# Diversification of spiders on volcanic islands from the Pacific Ocean

By

Darko Davor Cotoras

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy
in
Integrative Biology
in the
Graduate Division
of the
University of California, Berkeley

Committee in charge:
Professor David R. Lindberg, Co-Chair
Professor Rosemary G. Gillespie, Co-Chair
Professor Charles R. Marshall
Professor George K. Roderick

Fall 2014

Abstract

Diversification of spiders on volcanic islands from the Pacific Ocean

By

Darko Davor Cotoras

Doctor of Philosophy in Integrative Biology

University of California, Berkeley

Professor David R. Lindberg, Co-Chair

Professor Rosemary G. Gillespie, Co-Chair

What is the effect of isolation in the formation of new lineages? Volcanic islands provide an ideal scenario to explore this question due to their various degrees of isolation, environmental variability and generally understood geology. On the other hand, the dispersal abilities, diverse ecology, and known natural history of spiders make them a suitable group of study to address this big question.

Because, the formation of new clades on islands start with a colonization event, I first studied long distance dispersal across different Pacific archipelagoes for spiders of the genus *Tetragnatha*. Then, I focused on the early stages of diversification for this group in the Hawaiian archipelago using population genetics studies combining Sanger sequencing of mitochondrial genes and Exon Capture for nuclear markers. Finally, I studied the effect of isolation at the morphological level. In particular, by looking the convergent evolution of color polymorphisms among species present on islands and continental areas. For this I examine the case of the family Theridiidae.

Colonization of remote archipelagoes corresponds to a very rare event. The Long-jawed spiders, genus *Tetragnatha*, are widely distributed around the world including oceanic islands. In Hawai'i there are 37 species of one or two colonization events from North America. In the Marquesas there are 5 species and phylogenetic evidence suggests an origin somewhere in the Asia. For the species present in the Society Islands, Rapa Nui and *D. raptor* (Hawai'i), their biogeographic origin is still unsolved. In order to examine colonization patterns across the Pacific Rim, I used freshly collected and museum specimens for phylogenetic analysis. The DNA from historic collections was sequenced using newly design primer pairs to amplify overlapping short fragments of COI. The phylogenetic reconstruction shows two of the species from the Society Islands (*T. rava* and *T. tuamoaa*) and *D. raptor* in the same clade as other species from Asia. *T. moua* (Society Islands) is sister to *T. nitens*, which has a circumtropical distribution making strong biogeographic statements difficult. The species from the Marquesas appear linked to a basal polytomy, while the one from Rapa Nui is in the same clade as the Hawaiian species. Other new species from Hawai'i were added into the analysis, but their phylogenetic position within the Web building clade is not conclusive. The first chapter provides preliminary information about origin of *Tetragnatha* spiders in Pacific islands. The implementation of Next Generation Sequencing technologies and inclusion of more species could improve this understanding.

Once the founder event occurs lineages could follow different diversification trajectories. Adaptive radiation is one of them in which many species evolve on a short period of time occupying different ecological niches. The study of their temporal dynamic it is not a simple problem to address in natural conditions. However, the chronosequence of the Hawaiian Islands provides "snapshots" of different stages of the diversification process. Here, I propose a population-based mechanism, which associated with the changes on the landscape attempts to explain the patterns of species diversity. By comparing species present on the young (Big Island: *T. anuenue* and *T. brevignatha*) and middle age islands (Maui: *T. waikamoi* and *T. brevignatha*), I tested the following predictions: (1) higher genetic structure in the young island populations and (2) lower genetic structure in populations from the middle age island. To test these hypotheses I sequenced three mitochondrial genes (COI, ND1 and Cytb). For my first prediction I found that the mitochondrial haplotype distribution is not related to the specific volcano configuration, but rather with breaks in forest types associated with rain regimens. The second prediction was not supported. Instead, I found strong genetic isolation between populations on Maui. Finally, a time calibrated coalescent reconstruction of *T. brevignatha* provided evidence for a population expansion during Pleistocene glacial periods. It also suggests that the Windward/Leeward population break is related to more recent events than the original colonization of the island. Species evolving in Hawai'i seem to be affected not only by the geologic development of the islands, but also by more recent climatic events.

On this context, I studied the genetic structure of newly formed species. In the third chapter I examined the genetic signatures of young speciation events by applying the transcriptome-based Exon Capture approach to the study of three closely related species (*T. brevignatha, T. waikamoi* and *T. macracantha*) from Maui Nui and Big Island. The combined data showed the existence of 5 clades: *T. brevignatha* Big Island, *T. brevignatha* Maui, *T. macracantha*, *T. waikamoi* and a WaikBrevMac clade. The last one includes specimens from all three nominal species. For *T. brevignatha*, the molecular divergence between the populations of Maui and Big Island are similar to "species level" comparisons. The Big Island population shows a strong break between the Leeward and Windward populations. For *T. waikamoi* nuclear data suggests that the population break between Haleakalā and West Maui is less pronounced than the break based on mitochondrial data. In the case of *T. macracantha*, we found a separation between the Maui and Lana'i populations. Moreover, there are two very divergent lineages co-existing in Kīpahulu Valley. Lastly, the finding of the WaikBrevMac group adds a new layer of complexity where recent speciation events suggest divergent lineages with convergent (or are the ancestral) morphologies and no genetic admixture. The addition of species with different eco-morphologies will complement these findings and also inform about the genetic signatures of ecological speciation.

Finally, to examine the effects of isolation in morphology I addressed the phenomena of convergent evolution in a color polymorphism on island and continental species. This kind of convergence has been widely studied, as it corresponds to a unique situation where the convergence leads to diversity. In the spider family Theridiidae, the independent evolution of the abdominal color polymorphism has been well described in at least four species (*Enoplognatha ovata*, *E. latimana*, *Theridion grallator* and *T. californicum*). Among the shared characteristics on these cases of color polymorphism

are: (1) the presence of "Yellow" as the double recessive and more common variant, (2) single loci Mendelian inheritance (except in Big Island population of *T. grallator*), (3) constant frequencies of variants among poorly connected populations and (4) evidence of the action of natural selection. The genus *Selkirlkiella* from the temperate rainforest of southern South America has several species with some degree of color polymorphism. Here, I documented the color polymorphisms of *S. alboguttata* (Robinson Crusoe Island) and *S. luisi* (Valdivia, Chile). *Selkirlkiella alboguttata* displays 6 morphs and *S. luisi* two morphs. The "Yellow" morph was the more common in both species. Based on a molecular phylogeny, we confirm that this genus is closely related to *Enoplognatha*. The presence of color polymorphism in this genus appears to be an event of convergent evolution at the family level, while between species it is likely due to common ancestry. There is also evidence of other cases of color polymorphism in the family. Finally, this phenomenon seems to be associated with the "under leaf community".

Spiders are a very diverse group of organisms. Due to their exceptional dispersal abilities, they have colonized even the most remote islands on the planet. The implementation of new genetic methods as well as the collection of specimens from remote areas will reveal even more insights about the biology of these organisms. This knowledge will extremely valuable to expand the understanding of the effects of isolation in the generation of biodiversity.

*To my parents, who showed me the wonders of Life,*
*and my siblings with whom I share them.*

"I swept the ground with the beam from my headlamp for signs of life, and found-diamonds! At regular intervals of several meters, intense pinpoints of white light winked on and off with each turning of the lamp. They were reflections from the eyes of wolf spiders, members of the family Lycosidae, on the prowl for insect prey. When spotlighted the spiders froze, allowing me to approach on hands and knees and study them almost at their own level. I could distinguish a wide variety of species by size, color and hairiness. **It struck me how little is known about these creatures of the rain forest, and how deeply satisfying it would be to spend months, years, the rest of my life in this place until I knew all the species by name and every detail of their lives.** From specimens beautifully frozen in amber we know that the Lycosidae have survived at least since the beginning of the Oligocene epoch, forty million years ago, and probably much longer. Today a riot of diverse forms occupy the whole world, of which this was only a minutest sample, yet even these species turning about to watch me from the bare yellow clay could give meaning to the lifetimes of many naturalists."

- The Diversity of Life -

Edward O. Wilson, co-author of the
Theory of Island Biogeography

## ACKNOWLEDGEMENTS

"Gratitude is the memory of the heart", a friend once told me after giving me a ride from the Berkeley Marina. At the end of my Doctorate program, I have my heart full of gratitude and joy as when I look back I see how many people have helped me on this process. Along the whole way or at a particular stage different people have given me a hand, shared his/her knowledge or simply been there.

I would like to start acknowledging my advisors who received me as a student and gave me the intellectual freedom to explore my biological interests. The support to this freedom is perhaps what I appreciate the most and what has helped me to mature as a scientist. I want to acknowledge Rosie for her patience and constant interest in my questions and projects. Thank you for always be there to listen to me, review what I write (including the grammar!), discuss ideas, plan fieldwork, encourage me to do a question based research that fits on a bigger picture and look the spiders that I brought from the islands. I really admire the work that you have done with the Pacific *Tetragnatha* and I consider it as a model to follow. I want to acknowledge Dave for his sincere and wise words. Not only about biology, but also about life in general. Our conversations helped me to put things in perspective and integrate the complexity of the biological phenomena. I admire the breadth of your knowledge, which does not compromise the details and intimate understanding of the organisms. I will keep working to be a real naturalist like you are. I would also like to acknowledge Nipam for receiving me into his lab in my first year and having the wisdom to let me follow my interests as they shifted during that year. Another professor that I consider as an advisor, even if he has not been formally, is Charles Griswold. Thank you for teaching me about field entomology, helping me organize fieldwork (Rapa Nui) and letting me work with the samples from the Cal Academy. I have learnt a lot from you and I am very excited for all that we will learn together during the PostDoc!

My qualifying and dissertation committee have also been important in my process of intellectual growth. Thank you professors Tony Barnosky, Paul Fine, Charles Marshall and George Roderick for many hours of conversation about evolution and diversification. I have to honestly say that I really enjoyed the preparation for my Quals. I appreciate your time and how much you pushed my limits during the preparation and the actual exam. For my dissertation I want again to acknowledge Charles and George for overseeing my progress and helping me to improve my research.

I have a very special thank you for Pete Croucher (Illumina) and Michael Brewer (East Carolina University). Both accompanied me in the daily life of the lab, teaching me most of what I know about lab work and bioinformatics. I appreciate your dedication, good will and all the knowledge you shared with me. To make an idea into something real you need to know how to do it and you taught me that.

My lab mates from the EvoLab and LindLab were also an important support that accompanied my learning, successes, failures and frustrations. A special thanks for the people with whom I interacted more closely, from the EvoLab: Andy, Athena, Ashley, Brad, Cerise, Curtis, David, Eli, Ellie, Emily, Hannah, James, Jenn, Jennifer, Jing, Joanne, Jun, Kari, Kelly, Lauren, Leslie, Lisa, Margarita, Matt, Misha, Natalia, Pete O., Sonia, Susan and Traci; and from the Lindlab: Jenna, Jessie, Joey and Rose.

The majority of my research took place in Hawai'i and I had the good fortune to interact and learn from a lot of people that live or work there. Thank you for sharing your

knowledge, helping me to get to places, giving me advice, but especially thank you for giving me a hand when I needed it the most. You cannot imagine how happy I was when you answered my first e-mail or when you let me join your group in the field. Many mahalos to Laura Arnold, Timothy Bailey (HALE), David Benítez (HAVO), Katie Champlin (Limahuli Botanical garden), James Friday (UH Mānoa), Emory Griffin-Noyes (Limahuli Botanical Garden), Faith Inman-Narahari (UH Mānoa), Darcey Iwashita (UH Mānoa), Raina Kaholoaa (HALE), Jessie Knowlton (Michigan Tech), Rick Lapoint (U Arizona), Scott Laursen (UH Mānoa), Arthur Madeiros (USGS), Karl Magnacca (O'ahu Army Natural Resources Program), Patrick O'Grady (UC Berkeley), Rita Pregana (HAVO), Donald Price (UH Hilo), David Rankin (U Vermont), William Roderick (Stanford), Karen Uy (UH Hilo), Erin Wilson (UC Riverside) and the Kīpuka team. I want to give a special thank to Liz Morrill (SFSU) and Andy Rominger (UC Berkeley) for joining me on one and two, respectively, field seasons. I really enjoyed working with you, guys!

All the collecting work would have not been possible without the support and help of the people working on the different reserves. Thank you very much for protecting the land and letting me visit it. Mahalo to Steve Bergfeld (DOFAW Big Island), Pat Bily (TNC Maui), Tabetha Block (HETF), Shalan Crysdale (TNC Big Island), Lance DaSilva (DOFAW Maui), Dean Danae (Kahoma Ranch), Charmian Dang (NAR), Melissa Dean (HETF), Betsy Gagne (NAR), Elizabeth Gordon (HALE), Lisa Hadway (DOFAW Big Island), Paula Hartzell (Lana'i Resorts, LLC), Greg Hendrickson (Kealakekua Ranch), Mel Johansen (TNC Big Island), Pomaikaʻi Kaniaupio-Crozier (Maui Land and Pinneapple), Cynthia King (DLNR), Peter Landon (NAR Maui), Rhonda Loh (HAVO), Russell Kallstrom (TNC Moloka'i), Joey Mello (DOFAW Big Island), Ed Misaki (TNC Moloka'i), Elliot Parsons (Pu'u Wa'awa'a HETF), Lani Petrie (Kapapala Ranch), Shawn Saito (Parker Ranch), Joe Ward (Maui Land and Pinneapple) and Kawika Winter (Limahuli Botanical Garden).

As part of my dissertation research I had the chance to work in other South Pacific islands. The fieldwork in Rapa Nui was supported by my collaborators Luis Flores and Cristian Villagra, both from the Instituto de Entomología, UMCE. On the island we received the invaluable help from the personal of Parque Nacional Rapa Nui (CONAF); in particular Graciela Campbell, Ninoshka Cuadros, Anthony Dubois, Omar Durán Veriveri, Pedro Hito, Pedro Lazo, Sergio Manuheuroroa, Raúl Palomino Matta, Hotu Paté, Carlos Salinas, Enrique Tucki and John Tucki; and the Armada de Chile. Maururu!

For the work in the Robinson Crusoe Island I have to start with a big "Thank you so much!" to Aaron Ramirez, for his invitation to join his expedition. This invitation was the door to an incredible discovery, which turned into my last chapter. For the work on the island I am very grateful of Iván Leiva, director of the Parque Nacional Archipiélago Juan Fernández, the park rangers Ramón Schiller and Alfonso Andaur, and the guide Rosa María Schiller for their field assistance. Moreover, I appreciate all the help provided by Javiera Meza (CONAF V Región) and Lynne Hollyer (UC Berkeley Industry Alliances Office) in obtaining the permits. I also want to acknowledge Gustavo Hormiga (George Washington University) and Miquel Arnedo (Universitat de Barcelona) for specimens collected on in Robinson Crusoe Island, as well Elizabeth Arias for her collections in Valdivia, Chile. I also want to thank Andrés Muñoz for a discussion about

the plant community associated with this spider. ¡Muchas gracias a todos!

I also had the chance to work in the continental tropics of Central America. In Costa Rica, I worked at the Biological Stations of the Organization for Tropical Studies (OTS). On my visit there, I was supported by staff members from La Selva (Carlos de la Rosa and Ronald Vargas), Las Cruces (Rodolfo Quiros and Zak Zahawi) and San José (Francisco Campos and Barbara Lewis); as well as graduate students and other researchers (James Crall (Harvard University), Bryan Folt (Auburn University), Linus Günther (Museum für Naturkunde), Dave Janas (Wilson Botanical Garden) and Luis Solis (Universidad de Costa Rica (UCR))). I also appreciate the kind help provided by professor Gilbert Barrantes and Emilia Triana at Universidad de Costa Rica; and Carlos Víquez and Carolina Ávila from the Instituto Nacional de la Biodiversidad. In Panamá, I had the opportunity to work at the station of Barro Colorado Island of the Smithsonian Tropical Research Institute (STRI). My work there was facilitated by staff members (Oris Acevedo, Lil Camacho, Hilda Castañeda, and Orelis Orasemena,) and graduate students/researchers (Stefan Brändel (Universität Ulm), William Eberhard (STRI/UCR), Egbert Leigh (STRI), Owen McMillan (STRI) and Diomedes Quintero (Universidad de Panamá)). I really appreciate all the knowledge that they shared with me, which allowed me to put on a bigger context all what I know about the biota on remote islands. I am also thankful of friends who received me and gave me good advices on their respective countries (Ignacio Escalante (UC Berkeley), Edgar Figueroa (Municipalidad de Panamá) and Daniel Soto (UCR). Muchas gracias/Thank you/Danke.

On September 2012, I had the luck to visit the Senckenberg Museum. Danke curator Peter Jeager for teaching me about spider morphology and Julia Altman for helping me navigate through the collection. My stay with your group was very productive and pleasant. Related to this work, I also want to thank Henrik Krehenwinkel (Max Plank Evolutionary Biology) for his advice on working with museum DNA.

I am deeply grateful for all the support and help that I received in the work related with Next Generation Sequencing. Many people who selflessly shared his/her knowledge accompanied my introduction to this "new world". This knowledge has turned into something very important in my professional development, which undoubtedly helped me to move into the next step. Thank you Ke Bi (Computational Genomics Resource Laboratory, California Institute for Quantitative Biosciences. UC Berkeley), Jacob Crawford (UC Berkeley), Emiliano Méndez (Universidad de Magallanes), Rasmus Nielsen (UC Berkeley), Tyler Linderoth (UC Berkeley), Sonal Singhal (UC Berkeley), Line Skotte (U Copenhagen), Lydia Smith (Evolutionary Genetics Lab, Museum of Vertebrate Zoology. UC Berkeley), Barker DNA Sequencing Facility (UC Berkeley) and Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley (supported by NIH S10 Instrumentation Grants S10RR029668 and S10RR027303).

A lot of people also shared his/her knowledge related to different aspects of my research, for that reason I want to thank Juliane Casquet (Muséum national d'Histoire naturelle. France) for her help and discussions on the first versions of the *Tetragnatha* phylogeny; Fernando Álvarez-Padilla (Universidad Nacional Autónoma de México) and Darrell Ubick (California Academy of Sciences) for collaborating with the identification of some of the specimens; Jonathan Price (UH Hilo) for discussions about the age of the wet forest on Leeward Big Island; Amy Vandergast (USGS) for advice at the beginning of the project; Geoff Oxford (University of York) for discussions about the

developmental mechanisms of the color polymorphism in the Hawaiian Happy Face Spider, and Becca Carter for showing me the color patterns of the *Tetragnatha*. Moreover, I want to thank to the whole community of Integrative Biology (IB) dept., Environmental Sciences, Policy and Management (ESPM) dept. the Berkeley Natural History Museums and the Entomology dept. of the California Academy of Sciences. It has been enormously stimulating to interact with you all these years.

During my whole time at Berkeley I never had any administrative problem and that is something that should not be given for granted. For that reason I also want to thank all the people and institutions that supported me on that regard: Comisión Fulbright, Santiago; Fulbright Office, San Francisco, Unidad de Capital Humano avanzado, CONICYT, Berkeley International Office and the Integrative Biology dept.

My attendance to this PhD program and the realization of all the fieldwork and lab work would have never been possible without the support of several funding agencies. Thank you giving me this opportunity and supporting my education and research. My PhD was mainly funded by a Fulbright/CONICYT fellowship and a researcher position on the NSF Hawai'i Dimensions of Biodiversity Project. The fieldwork on Hawai'i and the visit to the Senckenberg Museum was funded by Graduate Research Allocation Committee (Integrative Biology dept. UC Berkeley), Summer Research Grant (Integrative Biology dept. UC Berkeley), Walker Grant (Essig Museum of Entomology), Graduate Division Travel Voucher Award, Sigma Xi grant and Research Grant Graduate Division UC Berkeley. The fieldwork in Rapa Nui was funded by the Tinker Grant (Center for Latin American Studies at UC Berkeley) and the Proyecto FONDECYT de Iniciación (Cristian Villagra and Luis Flores). The visit to the Juan Fernández archipelago was funded by Aaron Ramirez's grants (Tinker Gran -Center of Latinamerican Studies, UC Berkeley-, UC Berkeley Integrative Biology Summer Research Grant and the Graduate Research Fellowship Travel Grant -NSF-). The fieldwork in Central America was supported by the Pilot Research grant from OTS, Short term fellowship STRI, Walker Grant (Essig Museum, UC Berkeley), Graduate Research Allocation Committee (Integrative Biology dept. UC Berkeley) and Woodworth loan.

I also want to recognize people that supported me during part or my whole time in Berkeley, not necessarily in the development of the research, but with company and friendship. First, I want to acknowledge the constant and lovely support that my parents and siblings have given me during all these years. Your calls, messages and surprise packages always cheer me up and made me feel closer to you. Also, I want to thank a lot of friends that I had the fortune to meet on my time here. Some of them in IB, ESPM, the iHouse, the Chilean community on campus, Cal Academy, the Berkeley Marina and of course, to my roommates in the Milvia House; thank you all for making my time here so enjoyable!

Finally, I want to make a special recognition to the life of all the spiders collected as part of this research. It does not make sense to acknowledge them, because I am sure they would have been better of living in the forest. I just want to say that I believe that this research is important not only for us as humans and I will do everything in my power to take sure that we will learn as much as possible from their sacrifice.

At the end of this five and half years' journey through volcanic islands, rainforests, museums, field stations, labs, remote computer servers and classrooms my heart is full of gratitude and it is my desire to return all what I have received and more.

Thank you! - Mahalo nui loa - ¡Muchas gracias!

My first picture of a Hawaiian Happy Face spider. Probably, the first that I ever saw.
January 7th, 2010. Waikamoi Reserve. Maui.

# Curriculum Vitae

Name: Darko Davor Cotoras Viedma
Date of Birth: 03.16.1986
Place of Birth: Santiago, Chile
e-mail: darkocotoras@berkeley.edu
Web site: https://nature.berkeley.edu/evolab/?page_id=269

## 1.- Researcher positions

- September-November, 2014: Junior Specialist. NSF Hawai'i Dimensions of Biodiversity project. UC Berkeley
- January-August, 2014: Graduate Student Researcher. NSF Hawai'i Dimensions of Biodiversity project. UC Berkeley

## 2.- Education

- 2009 – 2014: PhD Integrative Biology. University of California, Berkeley
  Advisors: David Lindberg - Rosemary Gillespie
- 2008-2009: Master in Biological Sciences with mention in Ecology & Evolutionary Biology, Universidad de Chile (*Highest distinction*)
  Advisors: Miguel Allende – Verónica Cambiazo
- 2004-2007: Bachelor in Sciences with mention in Biology, Universidad de Chile. (*Highest distinction*. Class rank: 1st. Final grade: 6,5 (been 7,0 the maximum))
- 1991-2003: San Ignacio School

## 3.- Research interests:

Evolution & Biodiversity, Phylogenetics, Island Biogeography, Population Genetics, Evolutionary Developmental Biology (EvoDevo).

## 4.- Grants, fellowships and awards

- 2014: Hagey fund, California Academy of Sciences (Co-PIs: Charles Griswold and Brian Simison; US$36,028)
- 2014: Short Term fellowship. Smithsonian Tropical Research Institute (Fellow status)
- 2014 and 2012: Summer Research Grant. Graduate Division UC Berkeley (US$3,500)
- 2014-2011 (each year): Reshetko Family Scholarship Fund. UC Berkeley (US$1,700-3,000)
- 2014-2010 (each year): Walker Grant. Essig Museum of Entomology. UC Berkeley (US$1,000)
- 2013: Pilot research grant. Organization for Tropical Studies (US$2,200)
- 2012: Sigma Xi Research Grant (US$750)

- 2012: Best poster presentation. 36th Annual meeting American Arachnological Society.
- 2012: Tinker Grant. Center of Latinamerican Studies. UC Berkeley (US$2,196)
- 2012-2010 (each year): Research fellowship. UC Entomological Student Association. UC Berkeley (US$500)
- 2011: National Tropical Botanical Garden (NTBG) fellowship, McBryde Garden (US$1,250)
- 2011: Honorable Mention Border Crossing Essay Contest 2011. International House UC Berkeley
- 2010: 2nd Place Border Crossing Essay Contest 2010. International House UC Berkeley
- 2009: Fulbright/CONICYT Scholarship (4 years of full funding)
- 2008: Best Student Promotion 2008 Bachelor in Sciences with mention in Biology. Facultad de Ciencias. Universidad de Chile.
- 2008: Beca Líder, Fundación Carolina, España.
- 2007 and 2006: Academic Excellence Fellowship, Facultad de Ciencias, Universidad de Chile.
- 2004: Best entry score Bachelor in Sciences with mention in Biology. Facultad de Ciencias. Universidad de Chile

5.- Assistantships and teaching experience

- Fall 2014: Biogeography. Invited lecturer (Universidad de Costa Rica)
- Spring 2011: Biology 1B: Plant Diversity, Ecology and Evolution (UC Berkeley)
- Fall 2010: Evolution (UC Berkeley)
- Spring 2008: Developmental Biology (Universidad de Chile)
- Spring 2008: Evolution & Development. Invited lecturer (Universidad de Chile)
- Fall 2008 and 2009: General Zoology (Universidad de Chile)
- Fall 2007-2009: Comparative anatomy of vertebrates (Universidad de Chile)

6.- Undergraduate research experience

- Spring 2007: Research rotation: "Searching for common genes in vertebrate development of limb and tail" Developmental Biology Lab (PI: Miguel Allende)
- Spring 2006: Research rotation: "Comparison in the gene expression between tail and fin in zebra fish: Evolutionary implications" Developmental Biology Lab (PI: Miguel Allende)
- Spring 2005: Final work in Zoology II course. "Dietary differentiation between island-mainland in *Liolaemus pictus* (Duméril & Bibron, 1873)", Evolutionary Ecology Lab (PI: Rodrigo Medel)
- Fall-Spring 2004: Extracurricular activity in the Cellular Biology course. Developmental Biology Lab (PI: Miguel Allende)

7.- Field experience

- November/December 2014: Collecting spiders in Costa Rica and Panamá (6 weeks).

- June 2013: Collecting spiders in Kaua'i, Maui, Lana'i and Big Island (4 weeks).
- August 2012: Collecting terrestrial invertebrates in Rapa Nui (Isla de Pascua/Easter Island) (3 weeks).
- June 2012: Collecting spiders in Big Island, Molaka'i, and Maui (4 weeks).
- January 2012: Collecting spiders in Big Island of Hawai'i (3 weeks).
- August 2011: Collecting terrestrial invertebrates in Robinson Crusoe Island. Juan Fernández archipelago, Chile. (2 weeks).
- June 2011: Collecting spiders in Big Island of Hawai'i (4 weeks)
- August 2010: Collecting spiders in Big Island of Hawai'i (2 weeks).
- January 2010: Collecting spiders in Maui and Big Island of Hawai'i (2 weeks).
- January 2006: Field assistant in pollination studies in the Andes close to Santiago, Chile (day trips).

