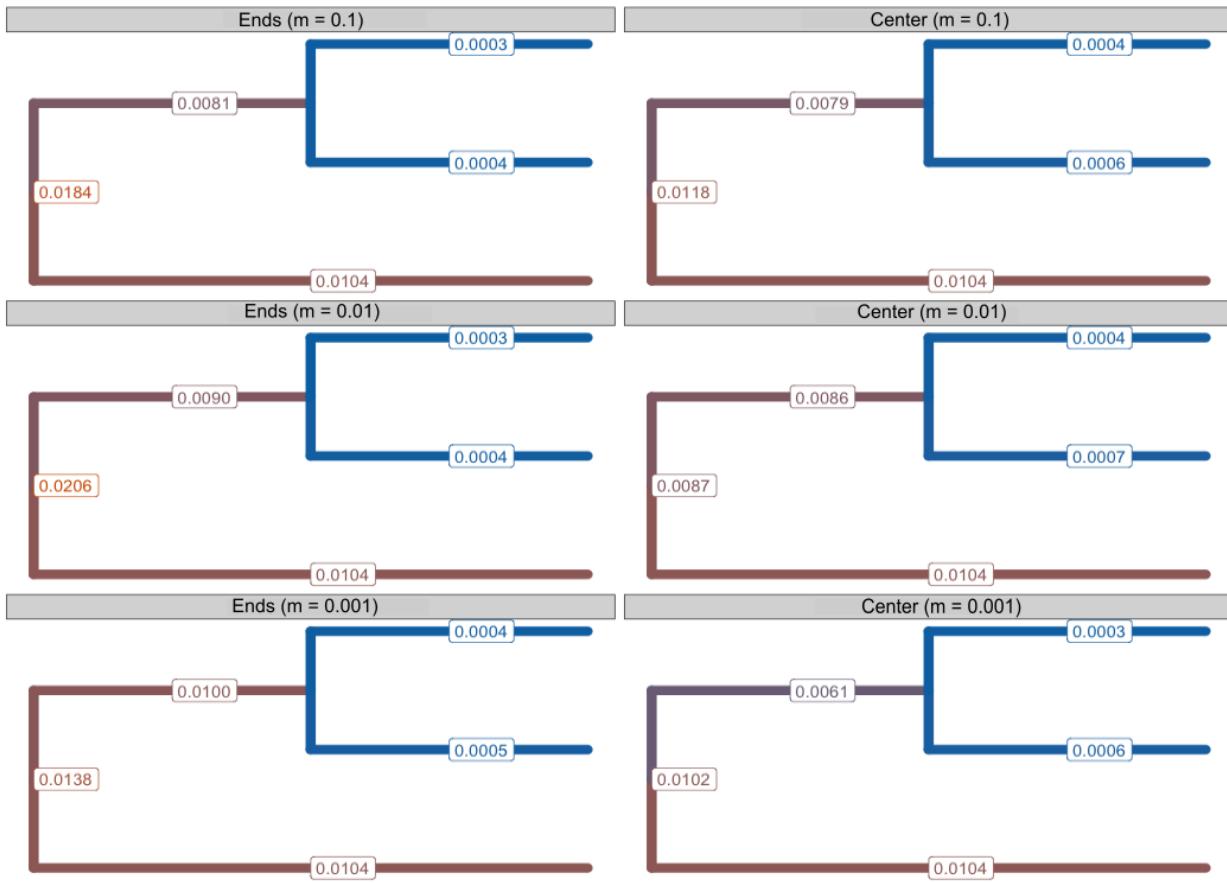
**Figure S3.1.** Plots of  $\log((T_{\text{MRCA}} - T_D) / T_D)$  against  $T_D / ND$  for each migration rate (0.1, 0.01, 0.001). Each point is the mean of 1000 simulations; lines around points are 95% confidence interval. Y-axis has been log-transformed to aid in visualizing differences between model/deme sampled.