8.- Publications

- Rominger A., Roesch Goodman K., Lim J., Valdovinos F., Armstrong E., Becking L., Bennett G., Brewer M. Cotoras D., Ewing C., Harte J., Martinez N., O'Grady P., Percy D., Price D., Roderick G., Shaw K., Gruner D., Gillespie Community assembly on isolated islands: Macroecology meets evolution. *Global Ecology and Biogeography* (*submitted*)
- Brewer M.S. Cotoras D.D., Croucher P.J.P., Gillespie R.G. 2014. New sequencing technologies, the development of genomics tools, and their applications in evolutionary Arachnology. *Journal of Arachnology*. 42:1-15
- Razeto-Barry P., Díaz J., Cotoras D. and Vásquez R. 2011. Molecular evolution, mutation size and gene pleiotropy: a geometric reexamination. *Genetics* 187: 877-885
- Glavic A., Molnar C., Cotoras D. and de Celis J.F. 2009. *Drosophila* Axud1 is involved in the control of proliferation and displays pro-apoptotic activity. *Mech Dev*. 126: 184-97

9.- Participation in Congress

- Cotoras D. and Gillespie R.G. (*accepted* Oral presentation) The role of National Parks and other protected areas on preserving adaptive radiations: The case of the Hawaiian *Tetragnatha* spiders. Science for parks, Parks for Science: The next Century. May 25$^{th}$-27$^{th}$, 2015. Berkeley, CA. USA
- Cotoras D. and Gillespie R.G. (Oral presentation) Preliminary identification of protected areas with high diversity of *Tetragnatha* spiders on Big Island and Maui Nui. 22$^{st}$ Hawai'i Conservation Conference. July 15$^{th}$ -17$^{th}$, 2014. Honolulu, HI. USA
- Cotoras D. and Gillespie R.G. (Oral presentation) Understanding the initiation of adaptive radiation using comparative phylogeography of spiders. Island Biology Conference. July 7$^{th}$-11$^{th}$, 2014. Honolulu, HI. USA
- Cotoras D. and Gillespie R.G. (Poster) Revealing unknown levels of diversity: Discovery of color change during life history of the endemic spider *Tetragnatha*

*kamakou.* 21st Hawai'i Conservation Conference. July 16th -18th, 2013. Honolulu, HI. USA

- Gillespie R.G., Ewing C.P., Gruner D.S., Roesch Goodman K., Harte J., Magnacca K.N., Martinez N.D., Nielsen R., O'Grady P.M., Percy D., Price D.K., Rabosky D., Shaw K.L., Rominger A.. and Cotoras D. (Poster) Assembly of Arthropod Communities in Hawaii: Can we predict the outcome given a modified dynamic? 21st Hawai'i Conservation Conference. July 16th -18th, 2013. Honolulu, HI. USA
- Cotoras D., Lindberg D. and Gillespie R. (Oral presentation) Phylogeography of the spider *Tetragnatha brevignatha* in the context of an adaptive radiation. Hawai'i Ecosystems Meeting. July 7th and 8th 2013. Hilo, HI. USA
- Cotoras D, Brewer M, Lindberg D. and Gillespie R. (Poster) Another event of independent evolution in color polymorphism on Theridiidae spiders in the Pacific Ocean? 36th Annual meeting American Arachnological Society meeting. July 20th-24th, 2012. Green Bay, WI. USA (*Best poster presentation*)
- Cotoras D., Lindberg D. and Gillespie R. (Oral presentation) Hypothesis about speciation processes on Hawaiian *Tetragnatha* spiders. Hawai'i Ecosystems Meeting. July 2nd and 3rd, 2012. Hilo, HI. USA
- Cotoras D., Lindberg D. and Gillespie R. (Oral presentation) Early stages of the *Tetragnatha* adaptive radiation in the Hawaiian archipelago. Hawai'i Ecosystems Meeting. June 30 – July 1st, 2011. Hilo, HI. USA
- Cotoras D., Casquet J. and Gillespie R. (Poster) Diversification patterns of *Tetragnatha* spiders in remote archipelagos on the Pacific Ocean. Evolution of Life on Pacific Islands and Reefs: Past, present, and future. May 26-30, 2011. Honolulu, HI. USA
- Cotoras D., Cambiazo V., Allende M. (Oral Presentation) Conserved gene networks as a tool to predict gene expression: An example from Animal appendages. Genetics, Development, and Molecular Evolution Supergroup Symposium. May 18. UC Berkeley.
- Cotoras D., Cambiazo V., Allende M. (Poster) Theoretical and experimental comparisons of gene network involved appendage development in arthropods (*Drosophila melanogaster*) and vertebrates (*Danio rerio*). 1st meeting of Chilean Scientists in USA: Nexos Chile-USA 2010. October 29. 2010, Washington D.C., USA
- Razeto-Barry P., Díaz J., Cotoras D., Vásquez R. (Poster) Molecular evolution, mutation size and gene pleiotropy in fluctuating environments: a geometrical model. IV Binational Meeting in Ecology. August 08-13, 2010. Buenos Aires, Argentina.
- Croucher P., Lam A., Cotoras D., Mody N. y Gillespie R.. (Poster) Phylogeography, diversity and colonization history of the Hawaiian happy-face spider. Annual symposium American Genetic Association: "Conservation Genomics". July 26-28, 2010. Hilo, HI. USA.
- Cotoras D., Cambiazo V., Allende M. (Poster) In the footsteps of *Urbilateria.* LI Annual meeting Chilean Society for Biology. November 26-29, 2008. Pucón, Chile
- Cotoras D., Cambiazo V., Allende M. (Poster) Evolution of the gene network involved in appendages development in arthropods (*D. melanogaster*) and vertebrates (*D. rerio*). XXII Annual meeting Chilean Society for Cell Biology. October 5-9, 2008. Pucón, Chile

10.- Thesis (Master's degree)

"In the footsteps of *Urbilateria*, or Evolution of the gene network involved in appendage development in arthropods (*Drosophila melanogaster*) and vertebrates (*Danio rerio*)"

11.- Invited Presentations

- La Selva Biological Station. Organization for Tropical Studies. November, 2014
- California Academy of Sciences. June, 2014
- "Chile Seminars". University of California, Berkeley. May, 2014
- Paleogenomics lab. University of California, Santa Cruz. April, 2014
- University and Japson Herbaria of the University of California, Berkeley. September, 2013
- "Chalk Talk" Integrative Biology dept. University of California, Berkeley. April, 2013
- Parque Nacional Rapa Nui. August, 2012
- Instituto de Entomología - Universidad Metropolitana de Ciencias de la Educación. July, 2012
- Pacific Coast Entomology Society. February, 2012
- Arachnology Dept. California Academy of Science. September, 2011
- Laboratorio de Ecología Evolutiva. Universidad de Chile. August, 2011
- Instituto de Entomología. Universidad Metropolitana de Ciencias de la Educación. August, 2011
- Instituto de Filosofía y Ciencias de la Complejidad. July, 2010

12.- Professional development and Certifications

- B3 Combination Helicopter/Airplane Safety. Interagency aviation training. (May 26$^{th}$, 2013)
- SSI Specialty Course: Enriched Air NITROX. Bamboo Reef School. San Francisco, CA (February 22, 2012)
- SSI Specialty Course: Diver Stress & Rescue. Bamboo Reef School. Monterey, CA (May 08, 2011)
- SSI Specialty Course: Dry Suit Diving. Bamboo Reef School. Monterey, CA (May 08, 2011)
- Wilderness First Aid. Foster Calm–First Aid and Leadership Trainings. Nevada City, CA. (April 02. 2011)
- Cardiopulmonary resuscitation (CPR*)* and automated external defibrillator (AED) for the Community and Workplace. American Safety & Health Institute. Foster Calm–First Aid and Leadership Trainings. Nevada City, CA. (April 02. 2011)
- PADI Advance SCUBA Diver. Seattle SCUBA School. Seattle, WA. (October 03, 2010)
- Webinar Short Course on Modeling Patterns and Dynamics of Species Occurrence. US Geological Survey and Proteus Wildlife Research Consultants. (August 23-27, 2010)

- NAUI Open water Diver. Cal Dive and Travel. Berkeley, CA (June 1st, 2010)
- Exploring genomes with ENSEMBL. Center for Genomics of the Cell (CGC). Santiago, Chile.
  (January 13-14, 2009)
- Workshop on Modeling of Genetic Regulatory and Metabolic Networks. Instituto de Sistemas Complejos de Valparaíso (ISCV). Valparaíso, Chile. (March 27-28, 2008)
- VI Complex Systems Summer School. Instituto de Sistemas Complejos de Valparaíso (ISCV). Valparaíso, Chile. (January 07-11, 2008)

13.- Outreach

- Earth Update planetarium show at Night Life. California Academy of Sciences (July 3rd, 2014)
- "What is a biologist? / Was ist ein biologe?" Arachno-Blog, Senckenberg Museum (April 11th, 2014)
- "Historias de diversificación y extinción en Tepito O Te Henua" Chile Indómito magazine. December 2013. 8: 81-86.
- "Field work after a mass extinction" California Academy of Sciences Blog (June 14th, 2013)
- "Humedal Rano Kau. Manavai de Biodiversidad" Interviewee in documentary from Parque Nacional Rapa Nui
- "Field work during a mass extinction" UCMP Blog (March 26th, 2012)
- "The eternal value of Natural History and the dazzling molecular promise" UCMP Blog. (November 23th, 2011)
- "El eterno valor de la historia natural y la encandilante promesa molecular" RedCiencia.cl (August 4th.2011)
- "Ciencia Básica VS Ciencia Aplicada: Una ilusión basada en la no comprensión" RedCiencia.cl (May 26th.2011)
- "Dinner with a Scientist" Oakland Zoo, CA. (June 3nd, 2010; May 4th, 2011)
- "¿Calentamiento global o cambio global? Conexioniberoamerica.wordpress.com (Feb 26th.2010)

14.- Professional Societies

- British Ecological Society
- European Society for Evolutionary Biology
- American Arachnological Society
- Phi Beta Kappa Academic Honor Society
- Sigma Xi, The Scientific Research Society

15.- Other interests

Skateboarding, Bodyboarding, Snowboarding, Mountain bike, Hiking, Mythology and Tropical ornamental fish

*UPDATED: December 2014*

## Table of contents

Chapter 4: Color polymorphism on *Selkirkiella* spiders (Theridiidae) from the Juan Fernández archipelago and the Valdivian temperate rainforest

# CHAPTER 1

## Using museum specimens to unravel the colonization origin of *Tetragnatha* spiders on Eastern Polynesia (Marquesas, Society Islands and Rapa Nui) and Hawai'i


## ABSTRACT

Colonization of remote archipelagoes corresponds to a very rare event. The Long-jawed spiders, genus *Tetragnatha*, are widely distributed around the world including oceanic islands. In Hawai'i there are 37 species of one or two colonization events from North America. In the Marquesas there are 5 species and phylogenetic evidence suggests an origin somewhere in the Asia. For the species present in the Society Islands, Rapa Nui and *D. raptor* (Hawai'i), their biogeographic origin is still unsolved. In order to examine colonization patterns across the Pacific Rim, we used freshly collected and museum specimens for phylogenetic analysis. The DNA from historic collections was sequenced using newly design primer pairs to amplify overlapping short fragments of COI. The phylogenetic reconstruction shows two of the species from the Society Islands (*T. rava* and *T. tuamoaa*) and *D. raptor* in the same clade as other species from Asia. *T. moua* (Society Islands) is sister to *T. nitens*, which has a circumtropical distribution making strong biogeographic statements difficult. The species from the Marquesas appear linked to a basal polytomy, while the one from Rapa Nui is in the same clade as the Hawaiian species. Other new species from Hawai'i were added into the analysis, but their phylogenetic position within the Web building clade is not conclusive. This study provides preliminary information about origin of colonization events for *Tetragnatha* spiders in the Pacific. The implementation of Next Generation Sequencing technologies and inclusion of more species could improve this understanding.

## INTRODUCTION

Colonization of remote archipelagoes corresponds to a very rare event, which produces new lineages (Losos and Ricklefs, 2009). The combination of the degree of isolation and area of the island results in different evolutionary patterns (Rosindell and Phillimore, 2011). Under conditions of low isolation, constant gene flow will prevent the formation of new species independent of island size. If the island is isolated, but its area small, it will typically produce single island endemics by anagenetic processes. On the other hand, high degrees of isolation and large areas allow for *in situ* diversification (cladogenesis). The most extreme case of this kind of diversification corresponds to adaptive radiations (Schluter, 2000).

The Long-jawed spiders, genus *Tetragnatha*, are widely distributed around the world. As they disperse through ballooning (Okuma and Kisimoto, 1981), most of the different island groups in the Pacific were colonized independently from the mainland (Gillespie, 2002). Even if every archipelago could be considered as a repetition of the same natural experiment, differences in area and isolation produced a wide variety of evolutionary outcomes. For these reasons the *Tetragnatha* spiders provide an ideal system to study evolutionary questions such as ecological shifts associated with isolation (Gillespie, 2004), rates of morphological evolution (Harmon et al, 2010), parallel systems of community assembly (Casquet et al, 2011) or differences in diversification patterns (Gillespie et al., 1997). However to address these questions a solid knowledge of the phylogeny and biogeographic affinities of the group is required.

Hawai'i represents one extreme with 37 described endemic species due to two separate adaptive radiations (Gillespie et al 1994). The Marquesas contain 5 endemic species as the result of a smaller adaptive radiation (Gillespie, 2003a), while the Society Islands have only three island endemics (Gillespie, 2003b). Rapa Nui, on the eastern side of Polynesia, has only one described endemic species (*T. paschae*) (Berland, 1924). On the other extreme are the Austral Islands, where there are no endemics (Claridge et al., 2007). Gillespie (2002) showed that the species present on the first three archipelagoes are not the closest relative to each other, presenting clear evidence of the lack of a stepping stone model of colonization.

Previous studies (Gillespie, 2002; Gillespie et al., 2008, Casquet et al., 2014) also show that the colonization of the Society Islands is different to the one in Hawai'i and the Marquesas in two main aspects. The first one is related with the number of species. Hawai'i as well as the Marquesas have more species than described for the Society Islands. This could be related with the fact that these two archipelagoes are more isolated than Society Islands. Also, the Hawaiian Islands (28, 311 $km^2$) are substantially bigger than the Societies (1, 590 $km^2$).

The second aspect in which they differ is the number of colonization events. In Hawai'i there are 3 clearly identifiable clades: the Spiny leg clade, the Web builder clade and the single species *T. hawaiiensis*, which is a habitat generalist and is present in all the islands. These three clades, even if well defined, cluster together in phylogenetic analysis and they are closely related to species from North America (Gillespie et al., 1994; Gillespie, 2002 and Gillespie et al, 2008). This pattern suggests up to three successful colonization events from a similar species pool or a single event, which diverged very early into at least three well-defined groups (Gillespie et al, 2008).

There is a fourth colonization event of Hawai'i within the Tetragnathidae family, and corresponds to the monotypic *Doryonychus raptor* (Simon, 1900). This species is restricted to the back of valleys on the Na Pali Coast of Kaua'i. Its assignment to a different genus is based on its very distinctive morphology (Simon 1900). *D. raptor* has an elongated claw that it uses to pierce its prey, which is a completely unique behavior in the family (Gillespie, 1991). Due to its restricted distribution in the oldest of the high Hawaiian Islands, it has been considered to be a relictual species of a previous Tetragnathidae colonization (Gillespie et al., 1994).

The Marquesas provide a simpler scenario where the five endemic species cluster as a monophyletic clade (Gillespie at el, 2008, Casquet et al 2014). It is also seen in the larger clade that contains species from Burma and South Africa (Casquet et al. 2014).

In contrast with Hawai'i and the Marquesas, the three species from Society Islands appear as a paraphyletic group. Two of the species (*T. rava* and *T. tuamoaa*) are sisters, while the third one (*T. moua*) is paired with the circumtropical species *T. nitens* (Gillespie at el, 2008, Casquet et al 2014). This suggests that in the Society Islands there is a pattern of diversification that includes a combination of independent colonization events and *in situ* speciation. The actual source of these two colonization events, as well as the one for the Marquesas, are still unclear due to the incomplete sampling of outgroups from the surrounding continental areas. Asia is particularly important as it corresponds to the source of many other groups of organisms, including humans (Gillespie et al., 2008; Bellwood, 2011).

On the other hand, Rapa Nui (Easter Island) presents a different situation in comparison to Hawai'i, Marquesas and Society Islands. It is a small island (163.3 $km^2$) and one of the most isolated locals on Earth. It also has been strongly impacted by human occupation (Polynesian, European and Chilean), which caused the alteration of most of the native environments (Mann et al., 2008). In 1924, Lucien Berland described the only endemic spider species from the island, *Tetragnatha paschae*. However, in subsequent years it has been considered possibly extinct (Baert et al, 1997) or as a synonym of other wide distributed species such as *T. mandibulata* (Gillespie *pers. comm.*).

The origin of *T. paschae* is potentially different from the Marquesan and Society species. Easter Island is 4,251 km from Tahiti and 3,526 km from the Chilean coast. As there is similar distance from Hawai'i to North America, it is conceivable that a portion of the Rapa Nui biota could come from South America. Indeed the extinct palm tree (*Paschalococos disperta*), which once covered most of the island, has been related to the Chilean palm (*Juvaea chilensis*) instead of other Indo-Pacific taxa (Dansfield et al., 1984).

In summary, for the Hawaiian species their origin has been established in North America (Gillespie et al., 1994). For the Marquesas there is some evidence that links them with Burma and South Africa (Casquet et al., 2014). For the Societies and Rapa Nui the source of colonization is still not clear. Then a key question to address before any comparative study of species among these islands is, what is the colonization source for the *Tetragnatha* from South Pacific Islands (Marquesas, Society Islands and Rapa Nui)? This answer will provide a more complete phylogenetic framework to ask comparative questions across volcanic archipelagoes with comparable geologic histories but with different patterns of colonization and modes of speciation.

Based on the biogeographic affinities of other groups, we hypothesize that the

Marquesas and the Society Islands were colonized from Southeast Asia, Different origins are not unexpected for the two groups present in Society Islands. For the species present in Rapa Nui, we expect a South American origin based on the distance from the continent and the fact that the mainland is an important source of propagules.

As part of the analysis we also included *D. raptor* in order to test whether it is effectively a unique genus or is nested within *Tetragnatha*. We expect to find it as a sister or nested within the other Hawaiian *Tetragnatha*, which have a North American origin (Gillespie et al., 1994).

In order to examine colonization patterns across the Pacific Rim, museum specimens provide a means for obtaining the necessary taxa throughout such a broad geographic area. However, not all museum specimens have been preserved for genomic work (Wandeler et al., 2007). Krehenwinkel and Tautz (2013) have successfully amplified DNA sequences from historical collections of the wasp spider *Argiope bruennichi.* For the mitochondrial gene Cytochrome oxidase subunit I (COI) they used a combination of primer pairs for short (135 bp) and long fragment (1,200 bp) amplification. To obtain polymorphic nuclear markers they did a single lane of 454-shotgun sequencing. These data allowed them to show a high genetic diversity on expanding populations and admixture with previously isolated lineages.

To include museum specimens into the current COI data set for Pacific *Tetragnatha*, we decided to take the short fragment amplification approach using Sanger sequencing to assemble long contigs for the phylogenetic reconstruction.

The goal of this paper is to take advantage of museum and field collected specimens to provide information about the colonization origin of *Tetragnatha* spiders on Eastern Polynesia (Marquesas, Society Islands and Rapa Nui) and Hawai'i (i.e. *D. raptor*).

## MATERIALS AND METHODS

### Collections

The taxa included on this study correspond to *Tetragnatha* species present worldwide with a particular emphasis on the Pacific Rim taxa. A total of 50 OTUs were downloaded from GenBank (Table 1), while another 33 correspond to museum or field collected specimens (Table 2). A total of 7 specimens were collected in the Hawaiian Islands and one in Rapa Nui. As part of the ingroup we also included *Doryonichus raptor*. Museum samples were obtained from the California Academy of Sciences (12), the Senckenberg Museum (6) and the Essig Museum of Entomology (7). Undescribed species were included with a provisional name in quotes, preceded by the abbreviation 'nsp'. These names are not valid and a proper description will be provided at a future date. Originally a total of 25 specimens from the Senckenberg Museums were tested in order to amplify DNA, but only 6 of them were finally included on the phylogeny (for original specimens Table 3).

**DNA sequences**

**i) Fresh Specimens**

DNA was extracted form 4 legs from the right side of each the spider. The mitochondrial gene COI was sequenced because it is well represented in species in the wider taxonomic range of interest. For the specimens preserved for molecular work a 648 bp region was amplified using the primers LCO and HCO (Folmer et al, 1994). The PCR reaction mix consisted of: 1 ul of each primer at 10 uM, 2 ul of AmpliTaq (Life Technologies) buffer 10x, 0,5 ul of $MgCl_2$ 25 mM, 11.7 ul $H_2O$, 1 ul BSA 1x, 0.2 ul DreamTaq (Thermo Scientific) and 1 ul of DNA extraction. The amplification profile started with 2 minutes at 95°C, followed by 35 repetitions of a cycle which started with 30 seconds at 95°C, then 45 seconds at 42°C and finally one minute at 72°C; there was an extra step of extension at 72°C for ten minutes. The PCR products were verified on an agarose/TBE gel. PCR products were cleaned using Axygen AxyPrep MagPCR Clean-up kit. DNA was sequenced directly in both directions through the cycle sequencing method using dye terminators (Sanger et al., 1977) using Life Technologies' BigDye Terminator v3.1 Cycle Sequencing Kit. Sequenced products were cleaned using Axygen AxuPrep Mag DyeClean kit and run out on ABI 3730XL DNA Analyzer automated sequencer. The sequences were edited and align with Geneious Pro 5.6.7 (Biomatters). The sequence alignment used the algorithm MAFFT (Katoh et al., 2002) with the default parameters.

**ii) Museum Specimens**

The museum specimens were obtained from the California Academy of Sciences (San Francisco, USA) and the Senckenberg Museum (Frankfurt, Germany). The specimens from the California Academy of Sciences were preserved in conditions for DNA work, so the same protocol as for the fresh specimens was utilized.

Most of the specimens from the Senckenberg Museum were preserved at room temperature in 70% ethanol. The DNA was extracted from 4 legs using the salt precipitation protocol from 5PRIME. As it does not require a centrifugation step, it is a more gentle method to obtain the already degraded DNA. We use glycogen as a DNA carrier (Krehenwinkel and Tautz, 2013) following the instructions of the manufacturer. We also designed and optimize six pairs of primers to amplify fragments between 200 bp and 300 bp.

**Primer Design**

The primers were designed using the web-based program Primer 3 (Koresaar and Remm, 2007; Untergrasser et al, 2012) using a consensus sequence generated mainly from Asia, the South Pacific and Europe specimens as a reference. In the case of ambiguities we selected the most common base. In the cases A and T were equally common we selected T (Ewing, *pers. comm.*).

The primer sequences are in the Table 4. The amplification profile started with 2 minutes at 95°C, followed by 24 repetitions of a cycle which started with 30 seconds at 95°C, then 45 seconds at the respective annealing temperature (Table 4) and finally one

minute at 72°C; there was an extra step of extension at 72°C for ten minutes. Once the PCR product was obtained the same protocol as for the fresh samples was applied. The primers where removed by hand using Geneious Pro 5.6.7 (Biomatters). The efficacy of the primers was tested with two Hawaiian species (*T. brevignatha* and *T. anuenue*) for which we had the sequence of COI amplified using HCO:LCO.

**Phylogenetic analysis**

Model parameters were estimated using PARTITIONFINDER (Lanfear et al., 2012). The sequence was partitioned by codon position and the favored models were: position 1: SYM+I+G; position 2: GTR+I+G; and position 3: HKY+G. As outgroup we used a clade with three *Pachygnatha* species. Previous studies of the whole Tetragnathidae family show that this is taxon is the sister genus of *Tetragnatha* (Álvarez-Padilla et al., 2009). The phylogenetic reconstruction was done using Mr. Bayes 3.2.2 software (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003) available on the remote server CIPRES Science Gateway version 3.3 (Miller et al., 2012). The phylogenetic reconstruction utilized 2 runs of 4 independent chains each. It was run for 275,000,000 generations with a sampling frequency of 1,000. The burn-in was set at 25%. Parameter convergence was determinate by checking the potential scale reduction factor (PSRF=1.0), standard deviation of split frequencies (<0.01) and estimating Effective Sample Size (ESS>200) by combining both runs using TRACER version 1.7.5. The phylogenetic tree was visualized on FigTree version 1.4.0 (http://tree.bio.ed.ac.uk/software/figtree/).

**RESULTS**

**Identification of *Tetragnatha* species from Rapa Nui**

During the fieldwork in Rapa Nui (August, 2012) specimens from the genus *Tetragnatha* were collected from the Rano Raraku crater (Fig. 1). The original description from Berland (1924) does not match the specimens collected, and based on cheliceral morphology, it is not *T. nitens*, which has been previously reported on the island (Baert et al., 1993). The morphology is also distinct from other South Pacific species (Gillespie *pers. comm.*) and continental species from Chile (Nicolet, 1849). The species with the closest cheliceral morphology corresponds to *T. similis*, but the orientation of the cheliceral teeth is different between the two species (Nicolet, 1949; Keyserling E. 1865).

**Primer design for small fragments**

The sequences for the six primer pairs used for short fragment amplification are given in Table 4. The assembled amplification product for *T. brevignatha* consisted of 414 bp with 97% of identical sites with respect to the sequence obtained with the universal primers from the same specimen. For *T. anuenue* the assembled amplification product was 459 bp with 98.2% of identity. These results validate the use of these primers for our historical samples.

**DNA amplification of specimens from Senckenberg Museum**

Our original target was a region of 813 bp of COI, using an assembly of six small overlapping fragments (length of each fragment in Table 4). The original set of samples was 25 specimens, for 9 of them it was possible to amplify COI with the regular primers (HCO:LCO), but the rest required to use of primer pairs to amplify small fragments. Table 5 shows that the amplification was variable depending on the fragment. As a general pattern Fragment 1 and 6 were the more commonly amplified, followed by fragments 4 and 5. Except for SM29 all the other regular primer samples were successful, and the other primer pairs were not utilized. Based on the quality of the sequences amplified and availability of sequences from the same species in GeneBank, we were able to include the 6 sequences from the Senckenberg Museum in the phylogenetic analysis.

**Phylogenetic reconstruction**

After running the phylogenetic analysis all parameters had a potential scale reduction factor equal to 1.0 and the Effective Sample Size for the combined runs was greater then 200. The standard deviation of split frequencies between runs was 0.015034. The authors of MrBayes consider a value <0.01 "very good" and values between 0.01 and 0.05 "adequate" (Ronquist et al., 2011). In order to take into account saturation on third positions (Garb et al., 2004), we did a per position partition of the COI sequence and tested for different models for molecular evolution at each site (see Methods).

The phylogenetic reconstruction (Fig. 2) recovers as a monophyletic clade of the *Tetragnatha* species and the specimens of *D. raptor* (posterior probability = 1). *D. raptor* is nested on a clade with species from Laos, Papua New Guinea, *T. nitens* (circumtropical), *T. valida* (eastern Australia and Tasmania), Madagascar and the species from Society Islands (World Spider Catalog, 2014). This pattern suggests a possible Asian origin for *D. raptor*.

A similar scenario can be argued for *T. nitens*, which even with its circumtropical distribution, it clusters with Asian species. Other studies also show that specimens from *T. nitens* from very different regions of the world always cluster as a monophyletic group (Gillespie et al., 2008; Casquet et al., 2014). This widespread species also likely originated in the Asian region.

This conclusion it is also extendible to the three species from Society Island. As in previous studies, *T. rava* and *T. tuamoaa* appear as sisters, while *T. moua* is on a different branch, but closely related to *T. nitens* (Gillespie et al., 2008). A species from Laos is sister to the pair *T. nitens/T. moua* and at the base of this group is a polytomy. Note that *T. nitens* has a wide distribution with populations across all the tropics, so even if as a species it clusters with Asian species, the population that originated *T. moua* could be from other region.

The four species from the Marquesas included in the analysis cluster form a monophyletic clade with good support, but it is positioned at the base of the ingroup polytomy. Therefore, it is not possible to resolve a biogeographic origin for the archipelago's fauna. A similar pattern is presented in some other Asia species.

Another clade, with less support (posterior probability = 0.6661), includes species from all over the world. The support values among the pairs of species are high and the divergences appear to be deep. This suggests that this group has mostly very distantly related species with numerous intermediate terminals missing. The lack geographical resolution also suggests this.

A clade that contains all of the Hawaiian species is present in this group and also has a polytomy at its base. The Web Building clade (in green) is not recovered as monophyletic. There is one clade that includes *T. perkinsi, T. filiciphilia, T. limu, T. stelarobusta, T. eurychasma* and *T. trituberculata*. On the other hand, *T. acuta* and *T. nsp. "eury-like"* are sisters, but instead are directly connected to the basal polytomy. Also directly connected to the basal polytomy are *T. nsp. "emerald ovoid", T. palludicola, T. nsp. "red white"* and *T. nsp. "line dots"*. The relative relationships between the species on the Web Building clade are partially congruent with previous publications (Blackledge and Gillespie, 2004), the much lower resolution here is likely due to the small fragment size of COI only (compared to 1,481 bases across COI, 16S and 12S in the previous work). We also included more species (*T. palludicola*, *T. trituberculata, Tetragnatha "line dots"* and *Tetragnatha "red white"*) from this group in the present study.

The major groupings within the spiny leg clade are recovered as in previous publications (Gillespie, 2004). This clade appears as sister to the Asian species (Sri Lanka to China) *T. geniculate.* Outside of this pair are four species from the New World. Three from North America (*T, laboriosa, T. shoshone* and *T. pallescens*) and one from Chile. Another species present in this clade corresponds to *T. montana*, which has a broad Palearctic distribution.

Interestingly, the species *T. nsp. "shirt tail"* corresponds to a sister species to *T. hawaiiensis*. This is the first molecular phylogeny in which a morphologically distinct taxon appears to be sister to this widespread species. However a taxonomic revision is needed for *T. hawaiiensis* (Gillespie *pers. comm.*), and it will likely be synonymized with *T. olindana* (Karsch, 1880). As part of the Hawaiian species group, but linked directly to the basal polytomy, is the specimen collected from Rapa Nui, suggesting a New World origin for this species as well.

## DISCUSSION

### Short fragment amplification

The present approach to design and optimize primers for short fragment amplification of historic DNA using Sanger sequencing was successful. But, due to the variable preservation conditions of the museum samples the amplification results were variable. A consistent pattern was the almost absent amplification of the area covered by the primer pairs number 2 and 3. As Fig. 3 shows that area corresponds to a high GC content region of the gene.

Krehenwinkel and Tautz (2013) amplified a region of 1,200 bp. The fragment (648 bp) used here is included within their region. Their sequences also cover the GC-rich region, but it is not as pronounced as the one in *Tetragnatha*. This could have contributed to our amplification problem or it is possible that their primers did not have an annealing site on this area. Interestingly the first gap on most of their specimens is

approximately 20 bp from the region with the highest GC content. In their approach four primer pairs produced amplification products of approximate 350 bp. Also, they had another combination that produced a long fragment (approx. 1,300 bp) and one that amplified a short one (approx. 100 bp).

Theoretical and empirical studies have shown that high GC content areas require short annealing times (3 to 6 seconds) in order to prevent the formation of ternary complex at erroneous binding sites (Mammedov et al., 2008). In early experiments different master mix combinations were tried included betaine to resolve secondary structures usually formed on GC-rich areas, however we did not vary annealing times.

Another difference between our work and Krehenwinkel and Tautz (2013) is that we did not have contemporary specimens to compare results with, which complicated primer development because of the lack of a reference COI sequence. Therefore, we had to create a consensus sequence from available "old world species" and then optimize the PCR conditions for each fragment on each specimen.

**Colonization origin of *D. raptor* and *Tetragnatha* species from Marquesas and Society Islands**

The presence of *D. raptor* and *Tetragnatha* from Society Islands in the same clade as other species from Southeast Asia and the south Pacific might suggest an origin from this region of the world. However, due to the lack of support at the base of this clade it is not possible further resolve a biogeographic origin within this region.

This result is somewhat unexpected because *D. raptor* is a Hawaiian species. The fact that it does not aggregate with any island species or with New World species supports the possibility of an Asian origin. Thus, Hawai'i may have been colonized from both sides of the Pacific Ocean by Tetragnathidae spiders. While the number of species left from both colonizations is substantially different, it is difficult to make a case for differences in evolvability (Wagner and Altenberg, 1996) due to the lack of a geographic origin point. Their restricted distribution and presence on the oldest of the high islands, suggest that *D. raptor* is a relictual species of what could have been an ancient radiation that existed when some of the Northwestern Hawaiian islands had high elevation environments. More work is required to test this and this will required a broader taxonomic sampling, more nuclear markers and explicit models of diversification that take into account extinct species in molecular phylogenies. The last issue seems to be highly complex and object of debate (Morlon et al., 2010; Rabosky, 2010).

The suggested Asian origin for *Tetragnatha* from Society Islands follows an already well-documented trend present in other organisms (Gillespie et al., 2008). Likewise, the presence of *T. moua* as a sister to the widely distributed *T. nitens* has already been demonstrated (Gillespie et al. 2008, Casquet et al. 2014) and suggests the possibility of peripatric speciation produced by an isolated ancient population from Tahiti. However, although *T. nitens* clusters with Asian species, better sampling and more markers are required to distinguish which regional population gave rise to *T. moua*. Again as previously shown (Gillespie et al., 2008), *T. rava* and *T. tuamoaa* appear as sister species, but the fact that they are part of the polytomy makes it impossible to infer a more detailed biogeographic origin.

As previously demonstrated, the Marquesan species clustered as a monophyletic

group (Gillespie et al. 2008; Casquet et al. 2014), but their origin is still unresolved. Unlike Casquet et al. (2014) findings the samples from Burma and South Africa did not appear as sister to the Marquesan clade. Moreover, the large distances of the Marquesan species to any other outgroup suggest colonization from a site that has not yet been sampled.

Other species from Pacific Islands also present the same pattern of a likely Asian ancestry. For example, *T. macilienta* (Norfolk Is. to Society Is.) and *T. spp. "Smk8"* from Palau are sisters to a *Tetragnatha* species from Brunei. Another species from Palau, *T. spp. "LongBr"*, is present in a clade with *T. spp. "Pointy abdomen"* from Burma and *T. spp. "3Lng"* from South Africa. Note that the lack of a proper identification makes it impossible to rule out the possibility that these three specimens correspond to the same species with a wide distribution. Nonetheless, this group is sister to *T. javana*, which is distributed from Africa to Japan, including the Philippines and Indonesia (World Spider Catalog, 2014). Finally, *T. spp.* from New Zealand is sister to a specimen collected in the Enga Province of Papua New Guinea.

**Species identity and origin of *Tetragnatha* from Rapa Nui**

An unequivocal identification of the specimen from Rapa Nui was not possible. Due to the level of disturbance of Easter Island it is possible that this spider is the result of human introduction, but this hypothesis is hard to test given the lack of knowledge of the Chilean *Tetragnatha* fauna (only one extensive taxonomic publication Nicolet, 1849). Associated with this, the lack of a comprehensive sampling from Chile makes it impossible to have a reference collection or sequences to compare. Our study includes one sequence from a specimen collected in the X Region of Chile, but the Rapa Nui specimen is not closer to it than to any of other North American species. What is possible to say is that it is not closely related to any of the other South Pacific species, including the ones from Palau, Marquesas, Societies, New Zealand or the Pacific widespread *T. macilienta*. Also, it does not correspond to *T. mandibulata*.

The specimen from Easter Island could be a human introduction from the west coast of South America or a natural colonization from the same region. If it corresponds to a human introduction we would expect a very short branch length with respect to other species from that region. In the case it is a natural colonization the number of mutations should be higher than the ones found on a human introduction and they will be proportional to the time since colonization. Better collecting form Chile, Perú and Ecuador is need before this question can be addressed.

If the Rapa Nui species is a natural colonization it will be a very important finding as the island has only other 2 endemic species of insects living in the surface (*Lipsana insularispaschalis* (Diptera) and *Bidessus skottsnergi* (Coleoptera: Dytisidae)). Recently 8 new endemic and 2 natives species of arthropods were described from cave systems (Wynne et al., 2014).

**Phylogenetic position of new Hawaiian species**

The general phylogenetic structure of the Spiny leg clade is very similar to previous publications (Gillespie et al, 1994; Gillespie, 2004; Gillespie et al., 2008). In the

case of the Web Building clade there was a very poor resolution and some of the species are not members of the same clade. In our analysis we include more species (*T. palludicola*, *T. trituberculata, Tetragnatha "line dots"* and *Tetragnatha "red white"*) that the previous studies have not included (Blackledge and Gillespie, 2004). However, the relative relationships between species is suspect using only a single mitochondrial marker and there is also a likely effect of under sampling. However, these preliminary data suggest that *T. nsp. "shirt tail"* is a sister to *T. hawaiensis*. This should be taxonomically review more closely taking in consideration the identity of *T. olindana.*

In terms of affinities with other parts of the world, the single specimen of *T. geniculata* from Laos appears to be sister to the spiny leg clade. However, more specimens and data, in particular from nuclear markers are essential to further test this relationship. Interestingly this is not the first time that a species from a very remote location appears as a sister to the Hawaiian *Tetragnatha.* A recent publication (Casquet et al. 2014) showed the same situation with a species from Réunion Island (Lineage C). Like other members of the Spiny leg clade, this species does not spin a web.

*T. montana* also appears in the same clade as the Hawaiian species. That is very unexpected as the North American species has a Palearctic distribution. However, we think that this is an artifact of the COI marker. In other analyses this species appears connected to the basal polytomy of the larger Hawaiian group.

There are also other American species that do not appear to be sister taxa to the Hawaiian ones. They do not form a monophyletic clade, which suggest a more complex history of diversification. The lack of more appropriate markers and under sampling makes it difficult to further resolve the biogeographic patterns of this group. Three North American species cluster together (*T. versicolor, T. elongata* and *T. viridis*), while other two are found in another region of the tree (*T. straminea* and *T. extensa*). The two species with distribution in Central (*T. tenuis*) and South America (*T. guatemalensis*) are not reciprocally monophyletic or closely related. This part of the American continent is extremely under sampled and given its latitudinal gradient, area and habitats we suspect it will ultimately be shown to have a very high diversity of spiders.

## CONCLUSION

The present work provides preliminary information on patterns of colonization of Pacific Islands with particular emphasis on Eastern Polynesia (Marquesas, Society Islands and Rapa Nui) and Hawai'i. It supports with previous studies (Gillespie, 2002, Gillespie et al 2008, Casquet et al 2014) that concluded a lack of support for a stepping stone model of colonization, and rather a pattern of direct colonizations from multiple sources areas. The inclusion of outgroups, from the Americas, Asia, and the South Pacific, broadly agrees with previous results on the sources of colonization for different archipelagoes and islands. However, the limitation to a single small fragment of mitochondrial DNA restricts further assessment of regions of origin.

The fact that all the Hawaiian *Tetragnatha* have a North American origin agrees with previous studies (Gillespie et al., 1994). However, the relationships of *D. raptor* and the *Tetragnatha* from Society Islands with Asian species is a new result. This last archipelago appears colonized by two independent events: *T. rava* and *T. tuamoaa* probably from Asia, while *T. moua* presents a more complicated situation, as it is sister to

*T. nitens*, which has a circumtropical distribution. Other species from the South Pacific show a similar pattern with Asian relationships. The endemics from Marquesas Islands form a well-defined monophyletic group, which does not cluster with any of the taxa included in the analysis. Finally, the specimen collected on Rapa Nui appears to be a member of the same clade as Hawaiian and other New World species. The difficulty of identification and lack of a comprehensive sampling from the west coast of South America make it currently impossible to distinguish between a human introduction and a natural colonization event for this species.

These data presents a pattern where species from both coasts of the Pacific Rim have contributed to the colonization of remote Pacific islands and these events have produced radically different evolutionary outcomes, including single island endemics and large adaptive radiations. Further taxonomic sampling including more species form Asia and Central and South America is needed to resolve at finer scales sources populations. The colonization of *Tetragnatha* spiders on Pacific islands not only provides a system for biogeographic questions, but also to test models of diversification in a well-studied geographic context. With additional geographic samples and the implementation of Next Generation Sequencing technologies in should be possible obtain in a cost effective way thousands of independent markers and examine these patterns.

## ACKNOWLEDGEMENTS

The authors would like to acknowledge a big number of people and institutions whom collaborate at different stages of this research. The fieldwork in Hawai'i was supported by Laura Arnold, Timothy Bailey (HALE), David Benítez (HAVO), Katie Champlin (Limahuli Botanical garden), James Friday (UH Mānoa), Emory Griffin-Noyes (Limahuli Botanical Garden), Faith Inman-Narahari (UCLA), Darcey Iwashita (UH Mānoa), Raina Kaholoaa (HALE), Susan Kennedy (UC Berkeley), Jessie Knowlton (Michigan Tech), Rick Lapoint (U Arizona), Scott Laursen (UH Mānoa), Karl Magnacca (O'ahu Army Natural Resources Program), Elizabeth Morrill (SFSU), Patrick O'Grady (UC Berkeley), Rita Pregana (HAVO), Donald Price (UH Hilo), David Rankin (U Vermont), William Roderick (Stanford), Andrew Rominger (UC Berkeley), Karen Uy (UH Hilo), Erin Wilson (UC Riverside) and the Kīpuka team.

The permit processing and access to different reserves and private land was possible thanks to Steve Bergfeld (DOFAW Big Island), Pat Bily (TNC) Maui, Tabetha Block (HETF), Shalan Crysdale (TNC Big Island), Lance DaSilva (DOFAW Maui), Dean Danae (Kahoma Ranch), Charmian Dang (NAR), Melissa Dean (HETF), Betsy Gagne (NAR), Elizabeth Gordon (HALE), Lisa Hadway (DOFAW Big Island), Paula Hartzell (Lana'i Resorts, LLC), Greg Hendrickson (Kealakekua Ranch), Mel Johansen (TNC Big Island), Pomaikaʻi Kaniaupio-Crozier (Maui Land and Pinneapple), Cynthia King (DLNR), Peter Landon (NAR Maui), Rhonda Loh (HAVO), Russell Kallstrom (TNC Moloka'i), Joey Mello (DOFAW Big Island), Ed Misaki (TNC Moloka'i), Elliot Parsons (Pu'u Wa'awa'a HETF), Lani Petrie (Kapapala Ranch), Shawn Saito (Parker Ranch), Joe Ward (Maui Land and Pinneapple) and Kawika Winter (Limahuli Botanical Garden).

The fieldwork in Rapa Nui was supported by Luis Flores (Instituto de Entomología, UMCE), Cristian Villagra (Instituto de Entomología, UMCE), personal of Parque Nacional Rapa Nui (CONAF); in particular Graciela Campbell, Ninoshka

Cuadros, Anthony Dubois, Omar Durán Veriveri, Pedro Hito, Pedro Lazo, Sergio Manuheuroroa, Raúl Palomino Matta, Hotu Paté, Carlos Salinas, Enrique Tucki and John Tuki; and Armada de Chile. The permit processing was supported by Lynne Hollyer (UC Berkeley Industry Alliances Office).

The work at the Senckenberg Museum was facilitated by curator Peter Jeager and Julia Altman; while at the California Academy of Sciences by curator Charles Griswold.

We also appreciate the advices on working with museum DNA of Henrik Krehenwinkel (Max Plank Evolutionary Biology) and help with the phylogenetic analysis of Michael Brewer (East Carolina University) and Pete Croucher (Illumina). Juliane Casquet (Muséum national d'Histoire naturelle. France) was involved in early discussions on this topic. Fernándo Álvared-Padilla (Universidad Nacional de México) collaborate with the identification of some of the museum specimens. Gerge Roderick (UC Berkeley) and Charles Marshall (UC Berkeley) also contributed with constructive commentds on earlier versions of the manuscript.

Darko Cotoras' PhD program has been founded by a Fulbright/CONICYT fellowship and a researcher positon on a NSF Dimensions of Biodiversity Project. The fieldwork on Hawai'i and the visit to the Senckenberg Museum was funded by Graduate Research Allocation Committee (Integrative Biology dept. UC Berkeley), Summer Research Grant ((Integrative Biology dept. UC Berkeley), Walker Grant (Essig Museum of Entomology), Graduate Division Travel Voucher Award, Sigma Xi grant and Research Grant Graduate Division UC Berkeley. The fieldwork on Rapa Nui was funded by the Tinker Grant (Center for Latin American Studies at UC Berkeley) and the Proyecto FONDECYT de Iniciación (Cristian Villagra and Luis Flores).

## LITERATURE CITED

Álvarez-Padilla F, Dimitrov D, Giribet G, Hormiga G. 2009. Phylogenetic relationships of the spider family Tet- ragnathidae (Araneae, Araneoidea) based on morphological and DNA sequence data. Cladistics 25: 109–146

Baert, L., Lehtinen, P. & Desender, K., 1997. The spiders (Araneae) of Rapa Nui (Easter Island). Bulletin van het Koninklijk Belgisch Instituut voor Natuurwetenschappen, Entomologie, 67: 9-32

Baert, L., Lehtinen, P. & Desender, K., 1997. The spiders (Araneae) of Rapa Nui (Easter Island). Bulletin van het Koninklijk Belgisch Instituut voor Natuurwetenschappen, Entomologie, 67: 9-32

Bellwood P. 2011. Holocene Population History in the Pacific Region as a Model for Worldwide Food Producer Dispersals. Current Anthropology. 52: S363-S378

Berland L. 1924. Araignées de l'île de Pâques et des îles Juan Fernandez. The Natural History of Juan Fernandez and Easter Island. 3:. 419-437

Blackledge T.A. and Gillespie R.G. 2004. Convergent evolution of behavior in an adaptive radiation of Hawaiian web-building spiders. Proceedings of the National Academy of Sciences USA 101: 16228–16233

Casquet J., Bourgeois Y.X.C., Cruaud C., Gavory F., Gillespie R.G. and Thébaud C. 2014. Community assembly on remote islands: a comparison of Hawaiian and Mascarene spiders . Journal of Biogeography (early view)

Casquet J., Thébaud C. and Gillespie R. 2010. How to build an insular community? Unrevealing neutral and deterministic processes using spider communities. Proceedings of the 11th Pacific Science Inter-Congress – ISBN N°978-2-11-098964-2

Claridge E.M., Garb J.E., Gillespie R.G. and Percy D.M. 2007. Insects and spiders of the Austral Islands. *In* J.Y.Meyer, P. Bouchet, & A. Allison, eds, The terrestrial biodiversity of the Austral Islands (French Polynesia), Museum national de Histoire naturelle, Paris & Bishop Museum Press, Honolulu, HI.

Dransfield J., Flenley J.R., King S.M., Harkness D.D. and Rapu S. 1994. A recently extinct palm from Easter Island. Nature 312: 750 - 752

Drummond AJ, Ashton B, Buxton S, Cheung M, Cooper A, Duran C, Field M, Heled J, Kearse M, Markowitz S, Moir R, Stones-Havas S, Sturrock S, Thierer T, Wilson A (2012) Geneious v5.6, Available from http://www.geneious.com

Folmer O., Black M., Hoeh W., Lutz R. and Vrijenhoek R.1994. DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. Molecular Marine Biology and Biotechnology. 3: 294-299

Garb J.E., González A. and Gillespie R.G. 2004. The black widow spider genus *Latrodectus* (Araneae: Theridiidae): phylogeny, biogeography, and invasion history. Molecular Phylogenetics and Evolution 31: 1127–1142

Gillespie R.G. 1991. Predation through impalement of prey: The foraging behavior of *Doryonychus raptor* (Araneae: Tetragnathidae). Psyche 98: 337-350

Gillespie R.G. 2003a. Marquesan spiders of the genus *Tetragnatha* (Araneae, Tetragnathidae). The Journal of Arachnology 31: 62–77

Gillespie R.G. 2003b. Spiders of the Genus *Tetragnatha* (Araneae, Tetragnathidae) in the Society Islands. The Journal of Arachnology 31:157–172

Gillespie R.G. 2004. Community assembly through adaptive radiation in Hawaiian spiders. Science. 303: 356-359

Gillespie R.G., Claridge E.M. and Goodacre S.L. 2008. Biogeography of the fauna of French Polynesia: diversification within and between a series of hot spot archipelagos. Phil. Trans. R. Soc. B 363: 3335-3346

Gillespie R.G., Croom H.B. and Hast G.L. 1997. Phylogenetic relationships and adaptive shifts among major clades of *Tetragnatha* spiders (Araneae: Tetragnathidae) in Hawai'i. Pacific Science 51: 380-394

Gillespie R.G., Croom H.B. and Palumbi S.R. 1994. Multiple origins of a spider adiation in Hawaii. Proc. Nati. Acad. Sci. USA 91: 2290-2294

Huelsenbeck J.P. and Ronquist F. 2001. MRBAYES: Bayesian inference of phylogeny. Bioinformatics 17: 754-755

Karsch F. 1880. Mittheilung uber die von Herrn Dr O. Finsch während seiner polynesischen Reise gesammelten Myriopoden und Arachniden. Sitzungsberichte der Gesellschaft Naturforschender Freunde zu Berlin 1880: 77-83.

Katoh K., Misawa K., Kuma K. and Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 30: 3059-3066

Keyserling E. 1865. Beiträge zur Kenntniss der Orbitelae Latr. *Verhandlungen de*r Kaiserlich-Königlichen Zoologisch-Botanischen Gesellschaft in Wien 15: 799-856

Koressaar T. and Remm M. 2007. Enhancements and modifications of primer design program Primer3. Bioinformatics 23:1289-1291

Koressaar T. and Remm M. 2007. Enhancements and modifications of primer design program Primer3 Bioinformatics 23: 1289-1291

Krehenwinkel H. and Tautz D. 2013. Northern range expansion of European populations of the wasp spider *Argiope bruennichi* is associated with global warming–correlated genetic admixture and population-specific temperature adaptations. Molecular Ecology 22: 2232–2248

Lanfear, R., Calcott, B., Ho, S.Y. & Guindon, S. 2012. Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Molecular Biology and Evolution. 29: 1695–1701

Losos J.B. and Ricklefs R.E. 2009. Adaptation and diversification on islands. Nature 457: 830-836

Mammedov T.G., Pienaar E., Whitney S.E., TerMaat J.R., Carvill G., Goliath R., Subramanian A. and Viljoen H.J. 2008. A Fundamental Study of the PCR Amplification of GC-Rich DNA Templates. Comput Biol Chem. 32: 452–457

Mann D., Edwards J., Chase J., Beck W., Reanier R., Mass M., Finney B. and Loret J. 2008. Drought, vegetation change, and human history on Rapa Nui (Isla de Pascua, Easter Island). Quaternary Research 69: 16–28

Miller M.A., Pfeiffer W. and Schwartz T. 2012. "The CIPRES Science Gateway: Enabling High-impact Science For Phylogenetics Researchers With Limited Resources", XSEDE '12, Proceedings of the 1st Conference of the Extreme Science and Engineering Discovery Environment: Bridging from the eXtreme to the campus and beyond: pg: -, ACM, 978-1-4503-1602-6

Morlon H., Potts M. D. and Plotkin J. B. 2010. Inferring the dynamics of diversification: A Coalescent approach. PLoS Biology 8: e1000493

Nicolet A.C. 1849. Aracnidos. In: Gay, C. (ed.) Historia física y política de Chile. Zoología 3, 319-543

Okuma C. and Kisimoto R. 1981. Airborne spiders collected over the East China Sea. (in Japanese). Japanese Journal of Applied Entomology and Zoology. 25: 296–298

Rabosky D. L. 2010. Extinction rates should not be estimated from molecular phylogenies. Evolution 64: 1816–1824

Ronquist F. and Huelsenbeck J.P. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19: 1572-1574

Ronquist F., Huelsenbeck J. and Teslenko M. Draft MrBayes version 3.2 Manual: Tutorials and Model Summaries. November 15, 2011

Rosindell J. and Phillimore A.B. 2011. A unified model of island biogeography sheds light on the zone of radiation. Ecology Letters 14: 552–560

Sanger F., Nicklen S. and Coulson A.R. 1977. DNA sequencing with chain-terminating inhibitors. Proceedings of the National Academy of Sciences USA 74: 5463-5467

Schluter D. 2000. The Ecology of Adaptive Radiation. Oxford University Press, Oxford.

Simon E. 1900. Arachnida. In: Fauna Hawaiiensis, or the zoology of the Sandwich Isles: being results of the explorations instituted by the Royal Society of London promoting natural knowledge and the British Association for the Advancement of Science. London 2, 443-519

Untergrasser A., Cutcutache I., Koressaar T., Ye J., Faircloth B.C., Remm M. and Rozen S.G. 2012. Primer3 - new capabilities and interfaces. Nucleic Acids Research 40: e115

Wagner G.P. and Altenberg L. 1996. Perspective: complex adaptations and the evolution of evolvability. Evolution 50: 967–976

Wandeler P., Hoeck P.E.A. and Keller L.F. 2007. Back to the future: museum specimens in population genetics. TREE. 22: 634-642

World Spider Catalog (2014). World Spider Catalog. Natural History Museum Bern, online at http://wsc.nmbe.ch, version 15.5, access on August 30, 2014.

Wynne J.J., Bernard E.C., Howarth F.G., Sommer S., Soto-Adames F.N., Taiti S., Mockford E.L., Horrocks M., Pakarati L. and Pakarati-Hotus V. 2014. Disturbance Relicts in a Rapidly Changing World: The Rapa Nui (Easter Island) Factor. BioScience 64: 711-718

**Figure 1: *Tetragnatha* collected in Rapa Nui.** The collected specimen clusters with Hawaiian and American species in the phylogenetic analysis. However, it is not possible to determine if this corresponds to a natural colonization from the South America or a human introduction. Further sampling from the west coast of South America is needed to determine this.

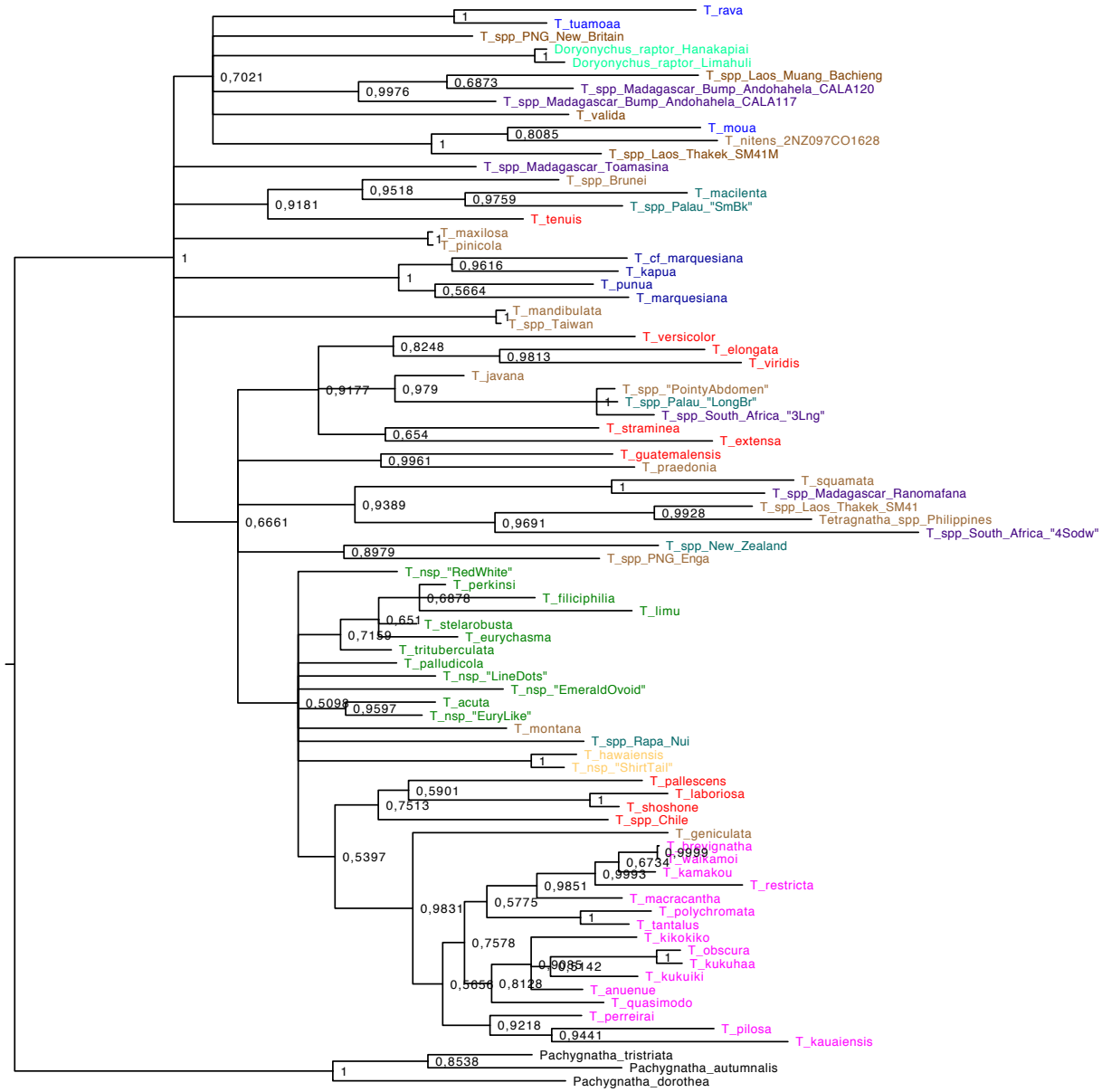



**Figure 2: Bayesian phylogenetic reconstruction using COI.** The details for the parameters are described in the Methods section. The support value in the nodes represents posterior probabilities. Values below 0.5 are not shown. The present reconstruction shows in different colors the geographic location of the different species. Algae green: Pacific species; Black: Outgroup, *Pachignatha*; Blue: Society Islands; Brown: Asia; Dark blue: Marquesas; Dark green: Web Building clade from Hawai'i; Light green: *Doryonoychus raptor*; Light purple: Spiny leg clade from Hawai'i; Orange: *T. hawaiiensis* clade; Purple: Africa; Red: America.

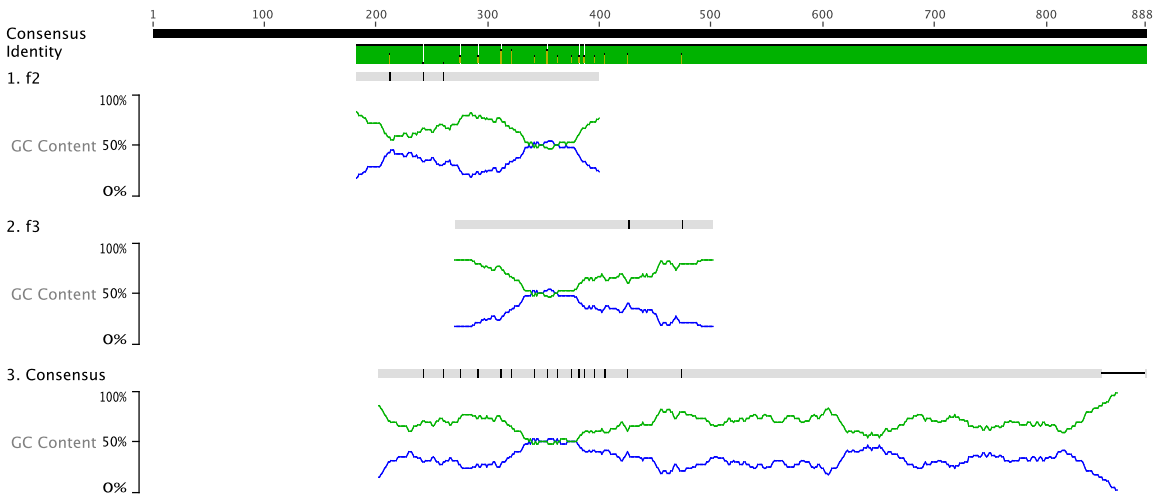



**Figure 3: GC content on the region of difficult amplification.** The figure shows in the upper portion the two fragments corresponding to the PCR product of the primer pairs 2 (f2) and 3 (f3). The sequences correspond to the sequences used as a template to produce the primers. The bottom sequence (Consensus) corresponds to a consensus sequence of all the sequences used for the phylogenetic analysis. The GC content is represented by the blue line and the AT by the green. Note that between 300 bp and 400bp there is a rich-GC area. This figure was produced using Geneious (Drummond et al., 2012).