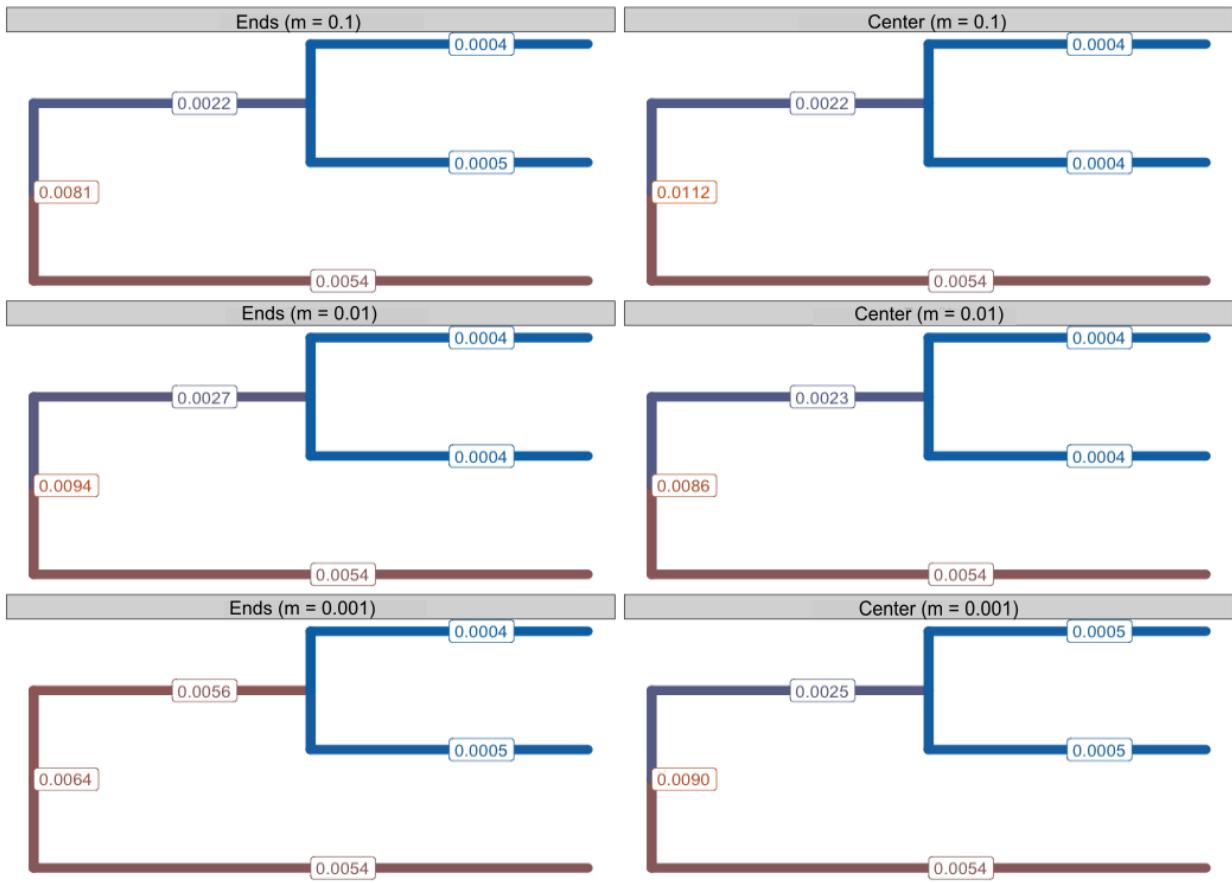
**Figure S3.2.** Boxplot of coalescent times ( $T_{\text{MRCA}}$ ) across the genome, where times have been converted into proportions of the population divergence time ( $T_D$ ); ns = “not significant”,  $p < 0.05$  (\*),  $p < 0.001$  (\*\*),  $p < 0.0001$  (\*\*\*),  $p < 0.00001$  (\*\*\*\*). Note that the y-axis differs between panels.



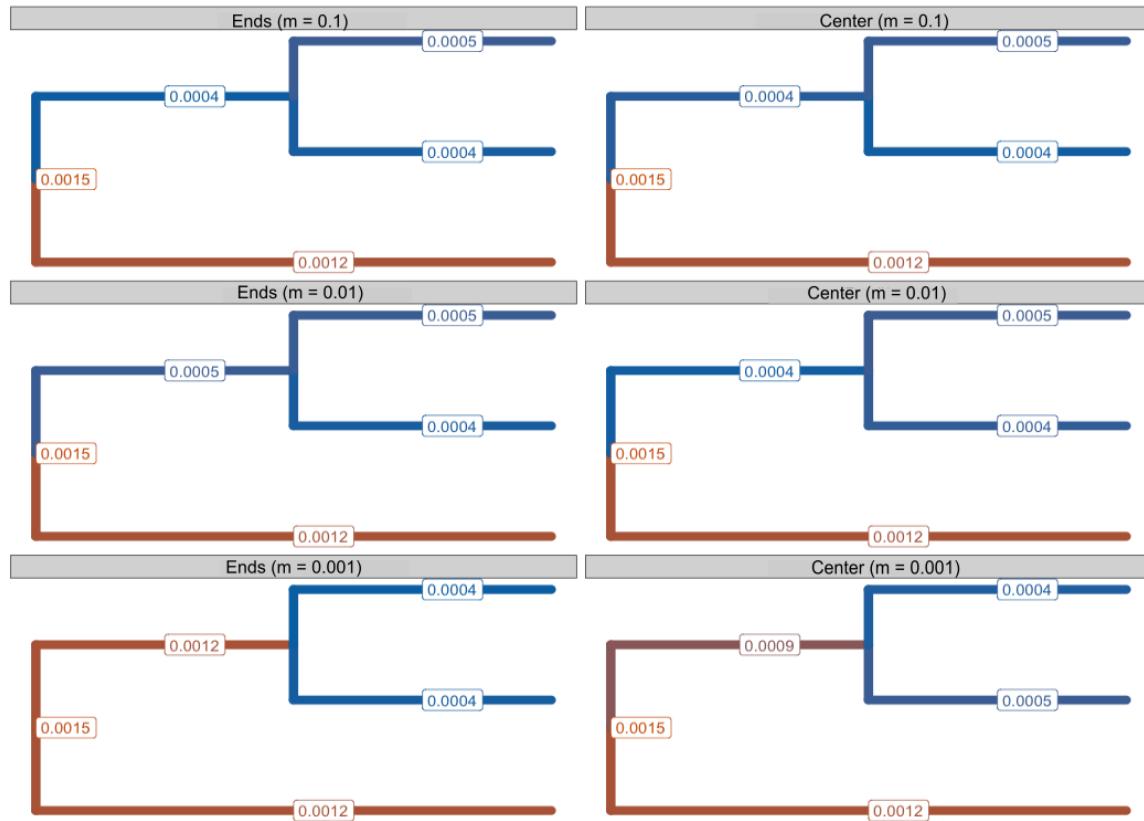
**Figure S3.3.** Genome-wide divergence times based on  $\pi_{12}$ . Divergence times are estimated as  $\pi_{12} / 2\mu$  and evaluated in 100 Kb windows. The y-axis is the scaled proportion of overestimation, where  $T_{\text{MRCA}}$  is the estimated age and  $T_D$  is the true age. The dashed line represents the value at which these two converge (i.e., 0). Center (blue), ends (orange). Note the y-axis differs between the panels.



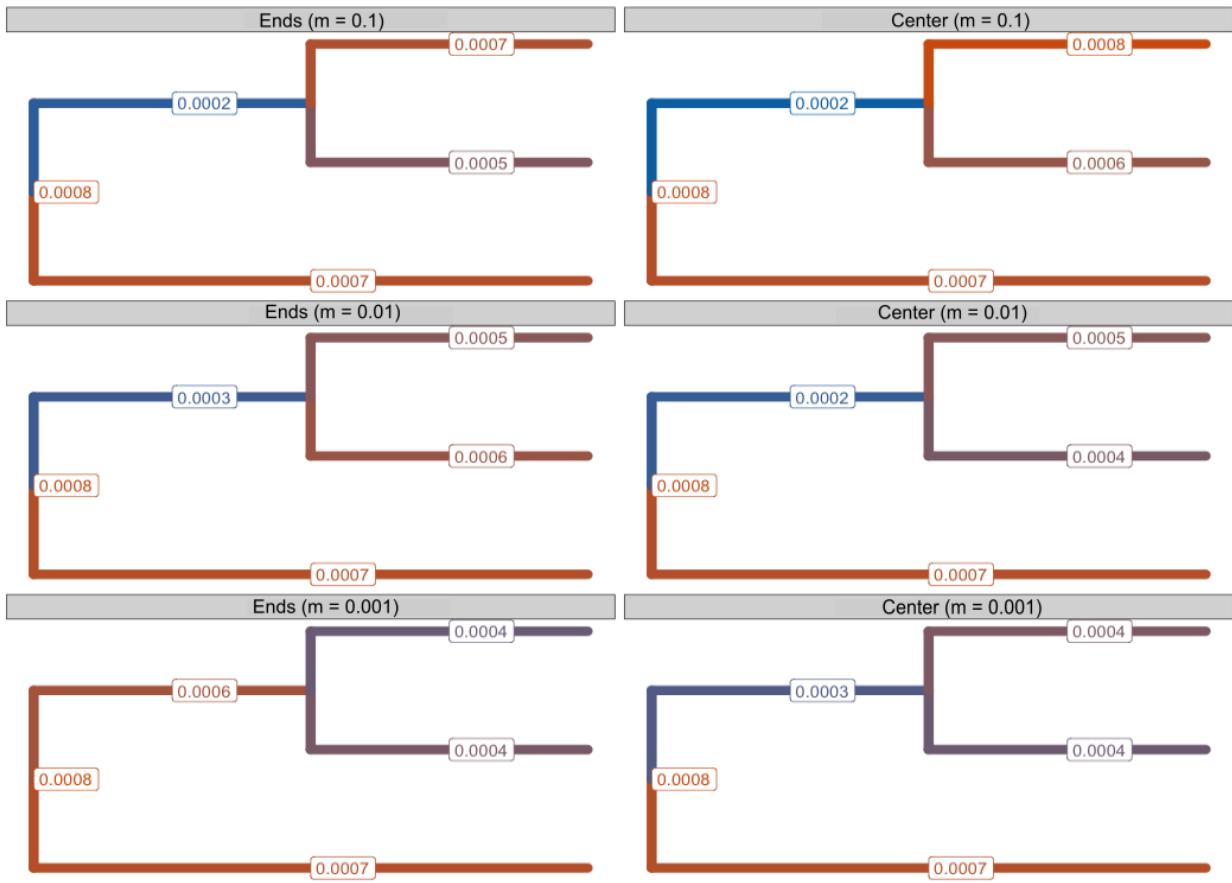
**Figure S3.4.** Estimates of  $\theta$  in SNAPP for  $T_D / ND = 50$ . Branch labels are estimated  $\theta$ .



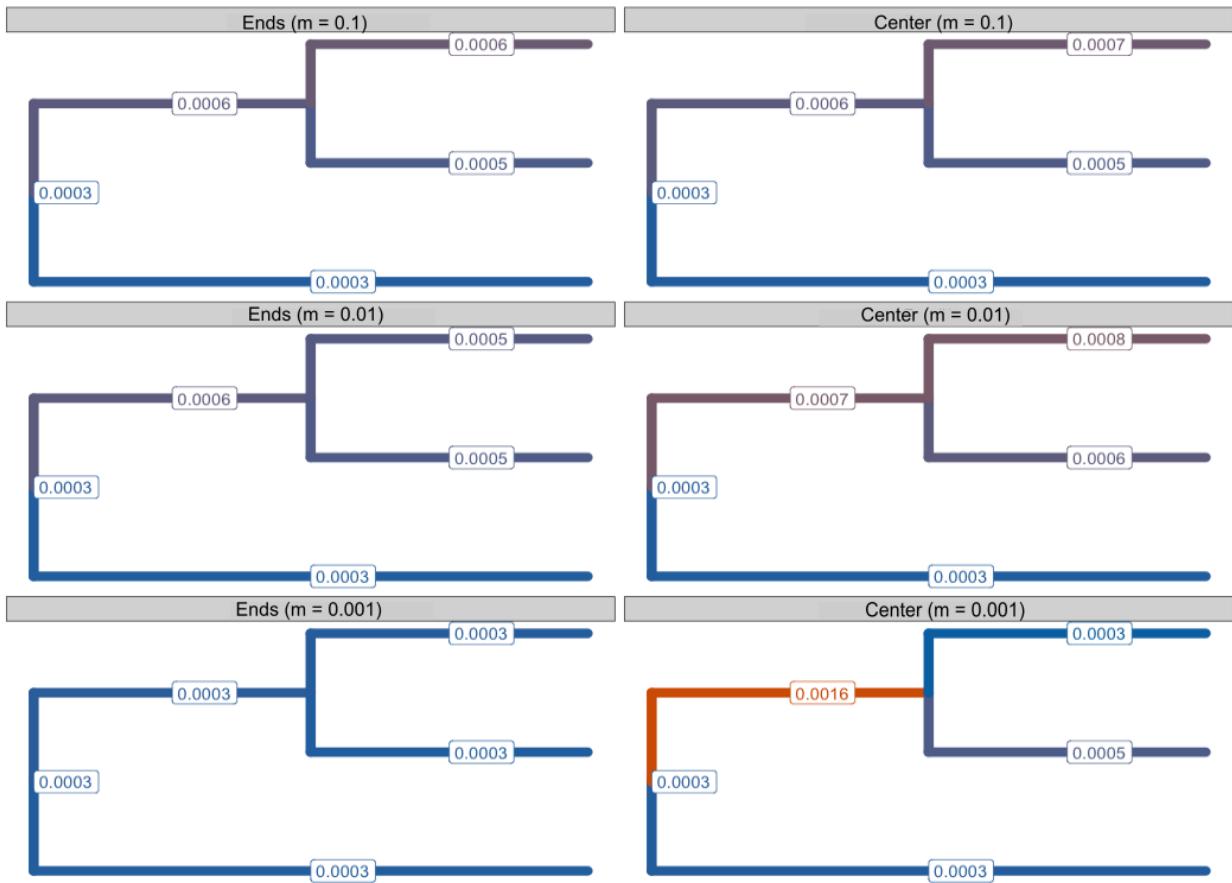
**Figure S3.5.** Estimates of  $\theta$  in SNAPP for  $T_D / ND = 25$ . Branch labels are estimated  $\theta$ .