**Table 1: Sequences used from GeneBank**

| Species Name | Geographic Origin | GeneBank accession number |
|---|---|---|
| *Pachygnatha autumnalis* | Canada | HQ924642 |
| *P. dorothea* | Canada | HQ924456 |
| *P. tristriata* | Canada | GU682654 |
| *Tetragnatha acuta* | Hawai'i, Hawai'i Island | AY803906 |
| *T. anuenue* | Hawai'i, Hawai'i Island | AY530502 |
| *T. brevignatha* | Hawai'i | DQ178962 |
| *T. cf. marquesiana* | Marquesas, Ua Huka | EU796924 |
| *T. elongata* | - | RBCH142 |
| *T. eurychasma* | Hawai'i, West Maui | AY803903 |
| *T. extensa* | Canada | KKCHE820 |
| *T. filiciphilia* | Hawai'i, East Maui | AY803902 |
| *T. guatemalensis* | USA | EU796928 |
| *T. hawaiensis* | Hawai'i, O'ahu | AY490338 |
| *T. javana* | Pakistan | JX294517 |
| *T. kamakou* | Hawai'i, Moloka'i | DQ182759 |
| *T. kapua* | Marquesas, Hiva Oa | EU796926 |
| *T. kauaiensis* | Hawai'i, Kaua'i | AY490333 |
| *T. kikokiko* | Hawai'i, Maui | AY490367 |
| *T. kukuhaa* | Hawai'i | AY490365 |
| *T. kukuiki* | Hawai'i, O'ahu | AY490362 |
| *T. laboriosa* | USA | EU796929 |
| *T. limu* | Hawai'i, O'ahu | AY803901 |
| *T. macilenta* | Society Islands, Mo'orea | EU796931 |
| *T. macracantha* | Hawai'i, Maui | DQ182767 |
| *T. marquesiana* | Marquesas, Nuku Hiva | EU796923 |
| *T. moua* | Society Islands, Tahiti | EU796906 |
| *T. nitens* | New Zealand | EU796912 |
| *T. nsp. "emerald ovoid"* | Hawai'i, O'ahu | AY803899 |
| *T. nsp. "eury like"* | Hawai'i, O'ahu | AY803904 |
| *T. obscura* | Hawai'i | AY490364 |
| *T. pallescens* | USA, California | AY803897 |
| *T. perkinsi* | Hawai'i, Hawai'i Island | AY803905 |
| *T. perreirai* | Hawai'i, O'ahu | AY490340 |
| *T. pilosa* | Hawai'i, Kaua'i | AY490341 |
| *T. pinicola* | - | JN817127 |
| *T. polychromata* | Hawai'i, O'ahu | AY490345 |
| *T. praedonia* | - | JN817124 |
| *T. punua* | Maruqesas, Nuku Hiva | EU796919 |
| *T. quasimodo* | Hawai'i, Hawai'i Island | AY530474 |
| *T. rava* | Society Islands, Tahiti | EU796902 |
| *T. restricta* | Hawai'i | DQ182764 |

| | | |
|---|---|---|
| *T. squamata* | - | JN817128 |
| *T. stelarobusta* | Hawai'i, Maui | AY803900 |
| *T. straminea* | Canada | HQ924664 |
| *T. tantalus* | Hawai'i, O'ahu | AY490368 |
| *T. tuamoaa* | Society Islands, Mo'orea | EU796904 |
| *T. versicolor* | - | FJ525317 |
| *T. viridis* | USA: Massachusetts | HQ979340 |
| *T. waikamoi* | Hawai'i, Maui | AY490374 |
| *Tetragnatha spp. Taiwan* | Taiwan | HQ441947 |

**Table 2: Sequences included on this study**

| Species name | Geographic origin | Origin of the sample | Preservation Condition |
|---|---|---|---|
| *Doryonychus raptor* Limahuli (2919) | Hawai'i, Kaua'i Limahuli Preserve | Field collected Darko Cotoras | -20°C, 95%ethanol |
| *D. raptor* Hanakapia'i (2944) | Hawai'i, Kaua'i Hanakapia'i | Field collected Darko Cotoras | -20°C, 95% ethanol |
| *Tetragnatha geniculata* (SM29) | Laos, Louangphabang | Senckenberg Museum (56685-121) | Room temperature, 70% ethanol |
| *T. mandibulata* (SM12) | Thailand, Trat | Senckenberg Museum (60430-121) | Room temperature, 70% ethanol |
| *T. maxilosa* (CA1) | Philipines | California Academy of Sciences (Ph0048) | -20°C, 95% ethanol |
| *T. montana* (SM16) | Germany, Rheinland Pfalz | Senckenberg Museum (59380-121) | Room temperature, 70% ethanol |
| *T. nsp. "Line dots"* (2265) | Hawai'i, Hawai'i island | Field collected Darko Cotoras | -20°C, 95% ethanol |
| *T. nsp. "Red white"* (905) | Hawai'i, Hawai'i island | Field collected Darko Cotoras | -20°C, 95% ethanol |
| *T. nsp. "Shirt tail"* (2082) | Hawai'i, Hawai'i island | Field collected Darko Cotoras | -20°C, 95% ethanol |
| *T. palludicola* (3071) | Hawai'i, Maui | Field collected Darko Cotoras | -20°C, 95% ethanol |
| *T. shoshone* | = | Essig Museum (2Ta084CO1628) | Sequence |
| *T. tenuis* | = | Essig Museum s003HCO2354030900 | Sequence |
| *T. trituberculata* (3077) | Hawai'i, Maui | Field collected Darko Cotoras | -20°C, 95% ethanol |
| *T. valida* | - | Essig Museum (2Aus00027) | Sequence |
| *Tetragnatha. spp.* Laos, Thakek SM41 | Laos, Thakek | Field collected Peter Jeager (Senckenberg Museum) | -20°C, 95% ethanol |
| Tetragnatha spp. Laos, Thakek SM41M | Laos, Thakek | Field collected Peter Jeager (Senckenberg Museum) | -20°C, 95% ethanol |
| *Tetragnatha spp*. Laos, Muang Bachieng (SM36) | Laos, Muang Bachieng | Field collectes Peter Jeager (Senckenberg Museum; SD458) | -20°C, 95% ethanol |
| *Tetragnatha spp.* PNG, New Britain (CD14) | Papua New Guinea, New Britain | Essig Museum (CD14) | Room temperature, 95% ethanol |

| | | | |
|---|---|---|---|
| *Tetragnatha spp.* Philippines | Philippines, Luzón Island | California Academy of Sciences (Ph0048 light dark) | -20°C, 95% ethanol |
| *Tetragnatha spp.* Madagascar, Bump Andohahela CALA117 | Madagasca, Bump Andohahela NP | California Academy of Sciences (HW0806) | -20°C, 95% ethanol |
| *Tetragnatha spp.* Madagascar, Bump Andohahela CALA120 | Madagasca, Bump Andohahela NP | California Academy of Sciences (HW0808) | -20°C, 95% ethanol |
| *Tetragnatha spp.* Chile | Chile, X Region | California Academy of Sciences (ChileRegXTTsp) | -20°C, 95% ethanol |
| *Tetragnatha spp.* PNG, Enga (CD15) | Papua New Guinea, Enga Province | Essig Museum (CD15) | Room temperature, 95% ethanol |
| *Tetragnatha spp.* Madagascar, Ranomafana (CALA110) | Madagascar, Ranomafana NP | California Academy of Sciences (HW0816) | -20°C, 95% ethanol |
| *Tetragnatha spp.* Madagascar, Toamasina (CALA113) | Madagascar, Toamasina | California Academy of Sciences (HW0801) | -20°C, 95% ethanol |
| *Tetragnatha spp.* Brunei (CA2) | Brunei, Merimbum | California Academy of Sciences (TM002) | -20°C, 95% ethanol |
| *Tetragnatha spp.* New Zealand (CALA99) | New Zealand, Mt. Richmond Forest Park | California Academy of Sciences (CGNZ19) | -20°C, 95% ethanol |
| *Tetragnatha spp.* Palau "LongBr" | Palau | Essig Museum (PalauLongBr081CO2396Y884) | -20°C, 95% ethanol |
| *Tetragnatha spp.* Palau "SmBk" | Palau | Essig Museum (PalauSmBk077CO1675Y886) | -20°C, 95% ethanol |
| Tetragnatha spp. "Pointy Abdomen" (CALA102) | Burma, Moezingyi | California Academy of Sciences | -20°C, 95% ethanol |
| *Tetragnatha spp.* Rapa Nui (2874) | Rapa Nui | Field collected Darko Cotoras | -20°C, 95% ethanol |
| *Tetragnatha spp.* South Africa "3Lng" | South Africa | California Academy of Sciences (SAfrica3LngTlTSA5) | -20°C, 95% ethanol |
| *Tetragnatha spp.* South Africa "4Sodw" | South Africa | California Academy of Sciences (SAfrica4SodwBaTSA4) | -20°C, 95% ethanol |

**Table 3: Specimens used from the Senckenberg Museum**

| **Species name** | **Geographic origin** | **Collection number Senckenberg Museum** |
|---|---|---|
| *T. praedonia* (SM1) | Laos, Muang Bachieng | 60354 – 121 |
| *T. cf. montana* (SM2) | Germany, Hesse | 56146 - 121 |
| *T. ceylonica* (SM3) | Laos, Louangnamtha | 58708 - 121 |
| *T. obtusa* (SM4) | Tunisia: Jendouba (Jundubah) | 60657 - 121 |
| *T. nitens* (SM5) | Italy: Sardegna | 62452 - 121 |
| *T. dearmata* (SM7) | Germany, Baden-Wuerttemberg | 60653 - 121 |
| *T. pinicola* (SM8) | Germany, Rheinland-Pfalz | 59381 - 121 |
| *T. nitens* (SM9) | Spain, Canarias Islands | 60648 - 121 |
| *T. pinicola* (SM10) | Germany, Hesse | 56917 - 121 |
| *T. mandibulata* (SM12) | Thailand, Trat | 60430 - 121 |
| *T. lauta* (SM14) | Laos, Louangnamtha | 58707 - 121 |
| *T. nitens* (SM15) | Spain, Baleares Islands | 60651 - 121 |
| *T. montana* (SM16) | Germany, Rheinland-Pfalz | 59380 - 121 |
| *T. praedonia* (SM17) | Japan, Kyoto | 56521 - 121 |
| *T. geniculata* (SM18) | Laos, Champasak | 60353 - 121 |
| *T. nitens* (SM19) | Greece, Lasithi | 60654 - 121 |
| *T. montana* (SM20) | Bulgaria, Ruse | 56547 - 121 |
| *T. extensa* (SM22) | Germany, Nordrhein-Westfalen | 56733 - 121 |
| *T. obtusa* (SM23) | Germany, Rheinland-Pfalz | 59385 - 121 |
| *T. hasselti* (SM26) | Laos, Viangchan | 62206 - 121 |
| *T. obtusa intermedia* (SM27) | Portugal, Algarve | 60655 - 121 |
| *T. mandibulata* (SM28) | Laos, Oudomxai | 62207 - 121 |
| *T. geniculata* (SM29) | Laos, Louangphabang | 56685 - 121 |
| *T. shoshone* (SM30) | Germany, Mecklenburg-Vorpommern | 63532 - 121 |
| *T. mandibulata* (SM33) | Laos, Viangchan | 58709 - 121 |

**Table 4: Primer sequences for short fragment amplification**

| Fragment (product size) | Primer | Sequence | Annealing temperature (°C) |
|---|---|---|---|
| F1 (217 bp) | Left | GATCTGCTATAGTAGGGACAGC | 55 |
| | Right | ATTCGAGGAAAAGCCATATC | |
| F2 (216 bp) | Left | TATTTTAATTGGGGGATTCG | 55 |
| | Right | TCTACAGATCTTCCTGAATGC | |
| F3 (300 bp) | Left | TTTTGATTACTACCCCCTTC | 42 |
| | Right | ATAGAAACCCCTCTTATTCG | |
| F4 (241 bp) | Left | GGGCTTCTTCTATTATAGGG | 53 |
| | Right | AAATGTTGAAACAAAACAGG | |
| F5 (205 bp) | Left | TGCAGTTTTATTACTTTTATCTTTAC | 53 |
| | Right | TCTAACAACATGAGAAACAATACC | |
| F6 (212 bp) | Left | AGGGGATCCTGTTTTGTTTC | 55 |
| | Right | CCACAGAAAATATATGATGAGC | |

**Table 5: Amplification result using six primer pairs for short fragments and the universal primer.**
In order to complete the gaps all the possible combinations of primers where tested. Green box: left and right read amplified; Red box: only left read amplified; Blue box: only right read amplified; Pound sign (#): amplification product produced using the Left primer for fragment 1 and the Right for fragment 4.

| Sample | F1 | F2 | F3 | F4 | F5 | F6 | FULL (HCO/LCO) |
|---|---|---|---|---|---|---|---|
| SM1 | | | | | | | |
| SM2 | | | | | | | |
| SM3 | | | | | | | |
| SM4 | | | | | | | |
| SM5 | | | | | | | |
| SM7 | | | | | | | |
| SM8 | | | | | | | |
| SM9 | | | | | | | |
| SM10 | | | | | | | |
| SM12 | ######### | ######### | ######### | ######## | | | |
| SM14 | | | | | | | |
| SM15 | ######### | ######### | ######### | ######### | | | |
| SM16 | | | | | | | |
| SM17 | | | | | | | 342 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| SM18 | | | | | | | 516 |
| SM19 | | | | | | | |
| SM20 | | | | | | | |
| SM22 | | | | | | | 494 |
| SM23 | | | | | | | 492 |
| SM26 | | | | | | | 506 |
| SM27 | | | | | | | 511 |
| SM28 | | | | | | | 503 |
| SM29 | | | | | | | 464 |
| SM30 | | | | | | | 498 |
| SM33 | | | | | | | |

# CHAPTER 2

## Effects of a dynamic landscape on the mitochondrial haplotype diversity of Hawaiian *Tetragnatha* spiders


## ABSTRACT

Adaptive radiations have been of interest for evolutionary biologists for a long time. The study of their temporal dynamic it is not a simple problem to address in natural conditions. However, the chronosequence of the Hawaiian Islands provides "snapshots" of different stages of the diversification process. Here, we propose a population-based mechanism, which associated with the changes on the landscape attempts to explain the patterns of species diversity. By comparing species present on the young (Big Island: *T. anuenue* and *T. brevignatha*) and middle age islands (Maui: *T. waikamoi* and *T. brevignatha*), we tested the following predictions: (1) higher genetic structure in the young island populations and (2) lower genetic structure in populations from the middle age island. To test these hypotheses we sequenced three mitochondrial genes (COI, ND1 and Cytb). For our first prediction we found that the mitochondrial haplotype distribution is not related to the specific volcano configuration, but rather with breaks in forest types associated with rain regimens. The second prediction was not supported. Instead, we found strong genetic isolation between populations on Maui. Finally, a time calibrated coalescent reconstruction of *T. brevignatha* provided evidence for a population expansion during Pleistocene glacial periods. It also suggests that the Windward/Leeward population break is related to more recent events than the original colonization of the island. Species evolving in Hawai'i seem to be affected not only by the geologic development of the islands, but also by more recent climatic events.

## INTRODUCTION

Adaptive radiations have been of interest for evolutionary biologists for a long time (reviewed on Schluter, 2000; Seehausen, 2006; Losos and Ricklefs, 2009). In general terms they correspond to the evolution of a large number of species in a very short period of time. This phenomenon is often associated with a particular ecological trait that allows for the "exploration" of different niches (Yoder et al., 2010). The trigger for adaptive radiations corresponds to the opening of an ecological space, either through colonization of a novel environment or by the development of a key innovation. The main factors that determine the places where they take place tend to be related to isolation and area. While the magnitude of the radiation likely depends on available ecological opportunity, how does a radiation get started in the first place? The study of the temporal dynamic of the radiation has been addressed theoretically (Gavrilets and Vose, 2005) and in laboratory experiments with bacteria (Meyer et al, 2011), but in natural conditions it is not a simple problem.

The isolation and geographic conditions of Hawai'i have produced adaptive radiations in multiple taxa (Price and Clague, 2002). This volcanic archipelago is the most isolated group of islands on Earth. The main Hawaiian Islands are much larger compared to other volcanic archipelagoes. They include a wide diversity of environmental conditions and microhabitats, including low and highland rainforest, dry forest, deserts and alpine environments (Ziegler, 2002). The interplay of these environmental and geographic conditions has transformed Hawai'i into a natural laboratory for evolution (Simon, 1987). In addition to the environmental and geographic conditions, the constantly changing landscapes also have played an important, and perhaps even more important, role.

The sequential structure of the ages of the islands provides "snapshots" of different stages of the diversification process. The "snapshot" that represents the initial stages of speciation will correspond to an island that is still volcanically active. At this time islands reach their largest area/altitude and their dynamic nature in many cases isolate populations. The dynamic and recent volcanic history of the Big Island (USGS Hawai'i Volcano Observatory http://pubs.usgs.gov/of/2007/1089/HawIsland_zone5_2007.pdf) makes it possible to examine population structure as species are diverging (Vandergast et al, 2004), and for this reason the Big Island has been called a "Crucible of Evolution" (Carson et al., 1990).

Intermittent lava flows fragment the forest and create a complex landscape structure, which could potentially contribute to increased genetic diversity (Carson and Templeton, 1984). Within this landscape, species boundaries are obscure. Population studies on *Drosophila* and *Tetragnatha* spiders in the Kīpuka region (south slope of Mauna Loa) of the Big Island showed a meta-population dynamic where repetitive episodes of local extinction and re-colonization have an important role in generating genetic diversity (Carson et al. 1990; Vandergast et al., 2004).

Another factor that contributes to the production of new species is the presence of five high elevation volcanoes (Ziegler, 2002). The high elevation areas provide habitat for rainforests. In the windward side there are four volcanoes (Kohala, Mauna Kea, Mauna Loa and Kīlauea) through which there is a continuous belt of rainforest, with an area of mesic habitat between Kohala and Mauna Kea (Jacobi, 198; Price et al., 2007).

This north to south distribution may eventually produce isolation by distance. In contrast, Hualālai and the west side of Mauna Loa are located in the rain shadow area of the island, but because of convective rains they also support rainforests. However, it is effectively isolated from the forest of the other volcanoes (Eldon et al., 2013). Therefore, on the spatial scale of the island of Hawai'i there might be a combination of isolation by distance and vicariant effects.

In the Hawaiian Islands, large lineages (Animals: *Platyini*, *Tetragnatha*, *Drosophila, Nesosydne, Laupala*; Plants: *Lobeliads*, *Cyrtandra*) peak in species diversity per unit of area on middle age islands (Maui or O'ahu), while small lineages show a constant increase in diversity in the direction of the oldest islands (Gillespie and Baldwin, 2010). The identification of the factors that contribute to the shape of this biodiversity dynamic will provide insights into the temporal dynamics of adaptive radiations.

High species diversity on middle age islands has been predicted in the general dynamic model of oceanic island biogeography (Whittaker et al., 2008). This model has been letter tested with data sets from different islands groups (Cameron et al., 2013) and using different regression models (Steinbauer et al., 2013). The general prediction is that biological diversity will track the carrying capacity of the island. Thus, species richness will show a parabolic trajectory because on young islands diversity will be low and it will reach a peak on middle age islands. Later, due to erosion, landslides and subsidence the diversity will decline as the islands become smaller.

Here, we propose a population-based mechanism to explain the rise of species diversity. Initially on a young volcanic island, as the Big Island, the different populations of a species will show high population structure due to geographic isolation. Then, as the island reaches its middle age, like Maui Nui, those isolated populations would diversify into new allopatric species.

The Hawaiian *Tetragnatha* spiders provide a model system to study adaptive radiations and their temporal dynamic (Gillespie et al., 1994; Gillespie, 2004). In the Hawaiian Islands there are more than 50 species in less than 5 million years. Most of these species are likely part of two independent adaptive radiations. One radiation produced large number of species and corresponds to the web-building clade (Blackledge and Gillespie, 2004). The second corresponds to the "spiny-leg" clade and shows dramatic ecological shifts and convergent ecomorphs across different islands. There are 4 described ecomorphs: green, maroon, small brown and large brown. On O'ahu and Maui, all of them are present, but on Kaua'i and Big Island the maroon ecomorph is missing. On some volcanoes there are more than one species per ecomorph, but never in sympatry. For example, on Haleakalā, Maui, there are 3 species of the green ecomorph (Gillespie, 2004).

The *Tetragnatha* spiders provide to an ideal taxon to answer the question: What is the effect of the dynamic geologic history of Hawai'i on the mitochondrial haplotype diversity of populations of *Tetragnatha* spiders? We hypothesize that by comparing the relative mtDNA haplotype diversity of species present on young (Big Island) and middle age (Maui) islands we will find the following general patterns:

1.- There will be a higher genetic structuration of species present on the Big Island, with respect to the ones on Maui, due to the geographic configuration of the island.

2.- There will be lower genetic structure of the Maui species, compared to the Big Island, because they are the result of isolated populations of an ancestral highly genetically structured species.

Note that the second prediction builds on the first one; so testing both of them will provide evidence to support the hypothesized mechanism. To answer the question about the role of the dynamic geologic history of Hawai'i in shaping the population structure, we will examine the mitochondrial haplotype diversity of three forest specialist *Tetragnatha* species from the Spiny-leg clade (Fig 1). In particular, the first prediction will use *T. anuenue* (Small brown ecomorph), which is present only on the Big Island in the volcanoes of Kohala, Mauna Kea, Mauna Loa (Kīpuka area and Ka'u) and Kīlauea. Here our expectation is to find highly structured populations, which are correlated with the different histories of the volcanoes.

For the second prediction we will examine *T. waikamoi* (Green ecomorph) endemic to Maui with populations on West Maui and Haleakalā (Upper Waikamoi). This species is closely related to other species representing the Green Ecomorph (*T. brevignatha* and *T. macracantha*). If our hypothesis of allopatric speciation within the island is correct, a natural expectation will be that these three species are the result of an allopatric speciation event. Then, the population structure within each species should be rather low. Our prediction is to find many shared haplotypes between the West Maui and Haleakalā populations and demographic signatures of population expansion in the volcano that was colonized the last.

Finally, we will study the population structure of *T. brevignatha* (Green ecomorph) on Maui (Lower Waikamoi) and on the Big Island (Mauna Kea, Mauna Loa (Kīpuka and Ka'u), Kīlauea and Hualālai). In this case we expect to see a pattern that combines both predictions in the context of the "Progression rule" (Funk and Wagner 1995). In other words, we expect that the population in Maui will be ancestral to the one present on the Big Island. This will be suggested by the genetic diversity of the Big Island which should be a subsample of the Maui population and also have a strong signature of population expansion since colonizing the Big Island. Also, as indicated by our first prediction, we expect to see a similar genetic structure for the Big Island population of *T. brevignatha* as that of *T. anuenue*; the distribution range of *T. anuenue* is contained within that of *T. brevignatha*.

## MATERIALS AND METHODS

### Collections

**S**pecimens of *T. anuenue, T. brevignatha* and *T. waikamoi* were collected during five field seasons (August 2010, June 2011, January 2012, June 2012 and June 2013). The sampling took place during the day and night using a beating sheet and by hand collecting. The visited sites corresponded to State, Federal and private land and access was obtained by 4WD vehicle, hiking and helicopter transportation. The exact collecting sites are given on Table 1. All the specimens were preliminary identified in the field and then confirmed in the lab. While alive they were photographed with a Nikon 3100 SLR camera with an 85 mm macro lens using the same black paper as a background. These

pictures were useful to corroborate identification later on. The samples were preserved in 95% ethanol at -20°C for genomic work. We also collected specimens of *T. macracantha* from Ko'olau Forest Reserve (Maui) for comparative purposes.

**DNA sequences**

The DNA was extracted form 4 legs from the right side of each the spider using the QIAGEN DNA easy kit. The mitochondrial genes Cytochrome oxidase subunit I (COI), NADH dehydrogenase 1 (ND1) and Cytochrome b (Cytb) were amplified for all the specimens. The master mix for all the PCR reactions consisted in: 1 µl of each primer at 10 µM, 2 µl of AmpliTaq buffer 10x, 0,5 µl of $MgCl_2$ 25 mM, 11.7 µl $H_2O$, 1 µl BSA 1x, 0.2 µl DreamTag and 1 µl of DNA extraction. In the cases were the PCR reaction was not successful we included betaine (Sigma-Aldrich) following the manufacturer instructions. This situation occurred in most of the samples of *T. anuenue* and *T. waikamoi*.

The amplification profile that we used for all the reactions started with 2 minutes at 95°C, followed by 40 repetitions of a cycle which started with 30 seconds at 95°C, then 45 seconds at the specific annealing temperature and finally one minute at 72°C; there was an extra step of extension at 72°C for ten minutes. The primer sequences and annealing temperatures are present in Table 2.

The PCR products were verified on an agarose/TBE gel. PCR products were cleaned using Axygen AxyPrep MagPCR Clean-up kit. DNA was sequenced directly in both directions through the cycle sequencing method using dye terminators (Sanger et al., 1977) using Life Technologies' BigDye Terminator v3.1 Cycle Sequencing Kit. Sequenced products were cleaned using Axygen AxuPrep Mag DyeClean kit and run out on ABI 3730XL DNA Analyzer automated sequencer. The sequences were edited, concatenated and align with Geneious Pro 5.6.7 (Biomatters). The sequence alignment used the algorithm MAFFT (Katoh et al., 2002) with the default parameters.

**Primer Design**

The primers for ND1 and Cytb were designed using the web-based program Primer 3 (Koresaar and Remm, 2012; Untergrasser et al, 2012) utilizing as a reference the respective sequences obtained from a whole transcriptome sequencing of a specimen of *T. brevignatha* from Big Island.

**Sequence analysis**

Each marker was analyzed separately and then concatenated. The concatenation was done using the program Geneious Pro 5.6.7 (Biomatters). For each alignment we collapsed identical haplotypes using the on-line tool DNA Collapser present on the web site FaBox 1.41 (Villesen, 2007). The files were exported in arp format and then analyze with Arlequin 3.5 (Excoffier et al., 2005) to estimate basic molecular diversity indices, number of haplotypes, gene diversity, number of polymorphic sites (S) and nucleotide diversity (pi).

To test for signal of recent demographic processes such as population expansion

or contraction we also performed two different neutrality tests (Tajima's D (Tajima, 1989) and Fu's Fs (Fu, 1993)). The statistical significance was assessed by comparing the observed values with the values obtained from a distribution of 10,000 coalescent simulations. Significant negative values of Tajima's D are interpreted as population expansion (or selective sweep), while positive values correspond to population contraction (or diversifying selection) (Eytan and Hellberg 2010). Because Fu's Fs is more sensitive on detecting population expansion the P-value for significance from a coalescent simulation is established at 0.02 instead of the common 0.05.

Population structure among sites, volcanoes and islands was determined with analysis of molecular variance (AMOVA) implemented in Arlequin 3.5 (Excoffier et al., 2005). As a complement, using the same program we also estimated the pairwise $\phi_{ST}$ with a significance threshold of 0.05 obtained from 1023 permutations.

Haplotypic networks for each marker were generated using the program HapNetwork (Salzburger et al., 2011). It requires as an input file a FASTA file and a phylogenetic tree with sequences. The FASTA file was generated using alignments from Geneious Pro 5.6.7 (Biomatters). For the phylogenetic reconstructions we used RAxML (Stamatakis, 2014) available on the remote server CIPRES Science Gateway version 3.3 (Miller et al., 2012). The data was partitioned per codon position, and the model GTR+GAMMA with 1000 bootstrap replicates was used. The logic behind partitioning the data is that it will calculate different parameters for each partition even if they use the same model of molecular evolution.

To study the population structure of *T. brevignatha* for possible support of the "Progression Rule" we did a time calibrated coalescent reconstruction. The appropriate models for molecular evolution were estimated using PARTITIONFINDER (Lanfear et al., 2012). The concatenated sequence (1,495 bp) had 6 partitions. The favored models were: COI position 1: HKY; COI position 2 and ND1 position 2: GTR+G; COI position 3: K80; Cytb position 1, Cytb position 2, ND1 position 1 and ND1 position 3: HKY+I; and Cytb position 3: TrN+G.

The phylogenetic reconstruction was done with BEAST 1.8 (Drummond et al., 2012) implemented in the remote server CIPRES Science Gateway version 3.3 (Miller et al., 2012). The nexus file was created using the program BEAUti 1.8 (Drummond et al., 2012). We tried two different sets of molecular clocks. The first corresponded to the standard rate for arthropods for COI (2.3%; Brower, 1994), and the rates for ND1 and Cytb were estimated. The model for the estimated molecular clocks (ND1 and Cytb) corresponded to a Lognormal relaxed clock (uncorrected). The second approach was to estimate directly the mutation rate for each marker. For this we considered all the available sequences generated from this study for each marker and using MEGA6 (Tamura et al., 2013) to calculate the uncorrected P-distance comparing Maui and Big Island. For the rate estimation, we considered as a calibration point 0.5 MY, which is the time by when Kohala (the oldest volcano of Big Island) was high enough to have rain forests present. The tree prior used was a coalescent tree with constant size. The analysis was run for 100,00,000 generations with a sampling frequency of every 1,000 trees. The parameter convergence of two independent runs was determined by estimating an Effective Sample Size bigger than 200 (ESS>200) on TRACER version 1.7.5. (Rambaut et al., 2014). Later, the two runs were combined using LogCombiner 1.8, prior to a burn in of 25% of the trees. The final tree was generated on TreeAnnotator 1.8 and visualized

on FigTree version 1.4.0 (http://tree.bio.ed.ac.uk/software/figtree/). The population of Maui was used as the outgroup.

## RESULTS

All the markers were amplified from the same specimens, but not all amplifications were successful. Markers amplified from each specimen are shown in Table 3 and a summary of the final totals in Table 4. For *T. anuenue* it was not possible to amplify ND1. Results are presented for each species separately.

### Big Island endemic: *T. anuenue*

The highest nucleotide diversity was found on Kīlauea. This pattern is present in the two sequenced markers and the concatenated data set (Tables 5, 6 and 7). Sample sizes for Kīlauea and Mauna Loa were similar and significantly higher than the specimens amplified in Kohala, Mauna Kea and Ka'u (collecting sites on Fig. 2a).

The haplotype network of COI (Fig. 2b) shows that most of the haplotypes occur on the windward volcanoes (Kohala, Mauna Kea, Mauna Loa and Kīlauea) and are shared by more than one site. The two most common haplotypes in the lower part of the diagram are more common in Mauna Loa (Kīpuka), while the ones in the upper part present higher incidence in Kīlauea (Pu'u Maka'ala). The only site that has endemic haplotypes is Ka'u. For Cytb (Fig. 2c), the pattern is similar. Most of the haplotypes are shared among localities on the windward area, with some of them more common at one site. As expected due to linkage in the mitochondrial genome the haplotype network of the concatenated sequences result shows a very similar pattern as the individual genes (Fig. 2d).

The grouping scheme in the analysis of molecular variance (AMOVA) that maximizes the variance among groups corresponds to the grouping of Leeward (Ka'u) vs Windward (all the rest) populations, instead of grouping by volcano (Table 8). This pattern is present in COI, Cytb and the concatenated sequence. The pairwise $\phi_{ST}$ analysis for COI (Table 9) shows significant differences between the pairs of: Kīlauea and Mauna Loa (0.12926); and Mauna Loa and Ka'u (0.38465). Note that the difference between Mauna Loa and Ka'u has more than a two-fold increase in comparison to the one between Kīlauea and Mauna Loa. Using the marker Cytb there is also a significant difference between Mauna Loa and Kīlauea (0.13301), but it is nearly four times smaller than the one between Kīlauea or Mauna Loa with Ka'u (Table 10).

In the neutrality tests, the only estimation that shows a significant signature of population expansion corresponds to the concatenated data set in the Mauna Loa population (Fu's Fs: -4.70771, P=0.01030) (Table 11). This same population shows values closer to a significant signature of population expansion using COI (Tajima's D: -1.49591, P=0.05480) and the concatenated data set (Tajima's D: -1.50956. P=0.05350). Grouping all the populations, it is not possible to detect a significant signal of population expansion for any marker or concatenated sequences using Tajima's D, but with Fu's Fs it is significantly detected for COI and the concatenated data set (Table 11).

**Maui endemic: *T. waikamoi***

*T. waikamoi* had only two populations. In terms of molecular diversity indices the pattern is similar across markers (Tables 12, 13 and 14). The population with the highest number of haplotypes and polymorphic sites (S) corresponds to Upper Waikamoi. While the nucleotide diversity of Pu'u Kukui is higher for COI and ND1, but not for Cytb. In the concatenated analysis the different indexes of genetic diversity are rather similar between both populations, however that could be an artifact produced by the small number of specimens which had the three genes amplified (Table 15).

The haplotype networks (Fig. 3) show a high diversity of closely related haplotypes in the Upper Waikamoi population and less diversity, but more divergence (except in Cytb), for the Pu'u Kukui population. There is a clear genetic break between these two populations. Only one COI haplotype (Fig. 3b) is shared between these two populations. Note that for the concatenated sequences the distance between different concatenated haplotypes is very high (Fig. 3e). This strong genetic break is supported by the estimations of $\phi_{ST}$, which ones are significant for all three markers (COI: 0.61996; Cytb: 0.55795; ND1: 056713).

Given that on this species there are just two populations and only one of them has two sites (Pu'u Kukui upper and Pu'u Kukui lower), there was no point in testing different grouping schemes (AMOVA). Note that the two sites on Pu'u Kukui do not have noticeable barriers. They just correspond to the lower and higher part of a boardwalk.