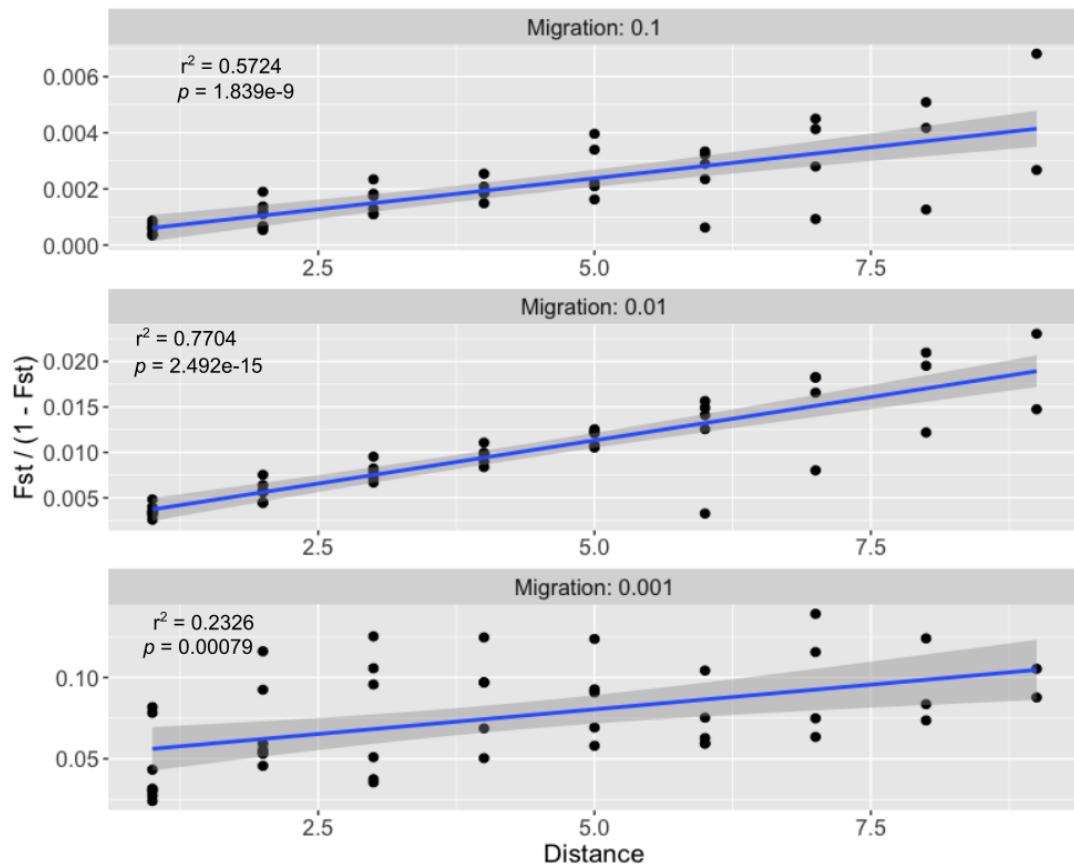
**Figure S3.6.** Estimates of  $\theta$  in SNAPP for  $T_D / ND = 10$ . Branch labels are estimated  $\theta$ .



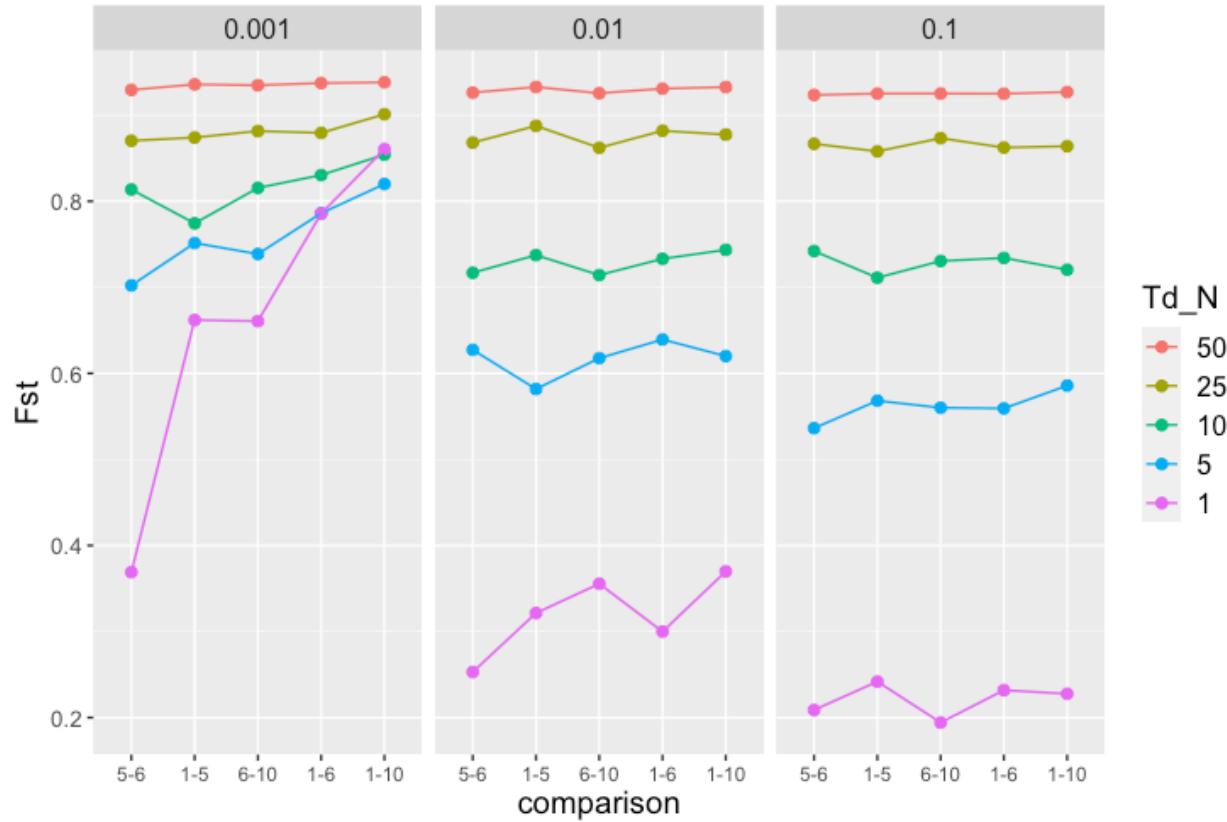
**Figure S3.7.** Estimates of  $\theta$  in SNAPP for  $T_D / ND = 5$ . Branch labels are estimated  $\theta$ .



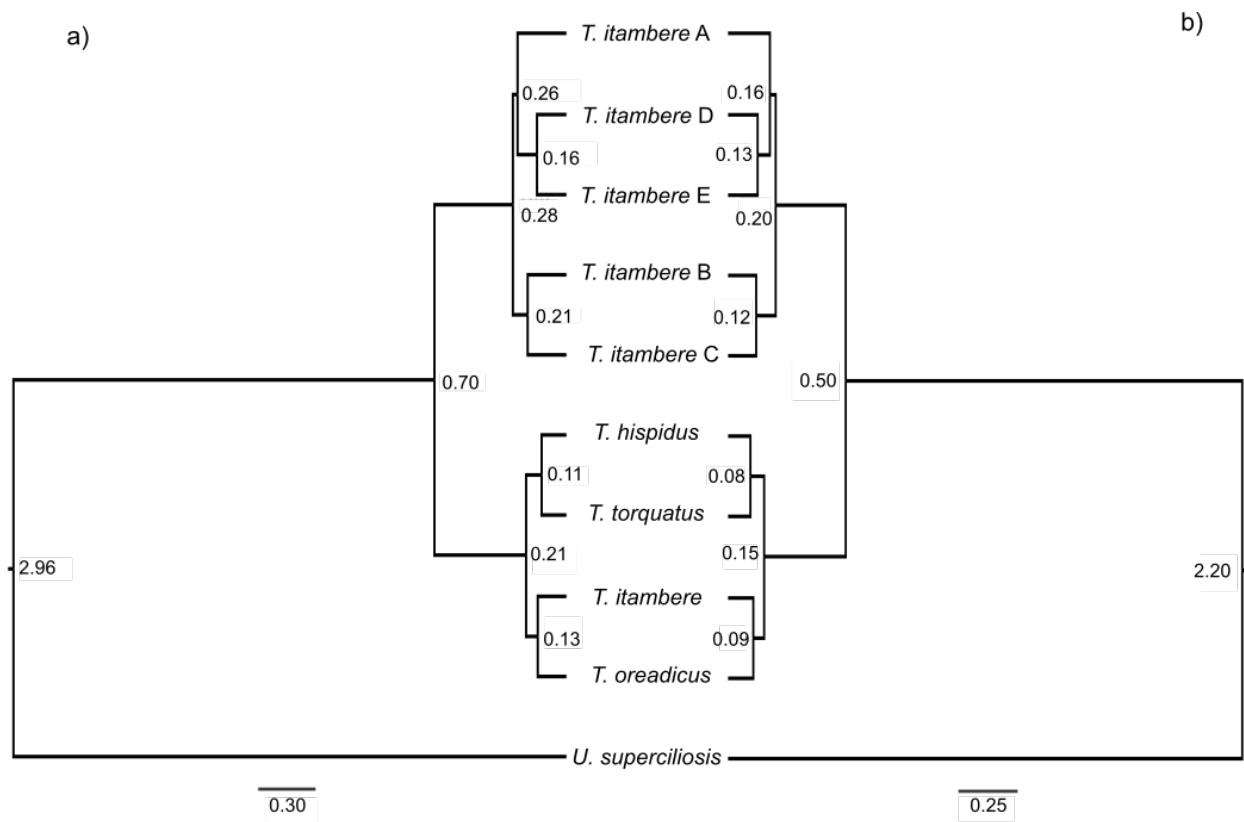
**Figure S3.8.** Estimates of  $\theta$  in SNAPP for  $T_D / ND = 1$ . Branch labels are estimated  $\theta$ .