The neutrality tests do not show a strong and uniform signal across the different mitochondrial genes (Table 16). For the Upper Waikamoi population there is a significant value for Fu's Fs on the COI marker, for the other markers and Tajima's D it is not significant. In the Pu'u Kukui population as a whole, Tajima's D detects a significant signal of population expansion in Cytb. The site Pu'u Kukui lower has the same signal for Cytb and detected with Tajima's D. However, it is not present in the other markers or in the Fu's F values.

**Big Island and Maui endemic: *T. brevignatha***

The *T. brevignatha* data set will be considered first in the comparison between Maui and the Big Island, and then within the Big Island. For all markers, the nucleotide diversity of the Maui population is higher than any population om the Big Island or the island as a whole (Table 17, 18 and 19). In the case of the concatenated data (Table 20) set the nucleotide diversity on Maui is very similar to Mauna Kea. This is in contrast with the number of haplotypes per volcano, where Maui is among the ones with the lowest numbers. The $\phi_{ST}$ values between Maui and the combined populations of the Big Island are high and significant for all the markers and the combined sequences (COI: 0.85344 (P=0.000); Cytb: 0.88955 (P=0.000); ND1: 0.84537 (P=0.000); Concatenated: 0.90042 (P=0.000)).

The haplotype networks make this pattern very evident (Fig. 4). There are no shared haplotypes between the two islands and many mutation steps separate them. In the COI (Fig. 4b) and ND1 networks (Fig. 4d) it is possible to see that within the specimens from Maui (purple) there are two well-defined groups. A group of 3 specimens

(represented by three circles in COI and two circles, one of them larger in ND1) correspond to the same specimens in both networks. It was not possible to amplify Cytb for those specimens, which is why they are not present in Fig. 4c. These samples substantially contribute to the increase in nucleotide diversity. However, even for Cytb (Fig. 4c) where they are not present, the Maui population has a higher nucleotide diversity (0.010006 +/- 0.005844), than the entire Big Island (0.005926 +/ 0.003470). This pattern it is also present in the concatenated network (Fig. 4e).

*T. brevignatha* is a paraphyletic species (Gillespie, 2004). It has been shown that the Maui populations are genetically closer to *T. macracantha*, than to the Big Island population. For that reason we considered the possibility that those three divergent specimens could correspond to *T. macracantha*. Initially they were identified based on morphology. The characters that distinguish these species are short chelicerae and short spines in the tibial segment of the first pair of legs on *T. brevignatha* in comparison to *T. macracantha* (Fig. 5). These characters were present and *T. macracantha* still has never been reported in Lower Waikamoi, but to test our species identity using molecular data we included it on the ND1 haplotype network of *T. brevignatha* and *T. macracantha*. In the case that these three *T. brevignatha* individuals corresponded to *T. macracantha* but with *T. brevignatha* characters, we would expected to find them nested within the *T. macracantha* clade. However, that is not the case. No *T. macracantha* haplotype is shared with any *T. brevignatha* specimens (Fig. 6). Indeed the vast majority of the Maui *T. brevignatha* are more closely in mutation steps to *T. macracnatha* than the identified divergent lineage. This suggests that the divergent specimens are indeed very distinct specimens within the diversity of *T. brevignatha*. More detailed information on the molecular diversity indexes (Table S1, S2, S3 and S4) and Neutrality tests (Table S5) for the Ko'olau population of *T. macracantha* is provided in the Supplementary material.

Considering only the Big Island populations, the volcano with the highest nucleotide diversity is Kīlauea (COI: 0.009236 +/- 0.005621; Cytb: 0.006843 +/- 0.004549; and ND1: 0.012574 +/- 0.008065). For the markers Cytb and ND1, the values of nucleotide diversity on Mauna Kea are very similar to the ones on Kīlauea (Cytb: 0.006679 +/- 0.003974; ND1: 0.011251 +/- 0.006554). In terms of numbers of haplotypes, Kīlauea does not have a high number and is very similar to Maui.

The sections of the haplotype networks that correspond to the Big Island specimens (Fig. 4) present the same pattern. It consists of a couple of very common variants and several closely related rare ones. There is not a clear geographic distribution among them. In fact, it seems like most of the haplotypes are shared among all of the volcanoes. The only exceptions are three haplotypes (one very common and two rare) of Cytb present on the leeward side of the Big Island (Honoaula and Kona Hema).

For the analysis of molecular variance (AMOVA) we tested different grouping schemes in order to identify the one that maximizes the percentage of the variance explained by the among groups variation (Table 21). For all the markers and the concatenated sequences the one that explains the major portion of the variance is grouping all the populations from Big Island together and comparing them to Maui (Lower Waikamoi). This in contrast to the situation in which, all the sites are grouped per volcano (Table 21, columns 1 and 2). We also tested the optimal arrangement considering only the volcanoes on the Big Island. We tried 3 different scenarios:

1.- Each volcano separated (Kohala, Hualālai, Mauna Kea, Mauna Loa, Kīlauea and Southwest Mauna Loa): Note, Mauna Loa is divided in two: the Kīpuka area and Southwest Mauna Loa (Ka'u and Kona Hema)

2.- Windward (Mauna Kea, Mauna Loa and Kīlauea) / South (Ka'u) / Leeward (Honoaula and Kona Hema)

3.- Windward-South (Mauna Kea, Mauna Loa, Kīlauea and Ka'u) / Leeward (Honoaula and Kona Hema)

The contrast of Windward-South / Leeward was the one that maximized the variance among groups for COI, Cytb and the concatenated data set. The results obtained on ND1 indicate that the percentage of the variance explained by the among groups differences is zero. Given this outcome, we kept every volcano as a unit (Haleakalā, Mauna Kea, Mauna Loa and Kīlauea), as well as Ka'u and merged Honoaula and Kona Hema (Leeward) for the estimations of pair-wise $\phi_{ST}$.

For COI all the populations are significantly different to Leeward and Haleakalā. In addition, Leeward is significantly different from Ka'u (Table 22). Note that the $\phi_{ST}$ values in the comparison with the Maui populations are at least 5 times larger than the ones for the Leeward group. The detected significant differences correspond to the favored group configuration that optimizes the among groups variation for this marker. Cytb shows the same pattern (Table 23). The main difference corresponds to the detection of a significant difference between Mauna Loa and Mauna Kea, however it is one order of magnitude smaller that the $\phi_{ST}$ values for the other significant comparisons. Finally, ND1 has the same pattern has Cytb but differs in that all the comparisons with Leeward are not significantly different from zero (Table 24). This corresponds to the AMOVA analysis for this marker where the percentage of the variation explained by among groups variation was nominally zero.

To understand the temporal dynamic of the colonization of the Big Island, we did a time calibrated coalescent reconstruction. We tried two time calibrations, the first one using the standard strict molecular clock for arthropod COI (2.3%; Brower, 1994) and estimated the others (ND1: 4.74%; Cytb: 6.68%). For the second calibration we estimated the rate of molecular evolution for each marker using as a calibration point the estimated time when Kohala (Big Island) was high enough to support rain forests (0.5 MY) (Price and Clague, 2002). The rates calculated were: COI: 13.4%; ND1: 17.2%, and Cytb: 12%. These rates are high, but fall into the upper bound of other rates reported for Hawaiian arthropods (Goodman et al., 2012; Magnacca *per. comm.*)

Both reconstructions (Fig. 7 and 8) have generally the same topology, however the divergence times are very different. In the case of the tree calibrated with 2.3% for COI, the split between Maui and Big Island is 3.1925 MY ago. In the tree that was calibrated with the calculated rates is approximate 0.6917 MY. The last estimation is more likely as it is known that the oldest emerged lava from Big Island are from 0.9 MY (Lipman and Calvert, 2011; Goodman et al., 2012). So, we used this reconstruction (Fig. 7) in further analyses.

Within the Maui clade it was possible to identify the divergence (0.5743 MY ago) between the majority of specimens and the second lineage that was seen in the

haplotype networks. The first divergence between the specimens collected on Big Island is 0.1493 MY ago, but the majority of the divergences date to around 50, 000 years ago. The support value for the most recent divergences is, as expected, very low. There is no evident grouping by volcano, except for a clade that includes all, except for single specimens from Kona Hema, Ka'u and Honoaula (posterior 0.8098) and one specimen from Mauna Kea.

## DISCUSSION

### Population structure on a young volcanic island

Our first prediction was that there would be a higher genetic structure within species present on the Big Island in comparison to species from Maui. We expected this to result from the presence of five volcanoes with isolated patches of rainforest. To test this prediction we studied the population structure of two species of the Spiny leg clade that are forest specialist: *T. anuenue* (Big island endemic) and *T. brevignatha* (endemic to Maui and Big Island). For the purpose of testing this hypothesis we consider the Big Island population of *T. brevignatha*.

*T. anuenue* presents a population structure that follows the geographic arrangement of the landscape, but not organized by volcano. Instead, there is a break between the windward volcanoes (Kohala, Mauna Kea, Mauna Loa and Kīlauea) and the population on the Leeward side (Ka'u). This arrangement is the one that maximizes the variance among groups (AMOVA) and the $\phi_{ST}$ values which are high and significant for all markers. The haplotype networks (Fig. 2) also exhibit this pattern, as there are no shared haplotypes between the windward volcanoes and the Ka'u population.

The haplotype networks lack structure among the windward volcanoes, and there is a predominance of only a few haplotypes and the presence of many rare ones. This general structure suggests a population expansion, however this is only detected by Fu's F in COI and combined data set. There is no signature of population expansion for Cytb or detected by Tajima's D. All Tajima's D values are negative, but not significant.

A very similar pattern is found in the Big Island populations of *T. brevignatha*. The haplotype network shows a few very common alleles and many rare ones (Fig. 4). In this case both neutrality tests show significant signals for population expansion for all the markers and concatenated sequences when considering the Big Island as a single big population (Table 25).

The genetic structure within the different populations on the Big Island is not related to particular volcanoes. It is associated with a break between Windward (Mauna Kea, Mauna Loa and Kīlauea), Leeward (Hualālai and west Mauna Loa) and South (Ka'u) populations. $\phi_{ST}$ estimations based on COI and Cytb show significant differences between each of the windward volcanoes and the Leeward populations (Honoaula and Kona Hema). The Ka'u population the only volcano that has significant differences in $\phi_{ST}$ values with Mauna Kea. This situation explains why in the AMOVA the supported grouping is the one in which Ka'u is grouped with the Windward volcanoes and the Leeward sites correspond to another group. Indeed on the haplotype network for Cytb (Fig. 4d), it is possible to see that there are three haplotypes form Leeward volcanoes that are not shared with any other site.

The analysis of molecular variance on ND1 shows that 0% of the variation is explained by among group variations when we tested for the alternative scenarios ("By Volcano only Big Island", "Windward/South/Leeward" or "Windward-South/Leeward"). It can be partially explained by the fact that there are no significant differences on $\phi_{ST}$ in the comparison of the Windward volcanoes with the Leeward populations. This is an anomalous situation with respect to the other markers. For ND1 the Honoaula population has only 1 polymorphic site, while there are 4 in COI and Cytb.

In terms of signals of demographic processes on the different volcanoes within the Big Island, it is possible to find a wide variety of results (Table 25). Mauna Kea shows a strong signal for population expansion, which is detected by both neutrality tests for COI and Cytb. In the concatenated analysis Fu's Fs also shows this signal. Kīlauea and Kona Hema do not have any significant signal. In the case of Ka'u, only Fu's Fs in the ND1 data set shows population expansion, and the population from Hualālai (Honoaula) shows population expansion in the COI (Fu's Fs) and Cytb (Tajima's D) data set. The demographic processes on Mauna Loa (Kīpuka) are highly variable due to the recent lava flows and the subsequent metapopulation dynamics (Vandergast et al., 2004).

In summary, the first hypothesis is partially supported given the genetic structure associated with the landscape configuration. However, it was simpler than hypothesized and instead of a by volcano grouping, we found a break associated with Windward and Leeward island sides.

**Population structure on a middle age volcanic island**

In considering whether genetic isolation between populations on a young volcanic island is sufficient to produce species through allopatry, a prediction is that in middle aged islands it would be possible to find many sister species produced by this mechanism. Each of these new species is expected to have low genetic diversity, as they are the result of ancient, highly genetically differentiated populations. To test our second prediction we studied a species endemic to Maui (*T. waikamoi*). It has only two populations, one on Haleakalā volcano (Lower Waikamoi) and the other on West Maui volcano (Pu'u Kukui). Our original expectation was to find many shared haplotypes between these two sites. Due to a founder effect, the volcano that was colonized later would be a subsample of the diversity of the original source. Associated with this colonization pattern would likely be a bottleneck signal associated with the colonization event . Alternatively, if the colonization was old enough we expected to detect population expansion.

The original expectation was not supported by the data. With the exception of one haplotype in the COI data set, there were no shared haplotypes between the two populations (Fig. 3). This is also reflected in very high and significant $\phi_{ST}$ values (COI: 0.61996; Cytb: 0.55795; ND1: 056713).

In terms of signatures of demographic processes, there is a significant signal for population expansion in the COI dataset in the Upper Waikamoi population (Fu's Fs: -13.72184, P=0.00050) and the same signal is present in the Cytb dataset in the Pu'u Kukui population as a whole (Tajima's D: -1.61160, P= 0.03630). As the signals are not present in other markers (including the concatenated sequences) or congruent between tests, it is difficult to reach a more definitive explanation. Also, the directionality of the

colonization is impossible to infer based on the genetic diversity given the fact that there are no common haplotypes.

As a comparison, studies of *Theridion grallator* (Hawaiian Happy face spider) populations of Pu'u Kukui and Waikamoi also show strong breaks in population connectivity (Croucher et al., 2012). For these two populations the estimated $\phi_{ST}$ value based on COI was 0.5577, while the pair-wise $\phi_{ST}$ between Kamakou (Moloka'i) and Waikamoi (Haleakalā) was 0.3722. The separation of the Pu'u Kukui population is very evident on a PCA analysis of allozymes and mtDNA., and STRUCTURE analysis of the nuclear data identifies Pu'u Kukui as a separate group and clusters Kamakou and Waikamoi together. Lastly, the mitochondrial haplotypes present on Pu'u Kukui are endemic to this area and not shared with any other site.

This high degree of genetic diversity has been also shown for other spiders' species (*Ariamnes, Tetragnatha* and *Theridion*) present in Maui Nui (Roderick et al., 2012), and has been linked to longer periods of isolation among populations present in this older group of volcanoes.

However, these examples are not directly applicable to the hypothesis that we tried to test here. Because we hypothesized the presence of newly formed allopatric species, we centered our attention on *T. waikamoi*, which is closely related to the other two green ecomophs of *Tetragnatha* (*T. macracnatha* and *T. brevignatha*). Our main assumption was that these three species were produced by allopatric speciation within Maui Nui on its earlier stages.

In summary, these mitochondrial data set suggests long-term isolation between different volcano populations on Maui. Mitochondrial markers are under strong effects of genetic drift given the small population size in comparison with nuclear genes (Ballard and Whitlock, 2004). In the context of our hypothesis 2, it is possible that the *T. waikamoi* population has low genetic diversity at the nuclear level, but the colonization of one of these two volcanoes was old enough to generate new variants in the mitochondrial genome. In order to test this scenario a set of nuclear markers are required.

**Divergent lineages on the Maui population of *T. brevignatha***

The COI (Fig. 4b) and ND1 (Fig. 4d) haplotype networks and the coalescent reconstruction (Fig. 7) show the existence of three specimens that are very distinct from the rest of the samples. Morphological examination, together with genetic analysis of the sister species *T. macracnatha*, shows that these specimens are not misidentified. The obvious conclusion is that they represent a lineage that lives in sympatry with the rest of the species, but is genetically distinct. A closer observation of the original specimens shows the presence of black pigmentation at the tip of the pre-tarsus. That is the only distinctive morphological trait that we were able to identify.

The discovery of this cryptic haplotype diversity adds more complexity to the understanding of the population diversity. What are the forces that prevent the admixture of these two mitochondrial lineages? Is this signal also present in the nuclear genome? A deep sequencing of nuclear markers would allow approaching these questions.

***T. brevignatha* on Big Island: Colonization history or recent climatic events?**

To test for the progression rule in the colonization of the Big Island (Funk and Wagner 1995), we did a time calibrated coalescent reconstruction of the Maui and Big Island populations of *T. brevignatha*. For dating the phylogeny we used a strict molecular clock that we calculated directly from the uncorrected P-distances of the sequences present on each island and using as the calibration point the time between the respective island ages (0.5 MY). This strategy gave us more reliable dates than the ones obtained using the standard 2.3% rate for COI in arthropods.

All the coalescent events on Big Island happen closer to the present than our initial expectation (similar pattern occurs in *T. grallator* (Croucher et al., 2012). The oldest divergence occurs at 0.1493 MY ago, but the majority are estimated to have occurred during the last 50,000 years. Moreover, most of the specimens present on the different volcanoes do not group as monophyletic clades. This is congruent with the large number of shared haplotypes across the different sites (Fig. 4). The general shapes of the haplotype networks suggest a population expansion, due to the presence of few very common haplotypes and a large number of rare ones. Indeed, grouping all the populations of Big Island together produces a very strong signal of population expansion in all markers and the concatenated sequences in the two tests performed. All of this evidence points towards a population expansion that happens at some point during the last 50,000 years.

A careful examination of the coalescent reconstruction shows that there is only one region that groups as a monophyletic clade (Posterior 0.8098), and corresponds to the Leeward side of the island (Honoaula and Kona Hema). All, but one of the specimens, are present in the same clade. Also, in the Cytb haplotype network there are three haplotypes that are exclusive to this region. Note that this is also consistent with the group arrangement that maximized the variance among groups in the AMOVA. This group corresponded to the Windward volcanoes group with Ka'u, while the Leeward sites group with a different cluster. This suggests a single colonization event with further isolation and the development of endemic mitochondrial alleles.

In *T. anuenue* the genetic structure between the volcanoes on the windward side is not strong and many haplotypes are shared among them. The haplotype network also presents a few common variants and a larger amount of rare ones. However, the neutrality tests do not pick up that signal as strongly as in *T. brevigntha.* Only, Fu's Fs for the COI data set and the concatenated sequences shows a signal of population expansion.

The *T. brevignatha* data suggests that there is was a colonization event on the Leeward side of the island about11,500 years ago. This event is likely related to more recent demographic events rather than with the original colonization of the island. It could be related to demographic events associated with the Pleistocene glacial periods. One of the populations from Leeward (Honoaula) presents significant signal of population expansion in two molecular markers (COI and Cytb). This signal is not present in the concatenated sequence, perhaps influenced by ND1, which also lacks this signal.

The pollen record from the Ka'au Crater (463 m. Ko'olau ridge, O'ahu) suggests that between 28,000 to 25,000 $^{14}$C yr B.P. conditions were dryer and colder than at present (Hotchkiss and Juvik, 1999). During the following two thousand years

(25,000–23,000 $^{14}$C yr B.P.) the weather remained dry, but became warmer. Then, between 23,000 – 20,000 $^{14}$C yr B.P. conditions became only moderately dry and colder again. This is consistent with the record of the Older Makanaka glaciation at the top of Mauna Kea (23,000 years ago), which indicates wetter and colder conditions than today for this high elevation site (Pigati et al., 2008). These periods of dry and colder weather could have had a strong effect on the wet forest communities, in particular the ones present on the Leeward side of the Big Island. Mauna Kea and Mauna Loa create a very strong rain shadow effect on Hualālai and the west side of Mauna Loa. The rain forest present in those areas is produced by convective rain. During the day the center of the island warms creating a low pressure, which pulls moist air from the ocean. As the clouds move up the slope, they release this moisture and produce the rain, which maintains the wet forest. This rain regime produces afternoon rains and an annual maximum rainfall during the summer months (Sanderson, 1993). For that reason the rain here is highly dependent on high atmospheric temperatures.

For the period between 20,000 – 16,000 $^{14}$C yr B.P. the pollen record shows a return to dry and colder conditions. But, between 16,000 – 9,000 $^{14}$C yr B.P. the climate turned warm and wet again. This is congruent with the data from the Younger Makanaka glaciation (13,000 years ago) in Mauna Kea (Pigati et al., 2008). This event was recorded to be significantly wetter than the previous ones. Indeed, the expansion of the ice cap in the Younger Makanaka event is related with an increase in precipitation, instead of a decrease in temperature as happened during the Older Makanaka (Pigati et al., 2008).

The increase in precipitation 13,000 years ago (Younger Makanaka glaciation) is also seen at low elevations (pollen record). This suggests that the rain came from high-altitude storm. In current conditions these frontal systems reach Hawai'i from the west and not from the northeast like the trade winds (Pigati et al., 2008).

However, other lines of evidence suggest that the trade winds during the last glacier periods were similar to the regimens present today, although their intensity could have been lower. The orientation of the sand dunes in the saddle area between Mauna Kea and Mauna Loa, as well as the tephra plumes on Mauna Kea indicate the presence of trade winds from the northeast similar to today's pattern (reviewed in Gavenda, 1992). However, it is also important to consider that the cooling of tropical waters could have produced a reduction in the moisture and a southward shift of the northeast Pacific subtropical anticyclone, which is the origin of the trade winds that reach Hawai'i. This would have resulted in a potential decrease in their frequency (Hotchkiss and Juvik, 1999).

Even if the oldest lava from Hualālai are from the late Pleistocene (USGS Hawai'i Volcano Observatory http://pubs.usgs.gov/of/2007/1089/HawIsland_zone5_2007.pdf), its wet forest plant community appears young. This is suggested by plant species richness among the main floristic regions of the Big Island: Windward (294), Kohala (252), Ka'u (224) and Kona (212). A similar rank order is found on the regional endemics: Windward (7), Kohala (6), Kona (1) and Ka'u (0). For strict wet forest endemics, the Windward and Kohala have 4 species each, while Kona only 1 and Ka'u none (Eldon et al., 2013). This low species diversity in general and absence of endemism, suggests that the plant community of Kona and Ka'u is younger than Windward and Kohala.

Eldon et al. (2013) also suggested that as Mauna Loa and Mauna Kea increased in size the strength of the convective rain would increase. In this way, the present wet forest areas of Kona and Ka'u would have been colonized initially by mesic species. However, Ka'u would have had an extra influx of wet forest specialist due to its proximity to the Windward area. That was not an option to the developing Kona wet forest, which is isolated by an arid region from the Kohala (its nearest wet forest area). It is therefore more likely that the wet forest specialists would have colonized Kona later on from Ka'u.

A more detailed chronosequence of the weather changes during the last glacier cycles is required for the Leeward side of the Big Island. However, based in the current evidence, it is possible to consider that between 28,000 to 23,000 $^{14}$C yr B.P. the conditions were not favorable for the development or maintenance of wet forest habits (Hotchkiss and Juvik, 1999). But during the Older (23,000 years ago) and especially the Younger Makanaka events (13,000 years ago; Pigati et al., 2008), more wet conditions would have favored the expansion of the rainforest and its fauna. Our coalescent reconstruction shows that the clade that includes specimens from Honoaula and Kona Hema date from 11,500 years ago with an error that ranges approximately between 22,700 and 4,500 years ago. Thus, we suggest that the colonization of *T. brevignatha* on the Leeward side of the island is correlated with the establishment of the extant wet forest on that area.

There is another arthropod that is a wet forest specialist and shows the same pattern of population structure. An analysis of molecular variance of COII in *Drosophila sproati* shows that the highest percentage of the variance explained among groups is obtained when grouping Kohala + Windward and Kona + Ka'u. In terms of genetic connectivity the highest degree of gene flow is between Kohala and the other Windward volcanoes. Then, there is a medium level of connection between Ka'u and Kona and the lowest is between Ka'u and Windward. There is no detectable gene flow between Kohala and Kona (Eldon et al., 2013).

In summary, the coalescent reconstruction and haplotype networks of *T. brevignatha* are congruent with a scenario of population expansion on the Big Island coincidental with the Pleistocene glacial periods. The colonization of the Leeward side of the island spears to be relatively young and possibly associated with the formation of the wet forest community. A similar pattern has been shown for *D. sproati* and the spider *T. anuenue.*

**Glacier presence in other Hawaiian volcanoes**

Mauna Kea is the volcano with the most clear evidence of glacial activity on the Big Island. Mauna Loa because of its similar elevation was probably also glaciated during the Pleistocene, but Holocene volcanic activity has erased that geologic evidence (Anslow et al., 2010). It has also been estimated that Haleakalā was higher than the snowline of Mauna Kea during the last Glacial Maximum. However, there is no conclusive geomorphological evidence to support the presence of a permanent ice cap (Porter, 2005).

The strongest argument in favor of an ice cap in Haleakalā corresponds to the unexpectedly large size of the adjacent valleys to the crater (Kaupo Valley, Ke'anae Valley and Kipahulu Valley) in consideration of their drainage area. Near the shoreline of

Kaupo Valley there are large accumulations of mudflows, which exceed expectations based on the size of the drainage and rainfall. Moreover, this valley is in the dry area of Haleakalā. Porter (2005) suggested that the mudflows are due to glacial outburst floods produced by subglacier eruptions

In the case of Mauna Loa and Haleakalā it is still more difficult to have a precise idea about the effect of the glacier periods in shaping the forest communities and the connectivity between populations. However, this factor adds another layer of complexity into the dynamic nature of the landscape of the Hawaiian Islands. The island biota is not only affected by the long term changes due to the geologic evolution of the island (millions of years), but it is also under more quickly and dramatic changes associated to glacier periods (tens of thousands of years).

## CONCLUSION

The motivation of this paper was to assess the mitochondrial haplotype diversity of species in the context of the dynamic landscape of the Hawaiian Islands. Our first hypothesis predicted high genetic structure associated to landscape configuration on the Big Island. We found that the mitochondrial haplotype distribution on this landscape is not related to the volcano patterns, but rather with breaks in forest formation associated with rain regimens. The second hypothesis predicted low genetic structure in newly formed allopatric species on Maui Nui. The mitochondrial data utilized on this study did not show this pattern, but demonstrate very strong genetic isolation among the volcanoes of Maui.

Lastly, a time calibrated coalescent reconstruction of *T. brevignatha* (Maui and Big Island) provided evidence for a recent population expansion coincidental with the Pleistocene glaciation. It also suggests that the Windward/Leeward population break is related to more recent events, which are much younger than the original colonization of the island.

The dynamic landscape of Hawai'i does not only change over millions of years. More recent events, such as the Pleistocene glaciations, may also have had an important effect on shaping the biota. This complex interplay between geologic and climatic forces make it necessary to have very precise maps of the substrate ages, paleo-climatic reconstructions and genetic data sets with multiple independent markers in order to study processes of generation of biodiversity. The production of high throughput sequencing data will provide a better knowledge of the role of admixture (meta-population dynamics) and isolation in the formation of new species in this always-changing landscape.

## ACKNOWLEDGEMENTS

The authors would like to aucknowledge a big number of people and institutions whom collaborate at different stages of this research. The fieldwork in Hawai'i was supported by Laura Arnold, Timothy Bailey (HALE), David Benítez (HAVO), Katie Champlin (Limahuli Botanical garden), James Friday (UH Mānoa), Emory Griffin-Noyes (Limahuli Botanical Garden), Faith Inman-Narahari (UH Mānoa), Darcey Iwashita (UH Mānoa), Raina Kaholoaa (HALE), Susan Kennedy (UC Berkeley), Jessie Knowlton (Michigan Tech), Rick Lapoint (U Arizona), Scott Laursen (UH Mānoa), Karl Magnacca (O'ahu Army Natural Resources Program), Elizabeth Morrill (SFSU), Patrick O'Grady

(UC Berkeley), Rita Pregana (HAVO), Donald Price (UH Hilo), David Rankin (U Vermont), William Roderick (Stanford), Andrew Rominger (UC Berkeley), Karen Uy (UH Hilo), Erin Wilson (UC Riverside) and the Kīpuka team.

The permit processing and access to different reserves and private land was possible thanks to Steve Bergfeld (DOFAW Big Island), Pat Bily (TNC) Maui, Tabetha Block (HETF), Shalan Crysdale (TNC Big Island), Lance DaSilva (DOFAW Maui), Dean Danae (Kahoma Ranch), Charmian Dang (NAR), Melissa Dean (HETF), Betsy Gagne (NAR), Elizabeth Gordon (HALE), Lisa Hadway (DOFAW Big Island), Paula Hartzell (Lana'i Resorts, LLC), Greg Hendrickson (Kealakekua Ranch), Mel Johansen (TNC Big Island), Russell Kallstrom (TNC Moloka'i), Pomaika'i Kaniaupio-Crozier (Maui Land and Pinneapple), Cynthia King (DLNR), Peter Landon (NAR Maui), Rhonda Loh (HAVO), Arthir Medeiros (USGS), Joey Mello (DOFAW Big Island), Ed Misaki (TNC Moloka'i), Elliot Parsons (Pu'u Wa'awa'a HETF), Lani Petrie (Kapapala Ranch), Shawn Saito (Parker Ranch), Joe Ward (Maui Land and Pinneapple) and Kawika Winter (Limahuli Botanical Garden).

We also appreciate the advices on lab and analytical work of Michael Brewer (East Carolina University), Pete Croucher (Illumina) and Barker DNA Sequencing Facility (UC Berkeley). Jonathan Price (UH Hilo) for discussions about the age of the wet forest on Leeward Big Island and Amy Vandergast (USGS) for advices at the beginning of the project. George Roderick (UC Berkeley) and Charles Marshall (UC Berkeley) also contributed with constructive commentds on earlier versions of the manuscript.

Darko Cotoras' PhD program has been founded by a Fulbright/CONICYT fellowship and a researcher positon on a NSF Dimensions of Biodiversity Project. The fieldwork on Hawai'i was funded by Graduate Research Allocation Committee (Integrative Biology dept. UC Berkeley), Summer Research Grant (Integrative Biology dept. UC Berkeley), Walker Grant (Essig Museum of Entomology), Sigma Xi grant and Research Grant Graduate Division UC Berkeley.

## LITERATURE CITED

Anslow F.S., Clark P.U., Kurz M.D., Hostetler S.W. 2010. Geochronology and paleoclimatic implications of the last deglaciation of the Mauna Kea Ice Cap, Hawaii. Earth and Planetary Science Letters. 297: 234–248

Ballard J.W.O. and Whitlock M.C. 2004. The incomplete natural history of mitochondria. Molecular Ecology 13: 729–744

Blackledge T.A. and Gillespie R.G. 2004. Convergent evolution of behavior in an adaptive radiation of Hawaiian web-building spiders. Proceedings of the National Academy of Sciences USA 101: 16228–16233

Brower A.V.Z. 1994. Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA Evolution. Proceedings of the National Academy of Sciences of the United States of America. 91: 6491-6495

Cameron R.A.D., Triantis K.A., Parent C.E., Alonso M.R., Ibanez M., Martins A.M.F., Ladle R.L. and Whittaker R.J. 2013. Snails on oceanic islands: testing the general dynamic model of oceanic island biogeography. Journal of Biogeography 40: 117–130

Carson H.L. and Templeton A.R. 1984. Genetic revolutions in relation to speciation phenomena: the Founding of new populations. Annual Review of Ecology and Systematics. 15: 97-131

Carson H.L., Lockwood J.P. and Craddock E.M. 1990. Extinction and recolonization of local populations on a growing shield volcano. Proceedings of the National Academy of Sciences USA 87: 7055–7057

Drummond AJ, Ashton B, Buxton S, Cheung M, Cooper A, Duran C, Field M, Heled J, Kearse M, Markowitz S, Moir R, Stones-Havas S, Sturrock S, Thierer T, Wilson A (2012) Geneious v5.6, Available from http://www.geneious.com

Drummond AJ, Suchard MA, Xie D & Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7 *Molecular Biology And Evolution* **29**: 1969-1973

Eldon J., Price J.P., Magnacca K. and Price D.K. 2013. Patterns and processes in complex landscapes: testing alternative biogeographical hypotheses through integrated analysis of phylogeography and community ecology in Hawai'i. Molecular Ecology 22:3613-3628

Excoffier L., Laval G. and Schneider S. 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online 1: 47–50

Eytan R.I. and Hellberg M.E. 2010. Nuclear and mitochondrial sequence data reveal and conceal different demographic histories and population genetic processes in Caribbean reef fishes. Evolution 64: 3380–3397

Fitzpatrick B.M. 2009. Power and sample size for nested analysis of molecular variance. Molecular Ecology. 18: 3961–3966

Folmer O., Black M., Hoeh W., Lutz R. and Vrijenhoek R.1994. DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. Molecular Marine Biology and Biotechnology. 3: 294-299

Folmer O., Black M., Hoeh W., Lutz R. and Vrijenhoek R.1994. DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. Molecular Marine Biology and Biotechnology. 3: 294-299

Fu Y.X. and Li W.H. 1993. Statistical tests of neutrality of mutations. Genetics. 133: 693–709

Funk, V.A., and Wagner W.L. 1995. Biogeographic patterns in the Hawaiian Islands. In: Hawaiian bio- geography: evolution on a hot spot Archipelago. Smithsonian Institution Press, Washington, DC

Gavenda R.Y. 1992. Hawaiian Quaternary Paleoenvironments: A Review of Geological, Pedological, and Botanical Evidence. Pacific Science 46: 295-307

Gillespie R.G. 2004. Community assembly through adaptive radiation in Hawaiian spiders. Science. 303: 356-359

Gillespie R.G., Croom H.B. and Palumbi S.R. 1994. Multiple origins of a spider radiation in Hawaii. Proc. Nati. Acad. Sci. USA 91: 2290-2294

Gillespie, R.G. and Baldwin, B.G. 2010. Island Biogeography of remote archipelagos: Interplay between ecological and evolutionary processes. In: The Theory of Island Biogeography at 40: impacts and prospects. Losos JB, Ricklefs RE (eds). Princeton University Press.

Goodman K.R., Welter S.C. and Roderick G.K. 2012. Genetic divergence is decoupled from ecological diversification in the Hawaiian *Nesosydne* planthoppers. Evolution 66: 2798–2814

Hotchkiss S. and Juvik J.O. 1999. A Late-Quaternary Pollen Record from Ka‘au Crater, O‘ahu, Hawai‘i. Quaternary Research **52:** 115–128

Jacobi J.D. 1989. Vegetation Maps of the Upland Plant Communities on the Islands of Hawai’i, Maui, Moloka’i, and Lana’i. Technical Report Number 61 Cooperative National Park Resources Studies Unit, University of Hawaii, Manoa, Honolulu, Hawaii.

Katoh K., Misawa K., Kuma K. and Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 30: 3059-3066

Koressaar T. and Remm M. 2007. Enhancements and modifications of primer design program Primer3. Bioinformatics 23:1289-1291

Lanfear, R., Calcott, B., Ho, S.Y. & Guindon, S. 2012. Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Molecular Biology and Evolution. 29: 1695–1701

Lipman, P.W.,and Calvert A.T. 2011. Early growth of Kohala Volcano and formation of long Hawaiian rift zones. Geology 39: 659–662

Losos J.B. and Ricklefs R.E. 2009. Adaptation and diversification on islands. Nature 457: 830-836

Miller M.A., Pfeiffer W. and Schwartz T. 2012. “The CIPRES Science Gateway: Enabling High-impact Science For Phylogenetics Researchers With Limited Resources", XSEDE '12, Proceedings of the 1st Conference of the Extreme Science and Engineering Discovery Environment: Bridging from the eXtreme to the campus and beyond: pg: -, ACM, 978-1-4503-1602-

Pigati J.S. Zreda M. Zweck C., Almasi P.F., Elmore D and Sharp W.D. 2008. Ages and inferred causes of Late Pleistocene glaciations on Mauna Kea, Hawai’i. Journal of Quaternary Science 23: 683–702

Porter S.C. 2005. Pleistocene snowlines and glaciation of the Hawaiian Islands. Quaternary International 138–139

Price J.P. and Clague D.A. 2002. How old is the Hawaiian biota? Geology and phylogeny suggest recent divergence. Proceedings of the Royal Society of London B 269: 2429–2435

Price J.P., Gon S.M. III, Jacobi J.D. and Matsuwaki D. 2007. Mapping plant species ranges in the Hawaiian Islands: developing a methodology and associated GIS layers. HCSU Technical Report, 008, 1–58.

Rambaut A, Suchard MA, Xie D & Drummond AJ (2014) Tracer v1.6, Available from http://beast.bio.ed.ac.uk/Tracer

Roderick G.K., Croucher P.J.P., Vandergast A.G. and Gillespie R.G. 2012. Species differentiation on a dynamic Landscape: shifts in metapopulation genetic structure using the chronology of the Hawaiian Archipelago. Evolutionary Biology 39: 192–206

Salzvurger W., Ewing G.B. and Von Haeseler A. 2011. The performance of phylogenetic algorithms in estimating haplotype genealogies with migration. Molecular Ecology 20: 1952–1963

Sanderson M. 1993. Prevaling Trade Winds. Weather and Climate in Hawai’i. University of Hawai’i Press. Honolulu, USA. 126 pp.

Sanger F., Nicklen S. and Coulson A.R. 1977. DNA sequencing with chain-terminating inhibitors. Proceedings of the National Academy of Sciences USA 74: 5463-5467

Schluter D. 2000. The Ecology of Adaptive Radiation. Oxford University Press, Oxford.

Seehausen O. 2006. African cichlid fish: a model system in adaptive radiation research. Proc. R. Soc. B 273: 1987–1998

Simon C. 1987. Hawaiian evolutionary biology: an introduction. TREE 2: 175-178

Stamatakis A. 2014. RAxML Version 8: A tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies. Bioinformatics.

Steinbauer M.J., Dolos K., Field R., Reineking B. and Beierkuhnlein C. 2013. Re-evaluating the general dynamic theory of oceanic island bio- geography. Frontiers of Biogeography 5: 185-194

Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123: 585–595

Tamura K., Stecher G., Peterson D., Filipski A., and Kumar S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Molecular Biology and Evolution 30: 2725-2729

Untergrasser A., Cutcutache I., Koressaar T., Ye J., Faircloth B.C., Remm M. and Rozen S.G. 2012. Primer3 - new capabilities and interfaces. Nucleic Acids Research 40: e115

Vandergast A. G., Gillespie R. G. and G. K. Roderick. 2004. Influence of volcanic activity on the population genetic structure of Hawaiian Tetragnatha spiders: fragmentation, rapid population growth and the potential for accelerated evolution. Molecular Ecology. 13: 1729-1743

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

Whittaker R. J., Triantis K. A. Ladle R. J. 2008. A general dynamic theory of oceanic island biogeography. Journal of Biogeography. 35: 977–994

Yoder J.B, Clancey E., Desroches S., Eastman J.M, Gentry L., Godsoe W., Hagey T.J., Jochimsen D., Oswald B.P, Robertson J., Sarver B.A.J., Schenk J.J., Spear S.F. and Harmon L.J. 2010. Ecological opportunity and the origin of adaptive radiations. J. Evol Biol. 23: 1581-1596

Ziegler A. C. 2002. Hawaiian Natural History, Ecology and Evolution. Honolulu, Hawai'i. USA: University of Hawai'i Press.

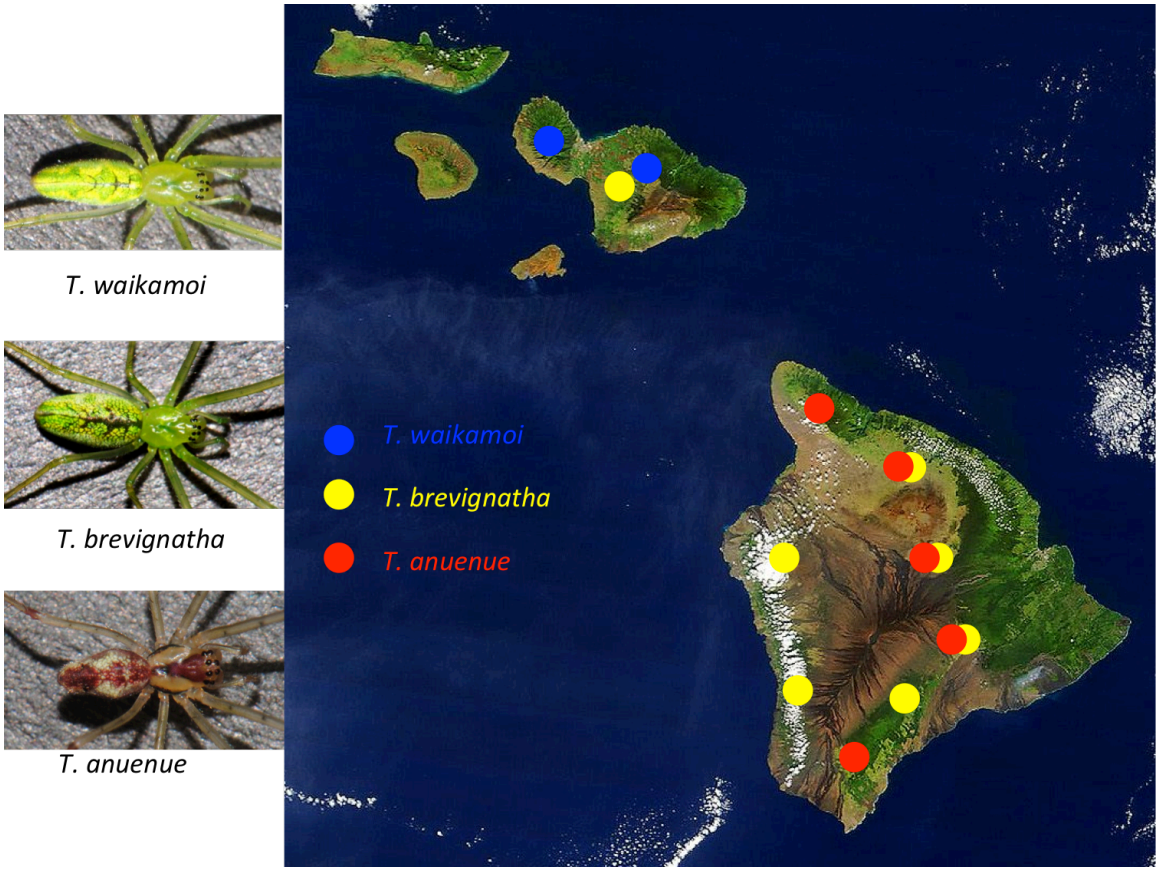



**Figure 1: Study system.** On this paper we are looking the population structure of three forest specialist *Tetragnatha* spiders. *T. waikamoi* is endemic to Maui (Pu'u Kukui (West Maui) and Upper Waikamoi (Haleakalā)). *T. anuenue* is endemic to the Big Island (Pu'u O'Umi (Kohala), Laupāhoehoe (Mauna Kea), Kīpuka in the Saddle road (Mauna Loa), Pu'u Maka'ala (Kīlauea), Ka'u (Mauna Loa)). *T. brevignatha* is present in Maui and the Big Island (Lower Waikamoi (Haleakalā), Laupāhoehoe (Mauna Kea), Kīpuka in the Saddle road (Mauna Loa), Pu'u Maka'ala (Kīlauea), Ka'u (Mauna Loa), Kona Hema (Mauna Loa) and Honoaula (Hualālai)).

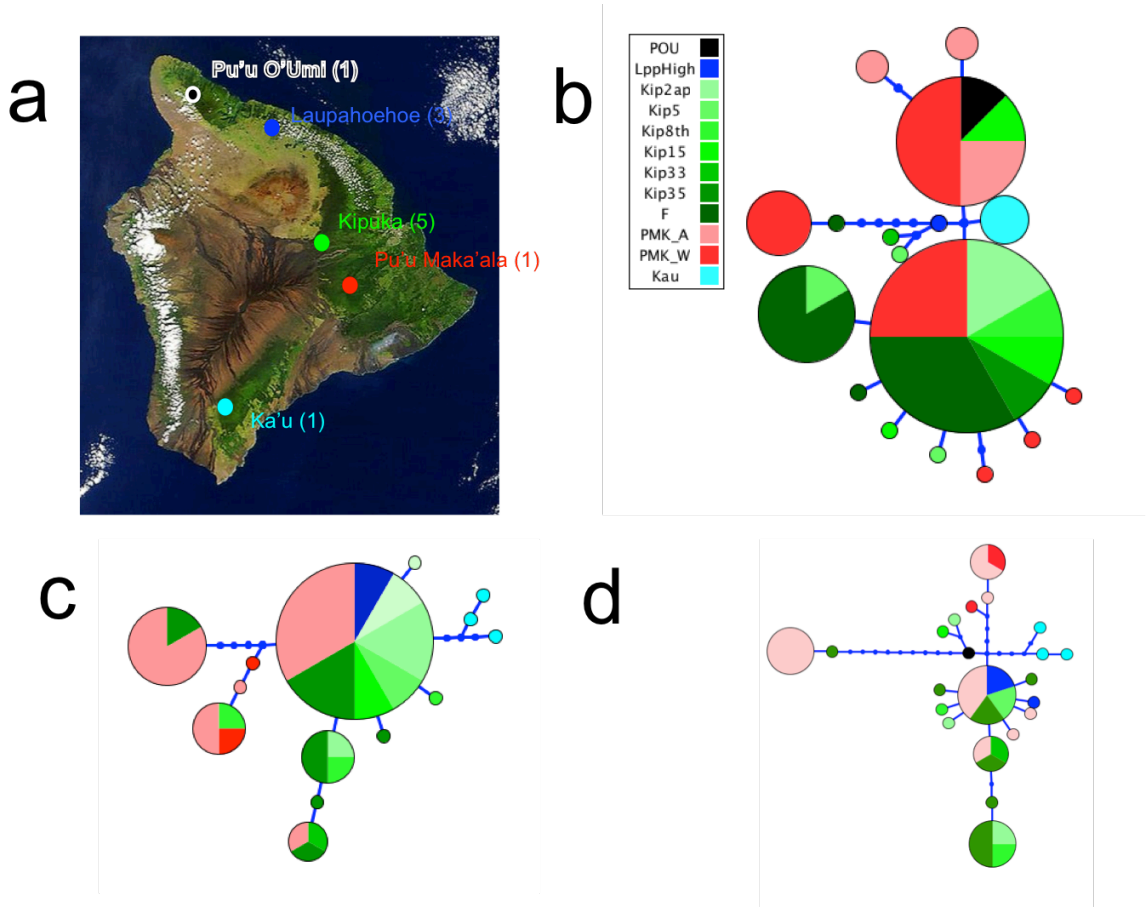



**Figure 2: Haplotype network for *T. anuenue*.** (a) Map with the sampling localities. In parenthesis the number of collecting sites, (b) COI haplotype network (47 samples, 652bp), (c) Cytb haplotype network (38 samples, 504bp) and (d) concatenated sequences haplotype network (36 samples, 1,156 bp). The colors correspond to the sampling locality and the size of the circles is proportional to the number of individuals.

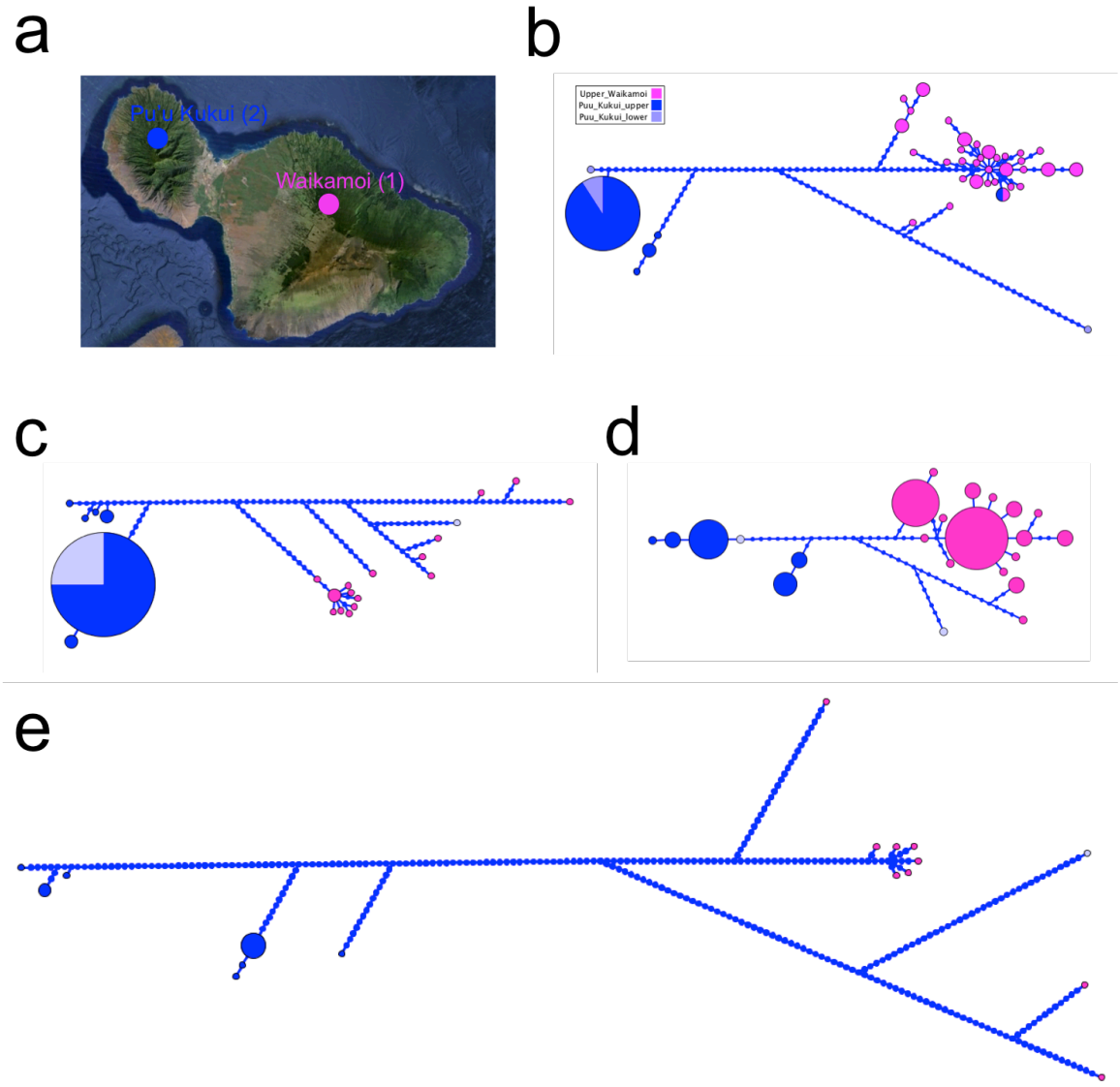



**Figure 3: Haplotype network for *T. waikamoi*.** (a) Map with the sampling localities. In parenthesis the number of collecting sites, (b) COI haplotype network (59 samples, 626bp), (c) Cytb haplotype network (41 samples, 481bp), (d) ND1 haplotype network (49 samples, 340bp) and (e) concatenated sequences haplotype network (21 samples, 1,447 bp). The colors correspond to the sampling locality and the size of the circles is proportional to the number of individuals.

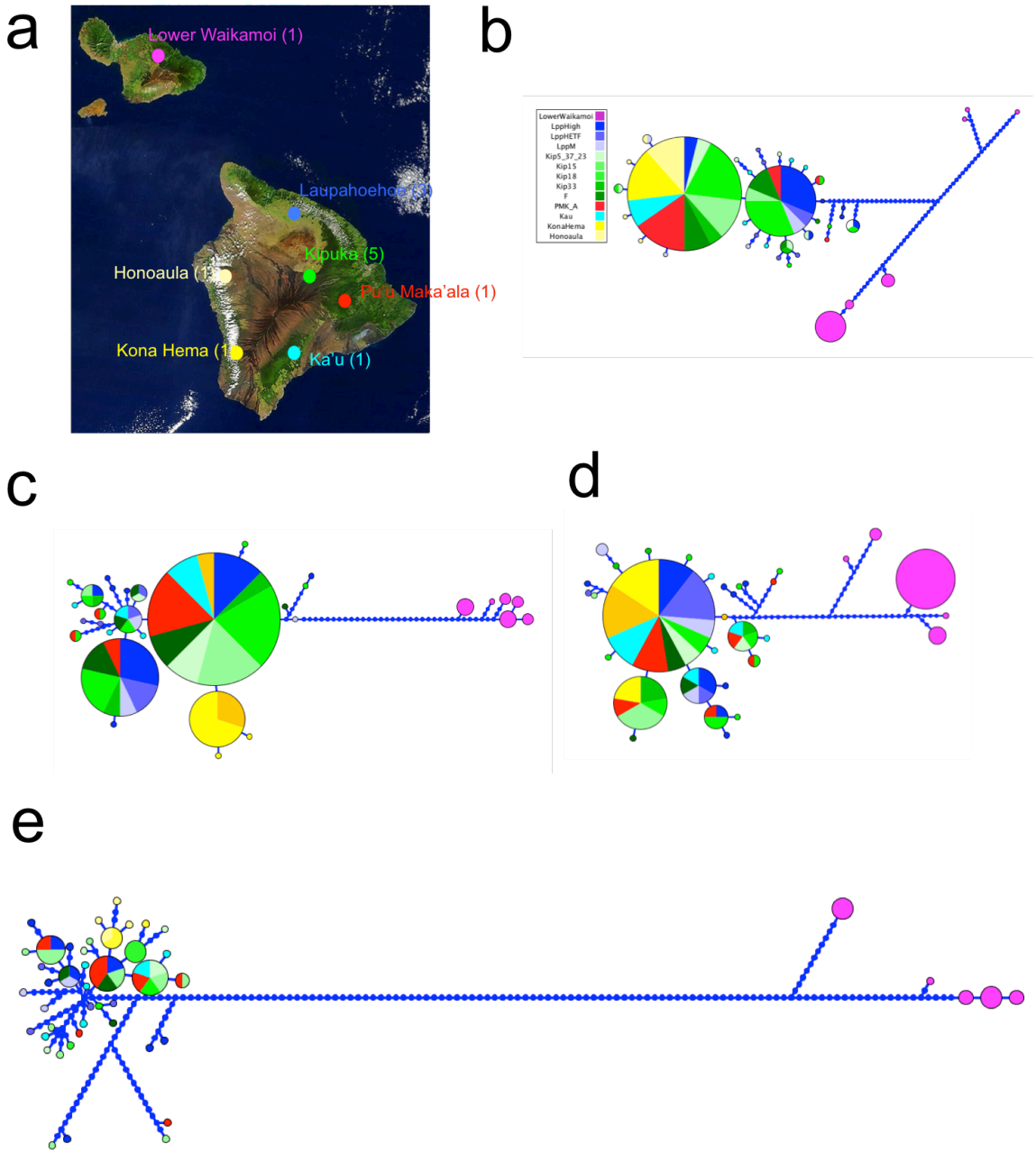
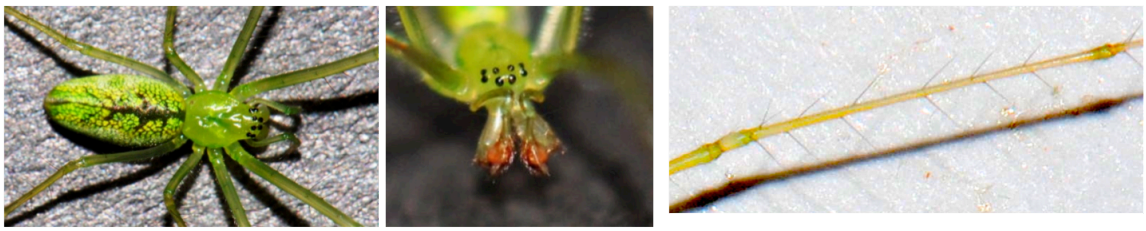



**Figure 4: Haplotype network for *T. brevignatha*.** (a) Map with the sampling localities. In parenthesis the number of collecting sites, (b) COI haplotype network (98 samples, 638bp), (c) Cytb haplotype network (95 samples, 501bp), (d) ND1 haplotype network (86 samples, 356bp) and (e) concatenated sequences haplotype network (71 samples, 1,495 bp). The colors correspond to the sampling locality and the size of the circles is proportional to the number of individuals.

*T. bravignatha:* Lower Waikamoi (Maui) and Big Island



*T. macracantha:* Kipahulu Valley (Maui), Ko'olau (Maui) and Lana'i

**Figure 5: Morphological comparison between *T. brevignatha* and *T. macracantha*.** *Left pictures*: dorsal view. *Central pictures*: Male chelicerae. *T. brevignatha* is short in comparison to *T. macracantha*. *Right pictures*: Spines in the tibial segment of the first pair of legs. The spines in *T. macracantha* can be as long as the cephalotorax, while the ones in *T. brevignatha* are very reduced.

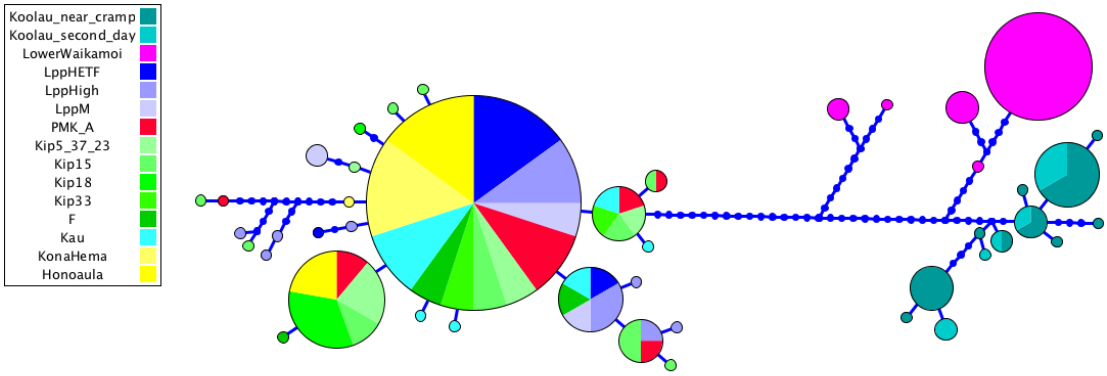



**Figure 6: Combined ND1 haplotype network of *T. brevignatha* and Ko'olau population of *T. macracantha*.** The part of the network corresponding to *T. brevignatha* is identical to the one presented in Fig. 4d. The samples corresponding to *T. macracantha* are colored turquoise and dark turquoise.

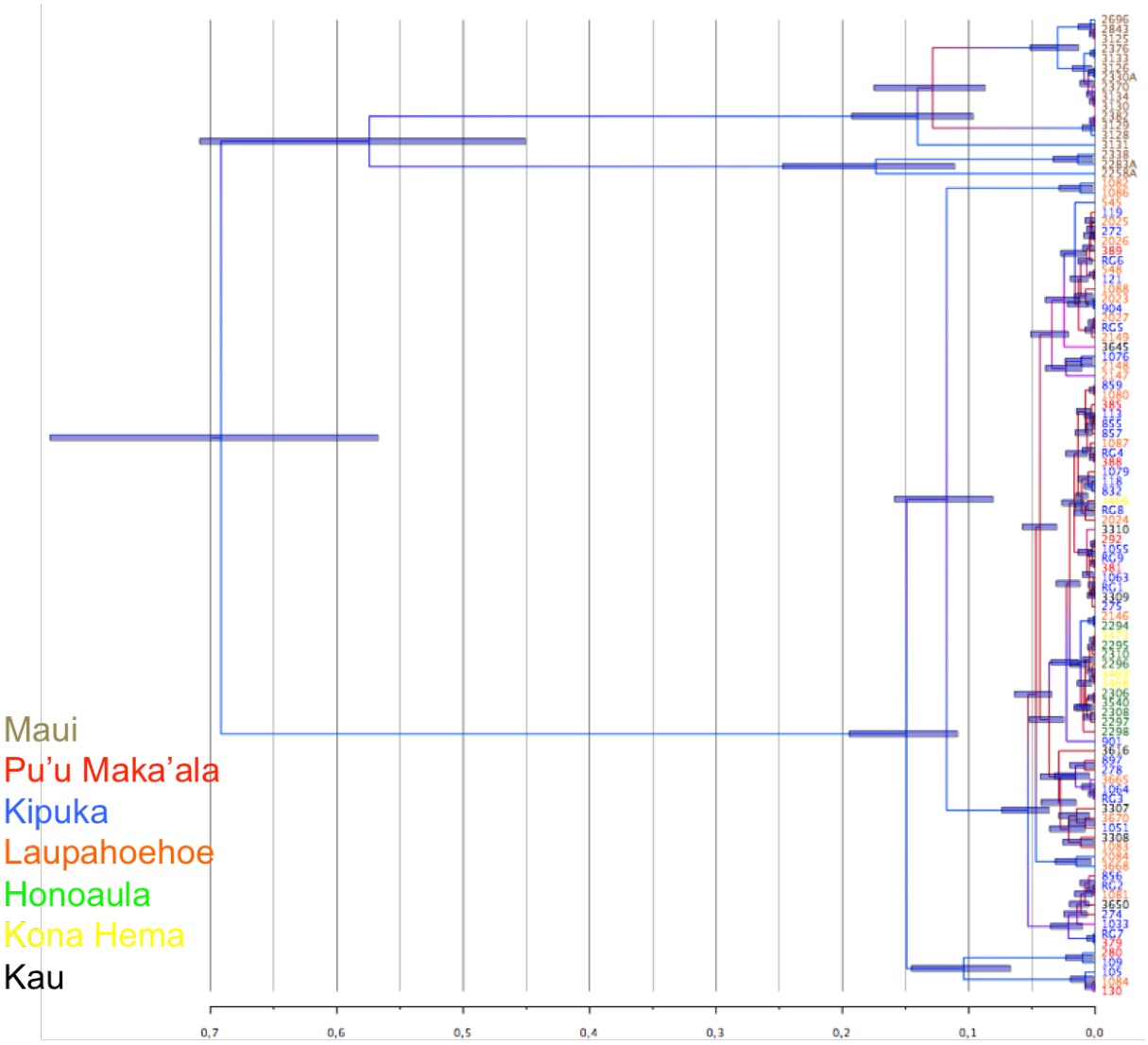



**Figure 7: Time calibrated coalescent reconstruction for *T. brevignatha* using calculated rates.** The rates correspond to COI: 13.4%; ND1: 17.2%; and Cytb: 12%. They were applied using a strict molecular clock. The geographic origin of the samples follows the color code with the locality names on the left side of the figure. The time scale is on millions of years. The color of the branches represents the posterior probability of the node ranging from blue (1.0) to red (0.5).