**Figure S3.9.** Isolation-by-distance plots for three migration rates in the ancestral population. Pairwise  $F_{ST}$  was calculated between each deme in the ancestral population prior to the split to verify that a pattern of IBD had occurred. Note that the y-axis differs between panels.



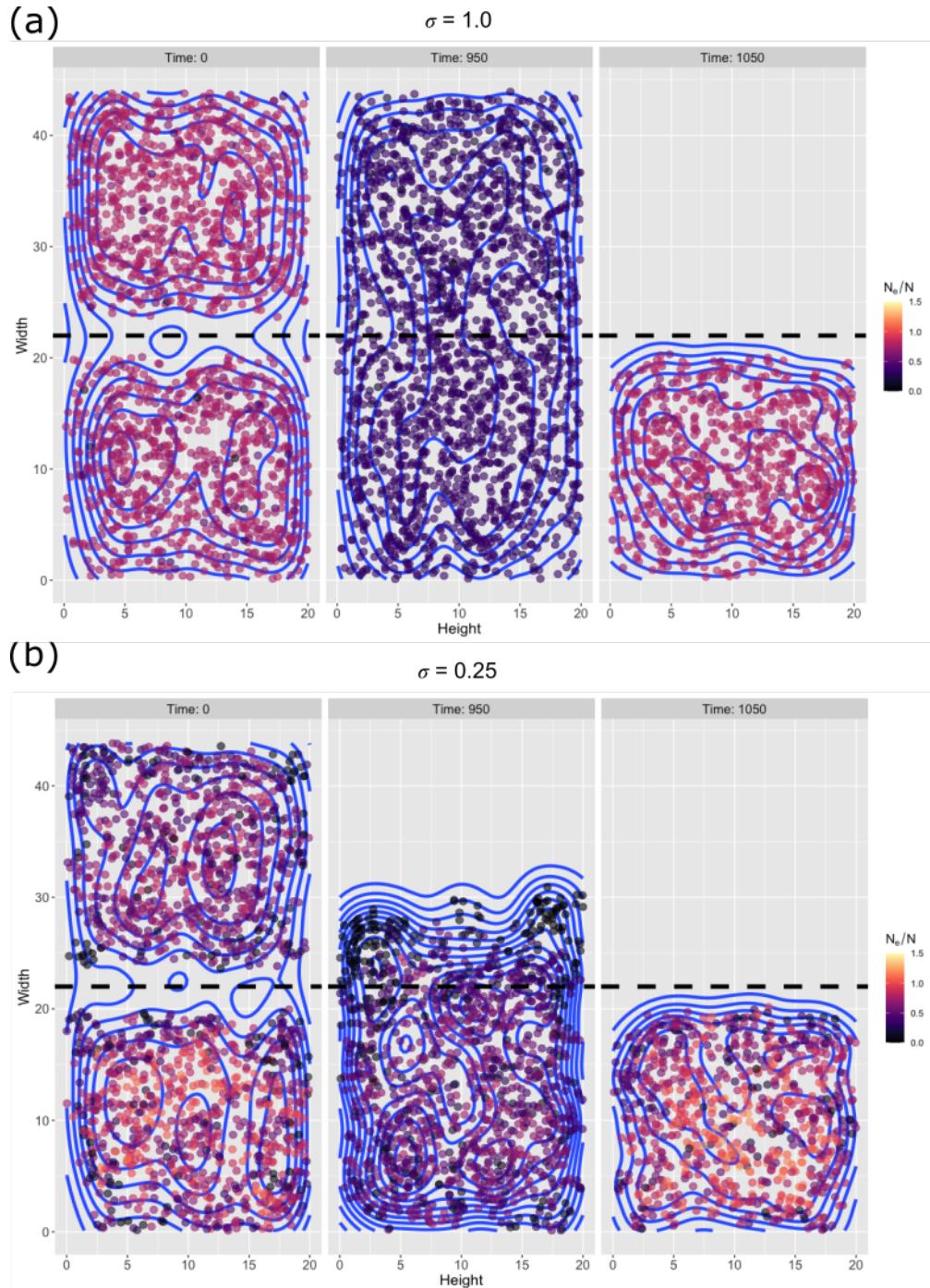
**Figure S3.10.** Pairwise  $F_{ST}$  for sampled end and center species. Comparisons are population numbers from Fig. 1; these are arranged linearly from 1–10.



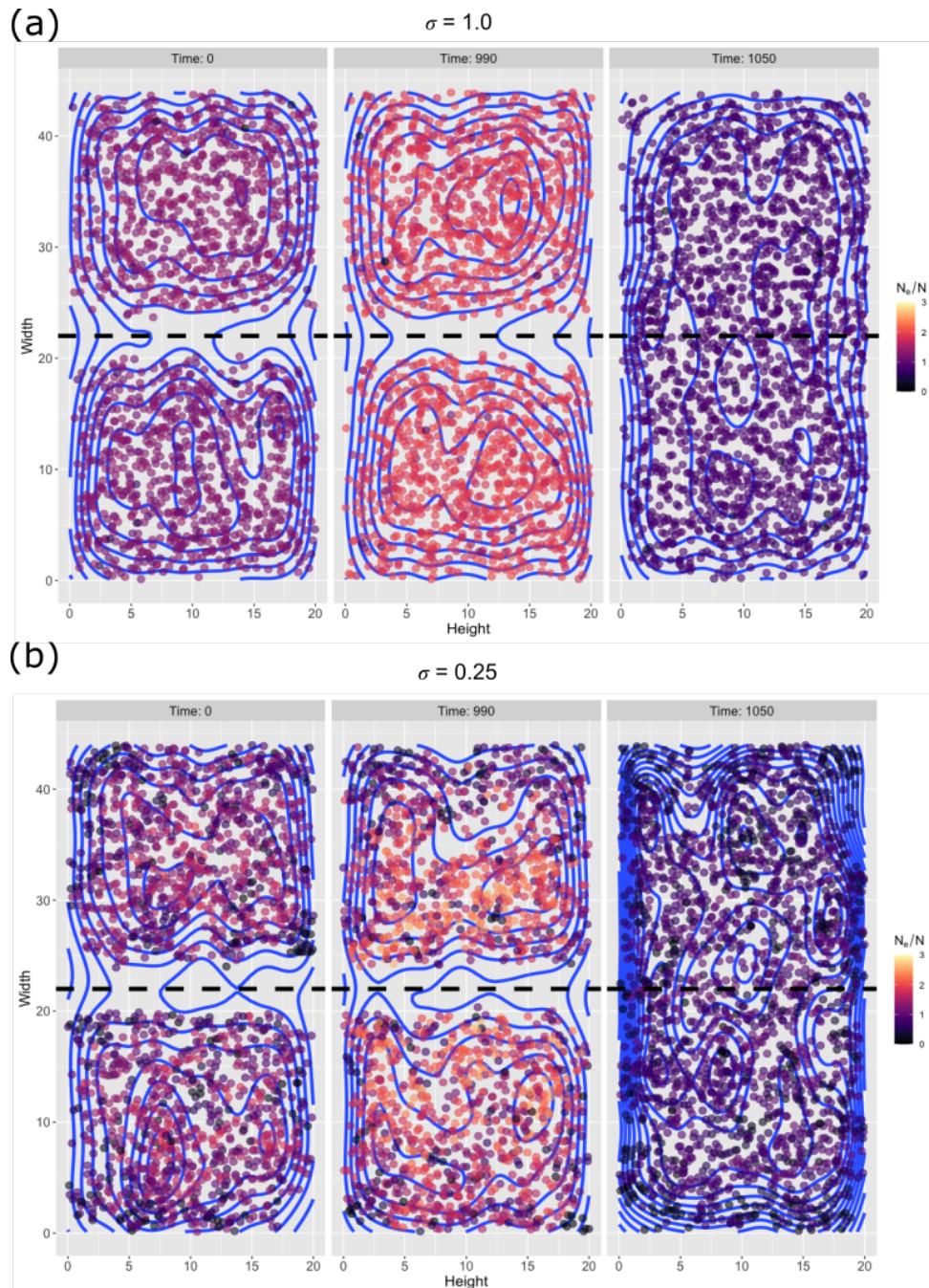
**Figure S3.11.** Estimated median heights in \*BEAST with a relaxed lognormal clock on all loci. Estimated node ages are in units of millions of generations (mga). Note that the scale bars are different between the two trees. A) “far” dataset; B) “near” dataset.

## CHAPTER IV

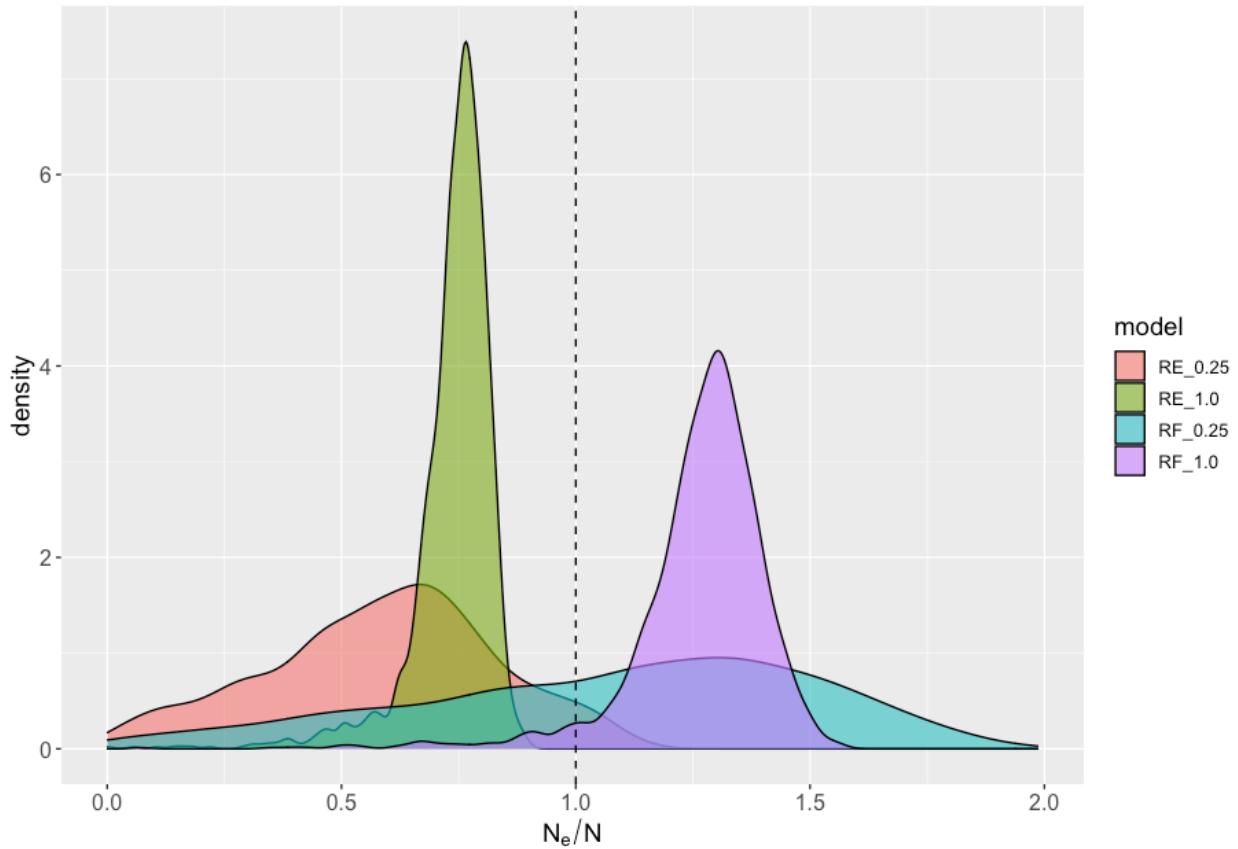
**Figure S4.1.** Distribution of  $N_e / N$  across the range for both clustered ( $\sigma = 0.25$ ) and unclustered ( $\sigma = 1.0$ ) models for the peripatric speciation scenario. Dotted line represents location of population split. Right panel is before the expansion; central panel is the expansion phase; left is after the expansion.