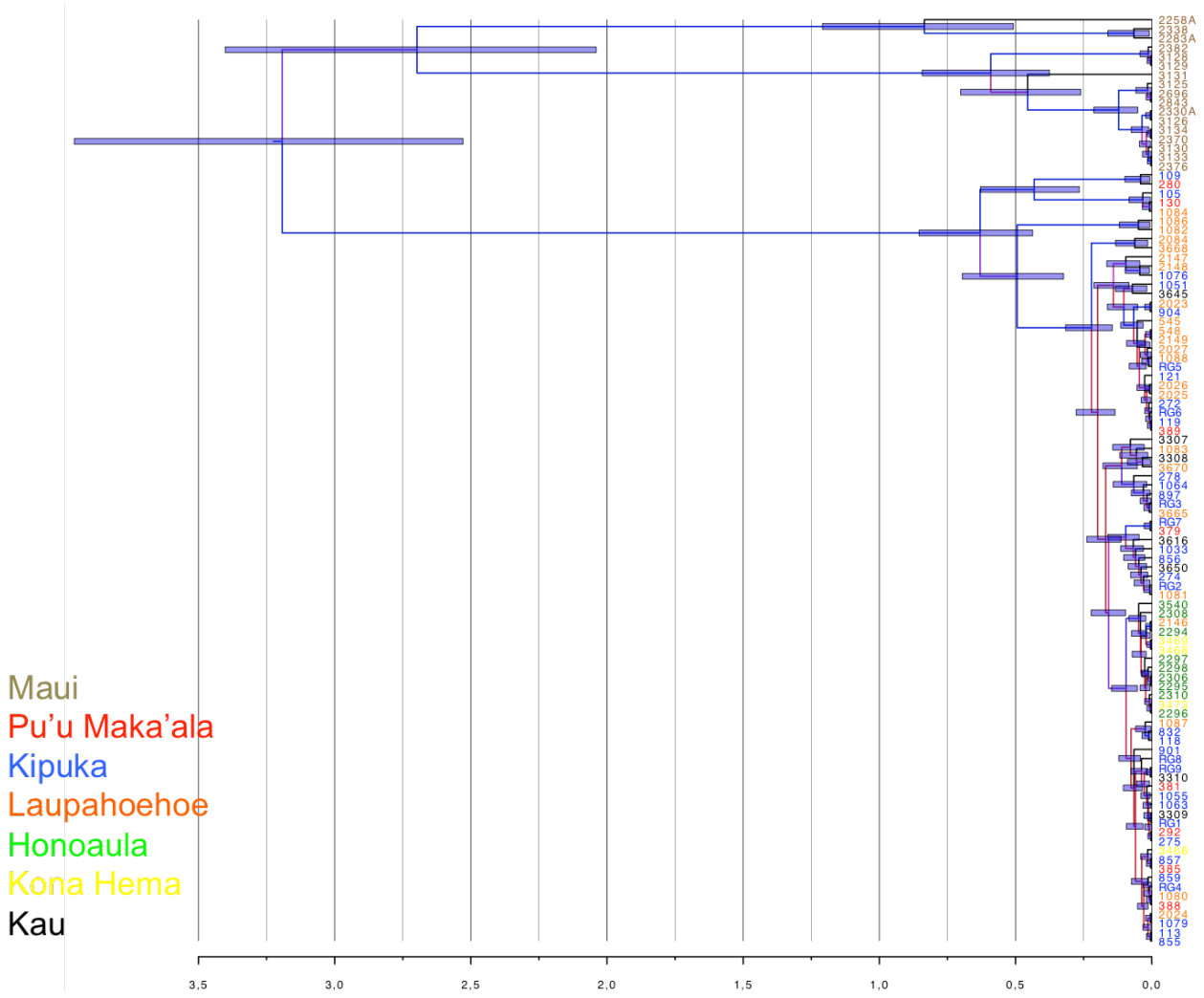



**Figure 8: Time calibrated coalescent reconstruction for *T. brevignatha* using 2.3% for COI mutation rate.** The other rates were estimated with the phylogeny and correspond to ND1: 4.74%; Cytb: 6.68%. The geographic origin of the samples follows the color code with the locality names on the left side of the figure. The time scale is on millions of years. The color of the branches represents the posterior probability of the node ranging from blue (1.0) to red (0.5).

**Table 1: Collecting sites**

| Island | Volcano | Site | GPS coordinate | Number of specimens | Specimens |
|---|---|---|---|---|---|
| *T. anuenue* | | | | | |
| Big Island | Kohala | Pu'u O'Umi NAR | N 20°40'06.15" W 155°43' 31.71" | 1 | 312 |
| | Mauna Kea | Laupāhoehoe HETF, High | N 19°54'57.62" W 155°18'22.91" | 1 | 1085 |
| | Mauna Loa | Kīpuka 15 | N 19°40'18.15" W 155°20'18.36" | 4 | 1069, RG10, RG20, RG21 |
| | | Kīpuka 8th | N 19°37'47.42" W 155°21'47.82" | 1 | 505 |
| | | Kīpuka 2' | N 19°38'56.71" W 155°21'58.40" | 2 | 564, 568 |
| | | Kīpuka 35 | N 19°38'54.73" W 155°22'20.59" | 1 | 769 |
| | | Kīpuka 5 | N 19°39'50.02" W155°20'51.65" | 3 | 896, 1031, 1034 |
| | | Kīpuka 33 | N 19°38'35.08" W 155°20'34.56" | 1 | 1062 |
| | | Forest (f2) | N 19°39'57.39" W 155°21'09.91" | 10 | 1052, 1071, 1078, RG12, RG13, RG14, RG16, RG17, RG19, RG23 |
| | | Ka'u FR, Kahuku access | N 19°40'42.50" W 155°40'42.51" | 4 | 3416, 3424, 3428, 3231, |
| | Kīlauea | Pu'u Maka'ala NAR, Army Road | N 19°33'03.36" W 155°13'51.54" | 6 | 284, 373, 377,380, 387, 3679 |
| | | Pu'u Maka'ala NAR, Wright Road | N 19°29'10.42" W 155°16'13.21" | 15 | 403, 427, 430, 431, 468, 469, 470, 473, 476, 479, 485, 490, 491, 494, 495 |

| | | | | | |
|---|---|---|---|---|---|
| *T. waikamoi* | | | | | |
| Maui | Haleakalā | Upper Waikamoi TNC | N 20°46'41.60'' W 156°13'43.21'' | 47 | 2335, 2385, 2632, 2633, 2634, 2635, 2636, 2637, 2638, 2639, 2640, 2641, 2642, 2643, 2644, 2645, 2646, 2647, 2648, 2649, 2650, 2651, 2652, 2653, 2654, 2741, 2742, 2743, 2854, 2855, 2856, 2857, 3280, 3282, 3283, 3284, 3285, 3286, 3618, 3620, 3622, 3624, 3627, 3631, 2308a, 2310a, 2311a |
| | West Maui | Pu'u Kukui upper | N 20°54'52.04" W 156°35'31.84" | 19 | 2765, 2766, 2767, 2768, 2769, 2770, 2771, 2772, 2773, 2774, 2775, 2776, 2777, 2778, 2779, 2780, 2781, 2782, 2783, |
| | | Pu'u Kukui lower | N 20°54'56.73" W 156°35'35.51" | 5 | 2818, 2820, 2823, 2829, 2834, |
| *T. brevignatha* | | | | | |
| Maui | Haleakalā | Lower Waikamoi TNC | N 20°48'24.22'' W 156°15'18.02'' | 15 | 2338, 2370, 2376, 2382, 2696, 2283a, 2330a, 3125, 3126, 3128, 3129, 3130, 3131, 3133, 3134, |
| Big Island | Hualālai | Honoaula FR, Makahi St. entry | N 19°43'07.96'' W 155°56'57.34'' | 6 | 2294, 2295, 2296, 2297, 2298, 2310, |
| | Mauna Kea | Laupāhoehoe HETF, HIPPNET | N 19°55'50.43" W 155°17'20.48" | 3 | 545, 548, 2084 |
| | | Laupāhoehoe HETF, High | N 19°54'57.62" W 155°18'22.91" | 14 | 1080, 1081, 1082, 1083, 1084, 1086, 1088, 2023, 2025, 2026, 2027, 3665, 3668, 3670 |
| | | Laupāhoehoe HETF, Maulua trail. Plot 32, Transect 31 | N 19°53'54.53" W 155°18'46.35" | 4 | 2146, 2147, 2148, 2149 |
| | Mauna Loa | Kīpuka 15 | N 19°40'18.15" | 13 | 105, 109, 113, 118, 119, 121, 272, 274, |

| | | | | | |
|---|---|---|---|---|---|
| | | | W 155°20’18.36” | | 275, 278, RG1 RG7, RG8, RG9 |
| | | Kīpuka 18 | N 19°40’27.03” W 155°19’59.45” | 4 | 855, 856, 857, 859 |
| | | Kīpuka 5+37+23 | N 19°39’50.19” W 155°21’09.72” | 6 | 832, 897, 901, 904, 1033, 1055 |
| | | Kīpuka 33 | N 19°38’35.08” W 155°20’34.56” | 2 | 1063, 1064 |
| | | Forest (f2) | N 19°39’57.39” W 155°21’09.91” | 6 | 1051, 1079, RG3, RG4, RG5, RG6, |
| | | Ka’u FR, Kapapala access | N 19°20’41.82” W 155°28’03.14” | 7 | 3307, 3308, 3309, 3310, 3616, 3645, 3650, |
| | | Kona Hema TNC | N 19°12’49.41” W 155°49’44.48” | 4 | 3466, 3468, 3469, 3473 |
| | Kīlauea | Pu’u Maka’ala NAR, Army Road | N 19°33'03.36" W 155°13'51.54" | 10 | 280, 292, 379, 381, 385, 381, 385, 388, 389, RG2 |
| | | ‘Ōla’a Unit, HAVO | N 19°27’37.26” W 155°14’50.33” | 1 | 130 |
| *T. macracantha* | | | | | |
| Maui | Haleakalā | Ko’olau FR | N 20°45'39.88" W 156°08'35.35" | 25 | 3147, 3154, 3157, 3158, 3163, 3164, 3165, 3166, 3167, 3168, 3170, 3171, 3172, 3175, 3180, 3181, 3182, 3184, 3193, 3194, 3199, 3203, 3217, 3226, 3233, 3237 |

TNC: The Nature Conservancy; NAR: Natural Area Reserve; FR: Forest; HETF: Hawai’i Experimental Tropical Forest; HAVO: Hawai’i Volcanoes National Park

**Table 2: Amplification primers**

| Gene | Primer name | Sequence | Species | Annealing temperature (°C) | Source |
|---|---|---|---|---|---|
| COI | HCO | TAAACTTCAGGGTGACCAAAAAATCA | *T. anuenue* | 42 | Folmer et al., 1994 |
| | LCO | GGTCAACAAATCATAAAGATATTGG | *T. brevignatha* and *T. waikamoi* | 52 | |
| Cytb | Ctb_F | TGCCTTGGGGCCAAATATC | *T. brevignatha* | 60 | This study |
| | Ctb_R | GCTCTTCGAAAACGAAAAGACAAC | *T. waikamoi* and *T anuenue* | 50 | |
| ND1 | ND1F | AGATAGCGGTAGGTACACCAAGG | *T. brevignatha* | 60 | This study |
| | ND1R | ATATCGGGGCCACACGAAG | *T. waikamoi* | 40 | |

**Table 3: Detail of the markers amplified on each specimen**

*T. anuenue*

| Sample | COI | Cytb_tRIle |
|---|---|---|
| 284 | 1 | 0 |
| 312 | 1 | 0 |
| 373 | 1 | 0 |
| 377 | 1 | 0 |
| 380 | 1 | 0 |
| 387 | 1 | 1 |
| 403 | 1 | 1 |
| 427 | 1 | 1 |
| 430 | 1 | 1 |
| 431 | 1 | 1 |
| 468 | 1 | 0 |
| 469 | 1 | 0 |
| 470 | 1 | 1 |
| 473 | 1 | 1 |
| 476 | 1 | 1 |
| 479 | 1 | 1 |
| 485 | 1 | 1 |
| 490 | 0 | 1 |
| 491 | 1 | 1 |
| 494 | 1 | 1 |
| 495 | 1 | 1 |
| 505 | 1 | 1 |
| 564 | 1 | 1 |
| 568 | 1 | 1 |
| 769 | 1 | 1 |
| 896 | 1 | 1 |
| 1031 | 1 | 1 |
| 1034 | 1 | 1 |
| 1052 | 1 | 1 |
| 1062 | 1 | 1 |
| 1069 | 1 | 0 |
| 1071 | 1 | 1 |
| 1078 | 1 | 1 |
| 1085 | 1 | 1 |
| 3416 | 1 | 0 |
| 3424 | 1 | 1 |
| 3428 | 1 | 1 |
| 3431 | 1 | 1 |
| 3679 | 1 | 1 |
| RG10 | 1 | 1 |

| | | |
|---|---|---|
| RG12 | 1 | 1 |
| RG13 | 1 | 0 |
| RG14 | 1 | 0 |
| RG16 | 1 | 1 |
| RG17 | 1 | 1 |
| RG19 | 1 | 1 |
| RG20 | 1 | 0 |
| RG21 | 1 | 1 |
| RG23 | 1 | 1 |

*T. waikamoi*

| Sample | COI | ND1 | Cytb_tRIle |
|---|---|---|---|
| 2335 | 0 | 1 | 1 |
| 2385 | 1 | 1 | 0 |
| 2632 | 1 | 0 | 1 |
| 2633 | 0 | 1 | 1 |
| 2634 | 1 | 0 | 0 |
| 2635 | 1 | 1 | 0 |
| 2636 | 0 | 0 | 1 |
| 2637 | 1 | 1 | 0 |
| 2638 | 1 | 1 | 1 |
| 2639 | 1 | 1 | 1 |
| 2640 | 1 | 0 | 0 |
| 2641 | 1 | 0 | 0 |
| 2642 | 1 | 1 | 0 |
| 2643 | 1 | 1 | 0 |
| 2644 | 1 | 1 | 0 |
| 2645 | 1 | 1 | 1 |
| 2646 | 1 | 1 | 1 |
| 2647 | 1 | 1 | 1 |
| 2648 | 1 | 1 | 1 |
| 2649 | 1 | 1 | 0 |
| 2650 | 1 | 0 | 1 |
| 2651 | 0 | 1 | 0 |
| 2652 | 1 | 1 | 0 |
| 2653 | 1 | 1 | 1 |
| 2654 | 1 | 1 | 0 |
| 2741 | 0 | 1 | 1 |
| 2742 | 1 | 1 | 0 |
| 2743 | 1 | 1 | 1 |
| 2765 | 1 | 1 | 1 |
| 2766 | 1 | 0 | 1 |
| 2767 | 1 | 1 | 1 |
| 2768 | 1 | 0 | 1 |
| 2769 | 1 | 0 | 1 |

| | | | |
|---|---|---|---|
| 2770 | 1 | 0 | 1 |
| 2771 | 1 | 1 | 1 |
| 2772 | 1 | 1 | 1 |
| 2773 | 0 | 0 | 1 |
| 2774 | 0 | 1 | 1 |
| 2775 | 0 | 0 | 1 |
| 2776 | 1 | 1 | 1 |
| 2777 | 1 | 1 | 1 |
| 2778 | 0 | 1 | 1 |
| 2779 | 1 | 1 | 1 |
| 2780 | 1 | 1 | 1 |
| 2781 | 1 | 1 | 1 |
| 2782 | 1 | 1 | 1 |
| 2783 | 1 | 1 | 1 |
| 2818 | 1 | 0 | 1 |
| 2820 | 1 | 1 | 1 |
| 2823 | 1 | 0 | 1 |
| 2829 | 0 | 1 | 1 |
| 2834 | 0 | 0 | 1 |
| 2854 | 1 | 1 | 0 |
| 2855 | 0 | 1 | 0 |
| 2856 | 1 | 1 | 0 |
| 2857 | 1 | 0 | 0 |
| 3280 | 1 | 0 | 0 |
| 3282 | 1 | 0 | 0 |
| 3283 | 1 | 0 | 0 |
| 3284 | 1 | 0 | 0 |
| 3285 | 1 | 1 | 0 |
| 3286 | 1 | 0 | 0 |
| 3618 | 0 | 1 | 0 |
| 3620 | 1 | 0 | 0 |
| 3621 | 1 | 0 | 0 |
| 3622 | 1 | 0 | 0 |
| 3624 | 1 | 1 | 0 |
| 3627 | 1 | 0 | 0 |
| 3631 | 1 | 1 | 0 |
| 2308A | 1 | 1 | 0 |
| 2310A | 1 | 1 | 1 |
| 2311A | 1 | 1 | 1 |

*T, brevignatha*

| Sample | COI | ND1 | Cytb_tRIle |
|---|---|---|---|
| 105 | 1 | 1 | 1 |
| 109 | 1 | 1 | 1 |
| 113 | 1 | 1 | 1 |
| 118 | 1 | 1 | 1 |
| 119 | 1 | 1 | 1 |
| 121 | 1 | 1 | 1 |
| 130 | 1 | 0 | 0 |
| 272 | 1 | 1 | 1 |
| 274 | 1 | 1 | 1 |
| 275 | 1 | 1 | 1 |
| 278 | 1 | | 1 |
| 280 | 1 | 1 | 1 |
| 292 | 1 | 1 | 1 |
| 379 | 1 | 1 | 1 |
| 381 | 1 | 1 | 1 |
| 385 | 1 | 1 | 1 |
| 388 | 1 | 1 | 1 |
| 389 | 1 | 1 | 1 |
| 545 | 1 | 1 | 1 |
| 548 | 1 | 0 | 1 |
| 832 | 1 | 1 | 1 |
| 855 | 1 | 1 | 1 |
| 856 | 1 | 1 | 1 |
| 857 | 1 | 1 | 1 |
| 859 | 1 | 1 | 1 |
| 897 | 1 | 1 | 0 |
| 901 | 1 | 1 | 1 |
| 904 | 1 | 0 | 1 |
| 1033 | 1 | 1 | 1 |
| 1051 | 1 | 0 | 1 |
| 1055 | 1 | 1 | 1 |
| 1063 | 1 | 1 | 1 |
| 1064 | 1 | 1 | 1 |
| 1076 | 0 | 0 | 1 |
| 1079 | 1 | 0 | 1 |
| 1080 | 1 | 1 | 1 |
| 1081 | 1 | 0 | 1 |
| 1082 | 1 | 1 | 1 |
| 1083 | 1 | 0 | 1 |
| 1084 | 1 | 1 | 1 |
| 1086 | 1 | 1 | 1 |
| 1087 | 0 | 1 | 1 |
| 1088 | 1 | 1 | 1 |

| | | | |
|---|---|---|---|
| 2023 | 1 | 1 | 1 |
| 2024 | 0 | 1 | 1 |
| 2025 | 1 | 1 | 1 |
| 2026 | 1 | 1 | 1 |
| 2027 | 1 | 1 | 1 |
| 2084 | 1 | 1 | 1 |
| 2146 | 1 | 1 | 0 |
| 2147 | 1 | 1 | 1 |
| 2148 | 1 | 1 | 1 |
| 2149 | 1 | 1 | 1 |
| 2294 | 1 | 0 | 1 |
| 2295 | 1 | 1 | 1 |
| 2296 | 1 | 1 | 1 |
| 2297 | 1 | 1 | 1 |
| 2298 | 1 | 0 | 1 |
| 2306 | 0 | 1 | 1 |
| 2308 | 0 | 0 | 1 |
| 2310 | 1 | 0 | 1 |
| 2338 | 1 | 1 | 0 |
| 2370 | 1 | 1 | 1 |
| 2376 | 1 | 1 | 1 |
| 2382 | 1 | 1 | 1 |
| 2696 | 1 | 1 | 1 |
| 2843 | 0 | 1 | 1 |
| 2283A | 1 | 1 | 0 |
| 2330A | 1 | 1 | 1 |
| RG1 | 1 | 1 | 1 |
| RG2 | 1 | 1 | 1 |
| RG3 | 1 | 1 | 1 |
| RG4 | 1 | 1 | 1 |
| RG5 | 1 | 1 | 1 |
| RG6 | 1 | 0 | 1 |
| RG7 | 1 | 0 | 1 |
| RG8 | 1 | 0 | 1 |
| RG9 | 1 | 0 | 1 |
| 3125 | 1 | 1 | 0 |
| 3126 | 1 | 1 | 1 |
| 3128 | 1 | 1 | 1 |
| 3129 | 1 | 1 | 1 |
| 3130 | 1 | 1 | 1 |
| 3131 | 0 | 1 | 1 |
| 3133 | 1 | 1 | 1 |
| 3134 | 1 | 1 | 1 |
| 3307 | 0 | 1 | 1 |
| 3308 | 1 | 1 | 1 |
| 3309 | 1 | 1 | 1 |

| | | | |
|---|---|---|---|
| 3310 | 1 | 1 | 1 |
| 3466 | 1 | 1 | 1 |
| 3468 | 1 | 1 | 1 |
| 3469 | 1 | 1 | 1 |
| 3473 | 1 | 1 | 1 |
| 3540 | 0 | 1 | 1 |
| 3616 | 1 | 1 | 0 |
| 3645 | 1 | 1 | 1 |
| 3650 | 1 | 1 | 1 |
| 3665 | 1 | 1 | 1 |
| 3668 | 1 | 1 | 1 |
| 3670 | 1 | 1 | 1 |

*T. macracantha*

| Sample | COI | ND1 | Cytb_tRIle |
|---|---|---|---|
| 3147 | 1 | 1 | 1 |
| 3154 | 1 | 1 | 1 |
| 3157 | 1 | 1 | 1 |
| 3158 | 1 | 1 | 1 |
| 3163 | 1 | 1 | 1 |
| 3164 | 1 | 1 | 1 |
| 3165 | 1 | 1 | 1 |
| 3166 | 1 | 1 | 1 |
| 3167 | 1 | 1 | 1 |
| 3168 | 1 | 1 | 1 |
| 3170 | 1 | 1 | 1 |
| 3171 | 1 | 1 | 1 |
| 3172 | 1 | 1 | 1 |
| 3175 | 1 | 1 | 1 |
| 3180 | 1 | 1 | 1 |
| 3181 | 1 | 1 | 0 |
| 3182 | 1 | 1 | 1 |
| 3184 | 1 | 1 | 1 |
| 3193 | 1 | 1 | 1 |
| 3194 | 1 | 0 | 0 |
| 3199 | 0 | 1 | 1 |
| 3203 | 1 | 1 | 1 |
| 3217 | 1 | 1 | 1 |
| 3226 | 1 | 1 | 1 |
| 3233 | 1 | 1 | 1 |
| 3237 | 1 | 1 | 1 |

*0: not sequenced*
*1: successfully sequenced*

**Table 4: Summary of the markers amplified**

| Species | Gene | Specimens | Nucleotides (bp) |
|---|---|---|---|
| *T. anuenue* | COI | 48 | 652 |
| | CytB | 37 | 504 |
| | Concatenated | 36 | 1,156 |
| *T. waikamoi* | COI | 59 | 626 |
| | CytB | 40 | 481 |
| | ND1 | 47 | 340 |
| | Concatenated | 21 | 1,447 |
| *T. brevignatha* | COI | 92 | 638 |
| | CytB | 94 | 501 |
| | ND1 | 84 | 356 |
| | Concatenated | 71 | 1,495 |
| *T. macracantha* | ND1 | 25 | 340 |

**Table 5: Molecular diversity indexes for COI in *T. anuenue***

| Site | n | H | Gene Diversity | S | Nucleotide diversity (average over loci) |
|---|---|---|---|---|---|
| **Kohala (Pu'u O'Umi)** | **1** | **1** | **1.0000 +/- 0.0000** | **0** | **0.000000 +/- 0.000000** |
| **Kīlauea** | **20** | **8** | **0.8632 +/- 0.0486** | **16** | **0.006587 +/- 0.003802** |
| PMK_A | 6 | 3 | 0.8000 +/- 0.1217 | 3 | 0.002454 +/- 0.001940 |
| PMK_W | 14 | 6 | 0.8352 +/- 0.0617 | 13 | 0.007079 +/- 0.004148 |
| **Mauna Loa*** | **22** | **10** | **0.8139 +/- 0.0640** | **14** | **0.003426 +/- 0.002189** |
| Kip15 | 4 | 4 | 1.0000 +/- 0.1768 | 4 | 0.003067 +/- 0.002573 |
| Kip8th | 1 | 1 | 1.0000 +/- 0.0000 | 0 | 0.000000 +/- 0.000000 |
| Kip2' | 2 | 1 | 0.0000 +/- 0.0000 | 0 | 0.000000 +/- 0.000000 |
| Kip35 | 1 | 1 | 1.0000 +/- 0.0000 | 0 | 0.000000 +/- 0.000000 |
| Kip5 | 3 | 3 | 1.0000 +/- 0.2722 | 6 | 0.006135 +/- 0.005214 |
| Kip33 | 1 | 1 | 1.0000 +/- 0.0000 | 0 | 0.000000 +/- 0.000000 |
| F | 10 | 4 | 0.7333 +/- 0.1005 | 7 | 0.002931 +/- 0.002056 |
| Ka'u_Kahuku | 3 | 1 | 0.0000 +/- 0.0000 | 0 | 0.000000 +/- 0.000000 |
| **Mauna Kea (Lpp_H)** | **1** | **1** | **1.0000 +/- 0.0000** | **0** | **0.000000 +/- 0.000000** |

*In bold the aggregated data for all the sites in the volcano.*
** Mauna Loa does not include Ka'u site.*

**Table 6: Molecular diversity indexes for Cytb in *T. anuenue***

| Site | n | H | Gene Diversity | S | Nucleotide diversity (average over loci) |
|---|---|---|---|---|---|
| **Kohala (Pu'u O'Umi)** | **-** | **-** | **-** | **-** | **-** |
| **Kīlauea** | **15** | **6** | **0.8190 +/- 0.0636** | **8** | **0.007332 +/- 0.004401** |
| PMK_A | 2 | 2 | 1.0000 +/- 0.5000 | 2 | 0.003968 +/- 0.004860 |
| PMK_W | 13 | 5 | 0.7821 +/- 0.0794 | 8 | 0.007377 +/- 0.004476 |
| **Mauna Loa** | **19** | **9** | **0.8830 +/- 0.0461** | **12** | **0.004989 +/- 0.003140** |
| Kip15 | 3 | 3 | 1.0000 +/- 0.2722 | 6 | 0.007937 +/- 0.006745 |
| Kip8th | 1 | 1 | 1.0000 +/- 0.0000 | 0 | 0.000000 +/- 0.000000 |
| Kip2' | 2 | 2 | 1.0000 +/- 0.5000 | 1 | 0.001984 +/- 0.002806 |
| Kip35 | 1 | 1 | 1.0000 +/- 0.0000 | 0 | 0.000000 +/- 0.000000 |
| Kip5 | 3 | 2 | 0.6667 +/- 0.3143 | 5 | 0.006614 +/- 0.005749 |
| Kip33 | 1 | 1 | 1.0000 +/- 0.0000 | 0 | 0.000000 +/- 0.000000 |
| F | 8 | 6 | 0.9286 +/- 0.0844 | 8 | 0.004889 +/- 0.003358 |
| Ka'u_Kahuku | 3 | 3 | 1.0000 +/- 0.2722 | 4 | 0.005291 +/- 0.004749 |
| **Mauna Kea (Lpp_H)** | **1** | **1** | **1.0000 +/- 0.0000** | **0** | **0.000000 +/- 0.000000** |

*In bold the aggregated data for all the sites in the volcano.*
** Mauna Loa does not include Ka'u site.*

**Table 7: Molecular diversity indexes for the concatenated sequences in *T. anuenue***

| Site | n | H | Gene Diversity | S | Nucleotide diversity (average over loci) |
|---|---|---|---|---|---|
| **Kohala (Pu'u O'Umi)** | **-** | **-** | **-** | **-** | **-** |
| **Kīlauea** | **14** | **8** | **0.8901 +/- 0.0603** | **20** | **0.007025 +/- 0.003895** |
| PMK_A | 2 | 2 | 1.0000 +/- 0.5000 | 3 | 0.002595 +/- 0.002997 |
| PMK_W | 12 | 7 | 0.8788 +/- 0.0751 | 19 | 0.007235 +/- 0.004059 |
| **Mauna Loa** | **18** | **12** | **0.9346 +/- 0.0409** | **21** | **0.003245 +/- 0.001920** |
| Kip15 | 2 | 2 | 1.0000 +/- 0.5000 | 3 | 0.002595 +/- 0.002997 |
| Kip8th | 1 | 1 | 1.0000 +/- 0.0000 | 0 | 0.000000 +/- 0.000000 |
| Kip2' | 2 | 2 | 1.0000 +/- 0.5000 | 1 | 0.000865 +/- 0.001223 |
| Kip35 | 1 | 1 | 1.0000 +/- 0.0000 | 0 | 0.000000 +/- 0.000000 |
| Kip5 | 3 | 3 | 1.0000 +/- 0.2722 | 7 | 0.004037 +/- 0.003374 |
| Kip33 | 1 | 1 | 1.0000 +/- 0.0000 | 0 | 0.000000 +/- 0.000000 |
| F | 8 | 7 | 0.9643 +/- 0.0772 | 15 | 0.004078 +/- 0.002542 |
| **Mauna Kea (Lpp_H)** | **1** | **1** | **1.0000 +/- 0.0000** | **0** | **0.000000 +/- 0.000000** |
| Ka'u_Kahuku | 3 | 3 | 1.0000 +/- 0.2722 | 4 | 0.002307 +/- 0.002070 |

*In bold the aggregated data for all the sites in the volcano.*
** Mauna Loa does not include Ka'u site.*

**Table 8: Analysis of molecular variance (AMOVA) for *T. anuenue***

| | COI | | Cytb | | Concatenated | |
|---|---|---|---|---|---|---|
| **Source of variation (%)** | **Per Volcano** | **Wind./Lee.** | **Per Volcano** | **Wind./Lee.** | **Per Volcano** | **Wind./Lee.** |
| **Among groups** | 8.96 | 16.50 | 25.98 | 43.74 | 29.27 | 36.76 |
| **Among populations within groups** | 10.96 | 12.94 | -6.54 | 0.63 | -10.82 | 1.06 |
| **Among populations** | 80.08 | 70.56 | 80.56 | 55.63 | 81.56 | 62.18 |

*The definition of the groups is as it follows:*
**Per Volcano:** Kohala (Pu'u O'Umi), Kīlauea (Pu'u Maka'ala, Army Road + Pu'u Maka'ala, Wright Road), Mauna Loa (Kīpuka 15 + Kīpuka 8$^{th}$ + Kīpuka 2 + Kīpuka 35 + Kīpuka 5 + Kīpuka 33 + Forest), Mauna Kea (Laupāhoehoe, High), Southwest Mauna Loa (Ka'u, Kahuku Unit).
**Wind./Lee.:** Windward (Pu'u O'Umi + Pu'u Maka'ala, Army Road + Pu'u Maka'ala, Wright Road + Kīpuka 15 + Kīpuka 8$^{th}$ + Kīpuka 2 + Kīpuka 35 + Kīpuka 5 + Kīpuka 33 + Forest + Laupāhoehoe, High), Leeward (Ka'u, Kahuku Unit).