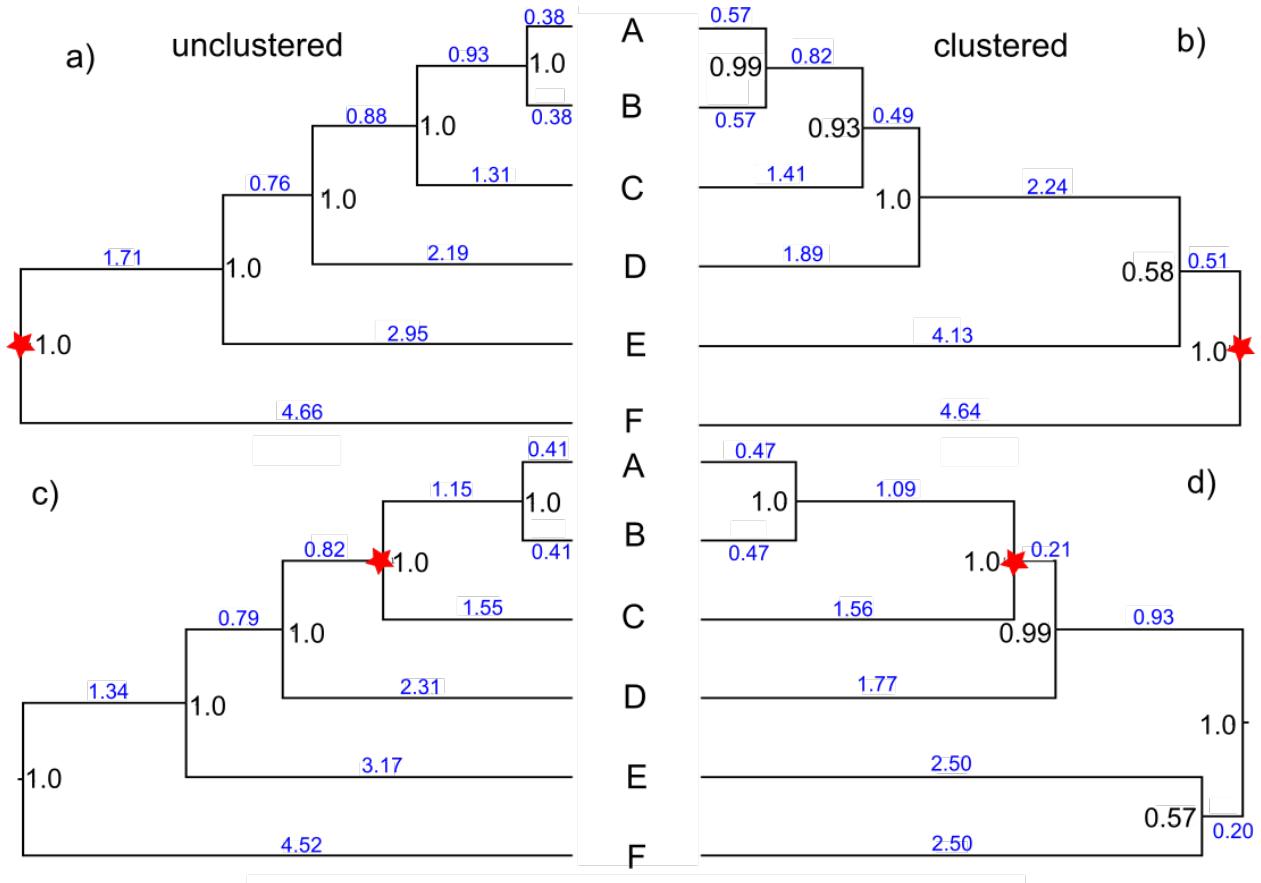
**Figure S4.2.** Distribution of  $N_e / N$  across the range for both clustered ( $\sigma = 0.25$ ) and unclustered ( $\sigma = 1.0$ ) models for the vicariant speciation scenario. Dotted line represents location of population split. Right panel is prior to the split; middle is 10 generations after; left is 990 generations after.



**Figure S4.3.** Density of  $N_e / N$  following the split for each model. RE\_0.25 = peripatric,  $\sigma = 0.25$ ; RE\_1.0 = peripatric,  $\sigma = 1.0$ ; RF\_0.25 = vicariant,  $\sigma = 0.25$ ; RF\_1.0 = vicariant,  $\sigma = 1.0$ . Dotted line represents when  $N_e = N$ .



**Fig. S4.4.** Consensus trees produced by \*BEAST2 with node calibrations signified by red stars. Node labels are posterior probabilities; branch labels are estimated lengths in substitutions.



**Fig. S4.5.** Consensus trees from the SNAPP analysis. Node labels are estimated divergence times in numbers of substitutions per site; branch labels are estimated  $\theta$  values.