**Table 9: Par-wise $\phi_{ST}$ estimations for *T. anuenue* based on COI.**

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| 1 | 0.00000 | | | | |
| 2 | -0.59064 | 0.00000 | | | |
| 3 | 0.18095 | 0.12926 | 0.00000 | | |
| 4 | 1.0000 | -0.30144 | 0.12245 | 0.00000 | |
| 5 | 1.0000 | 0.19064 | 0.38465 | 1.0000 | 0.00000 |

1: Kohala
2: Kilauea
3: MaunaLoa
4: MaunaKea
5: Kau_Kahuku

*Highlighted significant values (P<0.05).*

**Table 10: Par-wise $\phi_{ST}$ estimations for *T. anuenue* based on Cytb.**

| | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1 | 0.00000 | | | |
| 2 | 0.13301 | 0.00000 | | |
| 3 | -0.28904 | -0.70635 | 0.00000 | |
| 4 | 0.45923 | 0.48280 | 0.27273 | 0.00000 |

1: Kilauea
2: MaunaLoa
3: MaunaKea
4: Kau_Kahuku

*Highlighted significant values (P<0.05).*

**Table 11: Neutrality tests for populations of *T. anuenue***

| Site | COI | | Cytb | | Concatenated | |
|---|---|---|---|---|---|---|
| | **Tajima's D** | **Fu's Fs** | **Tajima's D** | **Fu's Fs** | **Tajima's D** | **Fu's Fs** |
| BIG ISLAND | -1.06226 (P= 0.14210) | -5.57181 (P= 0.01820) | -0.40768 (P= 0.38280) | -3.26474 (P= 0.09100) | -0.73457 (P= 0.25670) | -7.69862 (P= 0.00920) |
| Pu'u O'Umi | 0.00000 (P= 1.00000) | na | - | - | - | - |
| Kīlauea | -0.17733 (P=0.47660) | 0.27016 (P=0.57510) | 1.84831 (P=0.97930) | 0.83820 (P=0.68110) | 1.21758 (P=0.91120) | 0.93010 (P=0.67830) |
| PMK_A | 1.12414 (P= 0.88400) | 0.61510 (P= 0.56310() | 0.00000 (P= 1.00000) | na | 0.00000 (P= 1.00000) | 1.09861 (P=0.43210) |
| PMK_W | 0.51854 (P= 0.74030) | 1.28582 (P= 0.75200) | 1.71268 (P=0.97210) | 1.54760 (P=0.78690) | 1.44126 (P=0.95060) | 1.33823 (P=0.73150) |
| MaunaLoa | -1.49591 (P=0.05480) | **-3.67449 (P=0.01890)** | -0.97371 (P=0.17880) | *-2.56878 (P=0.06770)* | -1.50956 (P=0.05350) | **-4.70771 (P=0.01030)** |
| Kip15 | -0.78012 (P= 0.19890) | **-1.87180 (P= 0.02750)** | 0.00000 (P= 1.00000) | na | 0.00000 (P= 1.00000) | 1.09861 (P=0.42140) |
| Kip8th | 0.00000 (P= 1.00000) | na | 0.00000 (P= 1.00000) | na | 0.00000 (P= 1.00000) | na |
| Kip2' | 0.00000 (P= 1.00000) | na | 0.00000 (P= 1.00000) | 0.00000 (P=0.25070) | 0.00000 (P= 1.00000) | 0.00000 (P=0.25830) |
| Kip35 | 0.00000 (P= 1.00000) | na | 0.00000 (P= 1.00000) | na | 0.00000 (P= 1.00000) | na |
| Kip5 | 0.00000 (P=0.77360) | 0.13353 (P=0.27950) | 0.00000 (P=0.80340) | 2.35665 (P=0.80500) | 0.00000 (P= 0.76230) | 0.30830 (P=0.38650) |
| Kip33 | 0.00000 (P= 1.00000) | na | 0.00000 (P= 1.00000) | na | 0.00000 (P= 1.00000) | na |
| F | -0.96422 (P=0.19470) | 0.49191 (P=0.60410) | -0.97349 (P=0.20470) | -2.08596 (P=0.05360) | -0.94710 (P=0.20500) | -2.04857 (P=0.08100) |
| Lpp_H | 0.00000 (P= 1.00000) | na | 0.00000 (P= 1.00000) | na | 0.00000 (P= 1.00000) | na |
| Ka'u_Kahuku | 0.00000 (P= 1.00000) | na | 0.00000 (P=0.82740) | -0.34093 (P=0.18710) | 0.00000 (P=0.82980) | -0.34093 (P=0.18780) |

*Significant values in bold*

**Table 12: Molecular diversity indexes for COI sequences in *T. waikamoi***

| Site | n | H | Gene Diversity | S | Nucleotide diversity (average over loci) |
|---|---|---|---|---|---|
| **Upper Waikamoi** | 43 | 33 | **0.9889 +/- 0.0070** | 80 | **0.018506 +/- 0.009496** |
| **Pu'u Kukui** | 18 | 7 | **0.6340 +/- 0.1269** | 76 | **0.028086 +/- 0.014641** |
| Pu'u Kukui upper | 15 | 5 | 0.5619 +/- 0.1434 | 49 | 0.022775 +/- 0.012134 |
| Pu'u Kukui lower | 3 | 3 | 0.0 +/- 0.0 | 54 | 0.057508 +/- 0.043572 |

*In bold the aggregated data for all the sites in the volcano.*

**Table 13: Molecular diversity indexes for Cytb sequences in *T. waikamoi***

| Site | n | H | Gene Diversity | S | Nucleotide diversity (average over loci) |
|---|---|---|---|---|---|
| **Upper Waikamoi** | 16 | 16 | 1.0000 +/- 0.0221 | 81 | 0.058044 +/- 0.030045 |
| **Pu'u Kukui** | 24 | 7 | 0.5580 +/- 0.1177 | 53 | 0.016997 +/- 0.009095 |
| Pu'u Kukui upper | 19 | 6 | 0.6023 +/- 0.1242 | 20 | 0.012629 +/- 0.007018 |
| Pu'u Kukui lower | 5 | 2 | 0.4000 +/- 0.2373 | 41 | 0.033538 +/- 0.021136 |

*In bold the aggregated data for all the sites in the volcano.*

**Table 14: Molecular diversity indexes for ND1 sequences in *T. waikamoi***

| Site | n | H | Gene Diversity | S | Nucleotide diversity (average over loci) |
|---|---|---|---|---|---|
| **Upper Waikamoi** | 34 | 17 | 0.9144 +/- 0.0315 | 40 | 0.023189 +/- 0.012295 |
| **Pu'u Kukui** | 15 | 7 | 0.8571 +/- 0.0645 | 33 | 0.026611 +/- 0.014557 |
| Pu'u Kukui upper | 13 | 5 | 0.8077 +/- 0.0766 | 14 | 0.019306 +/- 0.010969 |
| Pu'u Kukui lower | 2 | 2 | 1.0000 +/- 0.5000 | 27 | 0.079412 +/- 0.080869 |

*In bold the aggregated data for all the sites in the volcano.*

**Table 15: Molecular diversity indexes for concatenated sequences in *T. waikamoi***

| Site | n | H | Gene Diversity | S | Nucleotide diversity (average over loci) |
|---|---|---|---|---|---|
| **Upper Waikamoi** | **9** | **9** | **1.0000 +/- 0.0524** | **146** | **0.036628 +/- 0.019890** |
| **Pu'u Kukui** | **12** | **8** | **0.8939 +/- 0.0777** | **176** | **0.036470 +/- 0.019130** |
| Pu'u Kukui upper | 11 | 7 | 0.8727 +/- 0.0891 | 120 | 0.027367 +/- 0.014554 |
| Pu'u Kukui lower | 1 | 1 | 1.0000 +/- 0.0000 | 0 | 0.000000 +/- 0.000000 |

*In bold the aggregated data for all the sites in the volcano.*

**Table 16: Neutrality tests for populations of *T. waikamoi***

| Site | **COI** | | CytB | | **ND1** | | Concatenated | |
|---|---|---|---|---|---|---|---|---|
| | **Tajima's D** | **Fu's Fs** | **Tajima's D** | **Fu's Fs** | **Tajima's D** | **Fu's Fs** | **Tajima's D** | **Fu's Fs** |
| **Upper Waikamoi** | **-1.34413 (P= 0.07150)** | **-13.72184 (P=0.00050)** | **0.69905 (P= 0.80690)** | **-3.59029 (P= 0.03940)** | **-0.70001 (P= 0.25700)** | **-1.90372 (P= 0.25440)** | **-0.06936 (P= 0.49840)** | **0.09750 (P= 0.32060)** |
| **Pu'u Kukui** | **-0.85272 (P= 0.20540)** | **7.44503 (P= 0.99170)** | **-1.61160 (P= 0.03630)** | **4.75086 (P= 0.95700)** | **-0.45903 (P= 0.35120)** | **2.68031 (P= 0.88380)** | **-0.44353 (P= 0.34940)** | **6.11348 (P= 0.98960)** |
| Pu'u Kukui upper | -0.23198 (P=0.45440) | 8.34715 (P= 0.99730) | 0.30466 (P= 0.66710) | 3.41095 (P= 0.92420) | 1.88789 (P= 0.98300) | 3.43419 (P= 0.93280) | -0.16065 (P= 0.47450) | 5.86208 (P= 0.98800) |
| Pu'u Kukui lower | 0.000 (P=0.67020) | 2.46656 (P= 0.55380) | -1.25387 (P= 0.00750) | 8.57139 (P= 0.99860) | 0.00000 (P= 1.00000) | 3.29584 (P= 0.59000) | 0.00000 (P= 1.00000) | N/A |

*In bold the aggregated data for all the sites in the volcano.*

**Table 17: Molecular diversity indexes for COI sequences in *T. brevignatha***

| Site | n | H | Gene Diversity | S | Nucleotide diversity (average over loci) |
|---|---|---|---|---|---|
| **Haleakalā (Lower Waikamoi)** | **15** | **6** | **0.7619 +/- 0.0961** | **59** | **0.030945 +/- 0.016270** |
| **BIG ISLAND** | **82** | **32** | **0.8621 +/- 0.0297** | **44** | **0.004856 +/- 0.002823** |
| **Mauna Kea** | **21** | **15** | **0.9000 +/- 0.0622** | **24** | **0.006076 +/- 0.003546** |
| Lpp | 6 | 6 | 1.0000 +/- 0.0962 | 7 | 0.004284 +/- 0.003042 |
| Lpp_H | 11 | 7 | 0.8182 +/- 0.1191 | 16 | 0.006897 +/- 0.004159 |
| Lpp_M | 4 | 4 | 1.0000 +/- 0.1768 | 6 | 0.004963 +/- 0.003849 |
| **Mauna Loa** | **33** | **12** | **0.8144 +/- 0.0508** | **26** | **0.004097 +/- 0.002499** |
| Kip15 | 15 | 7 | 0.8095 +/- 0.0753 | 21 | 0.005852 +/- 0.003507 |
| Kip18 | 4 | 2 | 0.5000 +/- 0.2652 | 1 | 0.000784 +/- 0.000972 |
| Kip5_37_23 | 6 | 6 | 1.0000 +/- 0.0962 | 9 | 0.005120 +/- 0.003533 |
| Kip33 | 2 | 2 | 1.0000 +/- 0.5000 | 2 | 0.003135 +/- 0.003839 |
| F | 6 | 4 | 0.8667 +/- 0.1291 | 3 | 0.001881 +/- 0.001598 |
| **Kīlauea** | **8** | **5** | **0.7857 +/- 0.1508** | **18** | **0.009236 +/- 0.005621** |
| Olaa | 1 | 1 | 1.0000 +/- 0.0000 | 0 | 0.000000 +/- 0.000000 |
| PMK_A | 7 | 4 | 0.7143 +/- 0.1809 | 12 | 0.005822 +/-0.003826 |
| **SW Mauna Loa** | **13** | **6** | **0.7692 +/- 0.1031** | **7** | **0.002653 +/- 0.001863** |
| Ka'u_Kapapala | 9 | 6 | 0.8889 +/- 0.0910 | 7 | 0.003135 +/- 0.002206 |
| Kona Hema | 4 | 1 | 0.0000 +/- 0.0000 | 0 | 0.000000 +/- 0.000000 |
| **Hualālai (Honoaula)** | **7** | **5** | **0.8571 +/- 0.1371** | **4** | **0.001791 +/- 0.001500** |

*In bold the aggregated data for all the sites in the volcano.*

**Table 18: Molecular diversity indexes for Cytb sequences in** *T. brevignatha*

| Site | n | H | Gene Diversity | S | Nucleotide diversity (average over loci) |
|---|---|---|---|---|---|
| **Haleakalā (Lower Waikamoi)** | **13** | **6** | **0.8846 +/- 0.0507** | **14** | **0.010006 +/- 0.005844** |
| **BIG ISLAND** | **82** | **24** | **0.8651 +/- 0.0251** | **31** | **0.005926 +/- 0.003470** |
| **Mauna Kea** | **22** | **13** | **0.8918 +/- 0.0550** | **22** | **0.006679 +/- 0.003974** |
| Lpp | 6 | 5 | 0.9333 +/- 0.1217 | 6 | 0.005190 +/- 0.003716 |
| Lpp_H | 13 | 8 | 0.8846 +/- 0.0699 | 17 | 0.007626 +/- 0.004609 |
| Lpp_M | 3 | 3 | 1.0000 +/- 0.2722 | 4 | 0.005323 +/- 0.004777 |
| **Mauna Loa** | **34** | **11** | **0.8004 +/- 0.0581** | **19** | **0.005707 +/- 0.003424** |
| Kip15 | 15 | 9 | 0.8762 +/- 0.0697 | 17 | 0.007452 +/- 0.004466 |
| Kip18 | 4 | 2 | 0.5000 +/- 0.2652 | 4 | 0.003992 +/- 0.003349 |
| Kip5_37_23 | 5 | 3 | 0.7000 +/- 0.2184 | 5 | 0.004391 +/- 0.003384 |
| Kip33 | 2 | 2 | 1.0000 +/- 0.5000 | 2 | 0.003992 +/- 0.004889 |
| F | 8 | 5 | 0.8929 +/- 0.0858 | 6 | 0.005061 +/- 0.003457 |
| **Kīlauea (PMK_A)** | **7** | **4** | **0.7143 +/- 0.1809** | **11** | **0.006843 +/- 0.004549** |
| **SW Mauna Loa** | **10** | **6** | **0.8667 +/- 0.0850** | **8** | **0.004391 +/- 0.002992** |
| Ka'u_Kapapala | 6 | 5 | 0.9333 +/- 0.1217 | 7 | 0.005057 +/- 0.003638 |
| Kona Hema | 4 | 2 | 0.5000 +/- 0.2652 | 1 | 0.000998 +/- 0.001237 |
| **Hualālai (Honoaula)** | **9** | **3** | **0.4167 +/- 0.1907** | **4** | **0.001774 +/- 0.001542** |

*In bold the aggregated data for all the sites in the volcano.*

**Table 19: Molecular diversity indexes for ND1 sequences in** *T. brevignatha*

| Site | n | H | Gene Diversity | S | Nucleotide diversity (average over loci) |
|---|---|---|---|---|---|
| **Haleakalā (Lower Waikamoi)** | **17** | **5** | **0.6397 +/- 0.1157** | **37** | **0.030114 +/- 0.016145** |
| **BIG ISLAND** | **68** | **27** | **0.8924 +/- 0.0264** | **31** | **0.008221 +/- 0.004838** |
| **Mauna Kea** | **20** | **11** | **0.8842 +/- 0.0535** | **18** | **0.011251 +/- 0.006554** |
| Lpp | 5 | 3 | 0.7000 +/- 0.2184 | 4 | 0.004494 +/- 0.003704 |
| Lpp_H | 11 | 9 | 0.9636 +/- 0.0510 | 17 | 0.015322 +/- 0.009013 |
| Lpp_M | 4 | 3 | 0.8333 +/- 0.2224 | 3 | 0.005150 +/- 0.004397 |
| **Mauna Loa** | **25** | **14** | **0.9167 +/- 0.0365** | **20** | **0.007416 +/- 0.004563** |
| Kip15 | 11 | 9 | 0.9636 +/- 0.0510 | 16 | 0.010930 +/- 0.006698 |
| Kip18 | 4 | 2 | 0.5000 +/- 0.2652 | 1 | 0.001408 +/- 0.001746 |
| Kip5_37_23 | 5 | 4 | 0.9000 +/- 0.1610 | 4 | 0.005056 +/- 0.004059 |
| Kip33 | 2 | 2 | 1.0000 +/- 0.5000 | 1 | 0.002809 +/- 0.003973 |
| F | 3 | 3 | 1.0000 +/- 0.2722 | 3 | 0.005618 +/- 0.005297 |
| **Kīlauea (PMK_A)** | 7 | 6 | 0.9524 +/- 0.0955 | 14 | 0.012574 +/- 0.008065 |
| **SW Mauna Loa** | **11** | **7** | **0.8182 +/- 0.1191** | **6** | **0.003473 +/- 0.002682** |
| Ka'u_Kapapala | 7 | 6 | 0.9524 +/- 0.0955 | 5 | 0.004548 +/- 0.003483 |
| Kona Hema | 4 | 2 | 0.5000 +/- 0.2652 | 1 | 0.001404 +/- 0.001741 |
| **Hualālai (Honoaula)** | **5** | **2** | **0.6000 +/- 0.1753** | **1** | **0.001685 +/- 0.001846** |

*In bold the aggregated data for all the sites in the volcano.*

**Table 20: Molecular diversity indexes for the concatenated sequences in** *T. brevignatha*

| Site | n | H | Gene Diversity | S | Nucleotide diversity (average over loci) |
|---|---|---|---|---|---|
| **Haleakalā (Lower Waikamoi)** | **12** | **5** | **0.8485 +/- 0.0586** | **33** | 0.008777 +/- 0.004789 |
| **BIG ISLAND** | **59** | **42** | **0.9801 +/- 0.0082** | **95** | **0.006519 +/- 0.003359** |
| **Mauna Kea** | **15** | **15** | **1.0000 +/- 0.0243** | **59** | **0.008791 +/- 0.004703** |
| Lpp | 4 | 4 | 1.0000 +/- 0.1768 | 15 | 0.005128 +/- 0.003620 |
| Lpp_H | 8 | 8 | 1.0000 +/- 0.0625 | 45 | 0.011037 +/- 0.006278 |
| Lpp_M | 3 | 3 | 1.0000 +/- 0.2722 | 11 | 0.004905 +/- 0.003944 |
| **Mauna Loa** | **24** | **18** | **0.9710 +/- 0.0208** | **59** | **0.006044 +/- 0.003215** |
| Kip15 | 11 | 10 | 0.9818 +/- 0.0463 | 52 | 0.009073 +/- 0.004988 |
| Kip18 | 4 | 2 | 0.5000 +/- 0.2652 | 6 | 0.002008 +/- 0.001570 |
| Kip5_37_23 | 4 | 4 | 1.0000 +/- 0.1768 | 14 | 0.004794 +/- 0.003401 |
| Kip33 | 2 | 2 | 1.0000 +/- 0.5000 | 5 | 0.003344 +/- 0.003664 |
| F | 3 | 3 | 1.0000 +/- 0.2722 | 8 | 0.003567 +/- 0.002943 |
| **Kīlauea (PMK_A)** | 7 | 6 | 0.9524 +/- 0.0955 | 37 | 0.007772 +/- 0.004601 |
| **SW Mauna Loa** | **9** | **8** | **0.9722 +/- 0.0640** | **19** | **0.003493 +/- 0.002116** |
| Ka'u_Kapapala | 6 | 6 | 1.0000 +/- 0.0962 | 17 | 0.004192 +/- 0.002677 |
| Kona Hema | 3 | 2 | 0.6667 +/- 0.3143 | 2 | 0.000892 +/- 0.000916 |
| **Hualālai (Honoaula)** | **4** | **4** | **1.0000 +/- 0.1768** | **4** | **0.001449 +/- 0.001197** |

*In bold the aggregated data for all the sites in the volcano.*

**Table 21: Analysis of molecular variance (AMOVA) for *T. brevignatha***

| Source of variation (%) | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| COI | | | | | |
| Among groups | 65.88 | 85.16 | -2.20 | 4.71 | 9.13 |
| Among populations within groups | -0.17 | -0.09 | 14.77 | 10.43 | 9.61 |
| Among populations | 34.29 | 14.94 | 87.43 | 84.86 | 81.26 |
| Cytb | | | | | |
| Among groups | 72.85 | 88.66 | 9.05 | 19.75 | 26.70 |
| Among populations within groups | -0.16 | 0.74 | 1.10 | -0.90 | -1.12 |
| Among populations | 27.31 | 10.60 | 89.85 | 81.15 | 74.42 |
| ND1 | | | | | |
| Among groups | 69.36 | 84.48 | 0.69 | -1.83 | -1.71 |
| Among populations within groups | -1.59 | -0.65 | 3.19 | 4.63 | 4.27 |
| Among populations | 32.23 | 16.17 | 96.12 | 97.21 | 97.43 |
| Concatenated | | | | | |
| Among groups | 76.90 | 89.87 | 5.40 | 7.24 | 12.01 |
| Among populations within groups | -0.04 | 0.01 | -2.76 | -1.42 | -1.34 |
| Among populations | 23.14 | 10.11 | 97.36 | 94.18 | 89.33 |

*The definition of the groups is as it follows:*

**1.- Per Volcano:** Haleakalā (Lower_Waikamoi), Hualālai (Honoaula), Mauna Kea (Laupāhoehoe + Laupāhoehoe, High + Laupāhoehoe, Maulua trail), Mauna Loa (Kīpuka 15 + Kīpuka 18 + Kīpuka 5-37-23 + Forest), Kīlauea (‘Ōla’a + Pu’u Maka’ala, Army Road), Southwest Mauna Loa (Ka’u + Kona Hema)

**2.- Maui / Big Island:** Maui (Lower_Waikamoi), Big Island (Honoaula + Laupāhoehoe + Laupāhoehoe, High + Laupāhoehoe, Maulua trail + Kīpuka 15 + Kīpuka 18 + Kīpuka 5-37-23 + Forest + ‘Ōla’a + Pu’u Maka’ala, Army Road +Ka’u + Kona Hema)

**3.- Per Volcano only Big Island:** Hualālai (Honoaula), Mauna Kea (Laupāhoehoe + Laupāhoehoe, High + Laupāhoehoe, Maulua trail), Mauna Loa (Kīpuka 15 + Kīpuka 18 + Kīpuka 5-37-23 + Forest), Kīlauea (‘Ōla’a + Pu’u Maka’ala, Army Road), Southwest Mauna Loa (Ka’u + Kona Hema)

**4.- Windward / South / Leeward:** Windward (Laupāhoehoe + Laupāhoehoe, High + Laupāhoehoe, Maulua trail + Kīpuka 15 + Kīpuka 18 + Kīpuka 5-37-23 + Forest + ʻŌlaʻa + Puʻu Makaʻala, Army Road), South (Kaʻu), Leeward (Honoaula + Kona Hema)
**5.- Windward-South / Leeward:** Windward-South (Laupāhoehoe + Laupāhoehoe, High + Laupāhoehoe, Maulua trail + Kīpuka 15 + Kīpuka 18 + Kīpuka 5-37-23 + Forest + ʻŌlaʻa + Puʻu Makaʻala, Army Road + Kaʻu), Leeward (Honoaula + Kona Hema)

**Table 22: Par-wise $\phi_{ST}$ estimations for *T. brevignatha* based on COI.**

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | 0.00000 | | | | | |
| 2 | 0.01878 | 0.00000 | | | | |
| 3 | 0.00538 | 0.02893 | 0.00000 | | | |
| 4 | 0.14302 | 0.10724 | 0.21707 | 0.00000 | | |
| 5 | 0.08319 | 0.04434 | 0.04108 | 0.34474 | 0.00000 | |
| 6 | 0.66288 | 0.80550 | 0.75348 | 0.73196 | 0.71004 | 0.00000 |

1: Kilauea
2: MaunaLoa
3: MaunaKea
4: Leeward
5: Kau
6: Lower_Waikamoi

*Highlighted significant values (P<0.05).*

**Table 23: Par-wise $\phi_{ST}$ estimations for *T. brevignatha* based on Cytb.**

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | 0.00000 | | | | | |
| 2 | -0.04837 | 0.00000 | | | | |
| 3 | 0.01140 | 0.03394 | 0.00000 | | | |
| 4 | 0.34249 | 0.27619 | 0.35276 | 0.00000 | | |
| 5 | -0.06449 | -0.04877 | 0.00282 | 0.45540 | 0.00000 | |
| 6 | 0.85315 | 0.88454 | 0.86879 | 0.89887 | 0.85977 | 0.00000 |

1: Kilauea
2: MaunaLoa
3: MaunaKea
4: Leeward
5: Kau
6: Lower_Waikamoi

*Highlighted significant values (P<0.05).*

**Table 24: Par-wise $\phi_{ST}$ estimations for *T. brevignatha* based on ND1.**

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | 0.00000 | | | | | |
| 2 | -0.03662 | 0.00000 | | | | |
| 3 | 0.00429 | 0.04990 | 0.00000 | | | |
| 4 | 0.05235 | -0.00732 | 0.05292 | 0.00000 | | |
| 5 | -0.01587 | 0.00429 | 0.03449 | 0.08149 | 0.00000 | |
| 6 | 0.71829 | 0.80650 | 0.76585 | 0.77374 | 0.74737 | 0.00000 |

1: Kilauea
2: MaunaLoa
3: MaunaKea
4: Leeward
5: Kau
6: Lower_Waikamoi

*Highlighted significant values (P<0.05).*

**Table 25: Neutrality tests for populations of *T. brevignatha***

| | **COI** | | **Cytb** | | **ND1** | | **Concatenated** | |
|---|---|---|---|---|---|---|---|---|
| Site | **Tajima's D** | **Fu's Fs** | **Tajima's D** | **Fu's Fs** | **Tajima's D** | **Fu's Fs** | **Tajima's D** | **Fu's Fs** |
| **Haleakalā (Lower Waikamoi)** | 0.38114 (P=0.70680) | 8.28261 (P=0.99700) | 0.46112 (P=0.72280) | 1.28957 (P=0.74590) | -0.08424 (P=0.52170) | 7.48365 (P=0.99470) | 0.90712 (P=0.85570) | 5.99761 (P=0.98880) |
| **BIG ISLAND** | **-2.09065 (P=0.00280)** | **-23.73919 (P=0.00000)** | **-1.63515 (P=0.02600)** | **-11.81769 (P=0.00040)** | **-1.84963 (P=0.00880)** | **-18.40718 (P=0.00000)** | **-1.84943 (P=0.00950)** | **-23.03343 (P=0.00000)** |
| **Mauna Kea** | **-1.59850 (P=0.03930)** | **-7.84269 (P=0.00040)** | **-1.67024 (P=0.03060)** | **-5.33696 (P=0.00570)** | -0.82352 (P=0.22750) | -2.59140 (P=0.10410) | -1.22182 (P=0.09980) | **-6.04194 (P=0.00780)** |
| Lpp | -0.63095 (P=0.34260) | **-3.51388 (P=0.00670)** | -0.06042 (P=0.48130) | -1.56520 (P=0.07470) | -1.04849 (P=0.15200) | 0.27642 (P=0.49250) | -0.62393 (P=0.36000) | 0.01708 (P=0.28100) |
| Lpp_H | -0.86708 (P=0.19770) | -0.66422 (P=0.33680) | -1.27799 (P=0.09860 | -1.47751 (P=0.18910) | -0.27367 (P=0.42420) | -2.51264 (P=0.06800) | -0.28392 (P=0.39040) | -1.22928 (P=0.15190) |
| Lpp_M | -0.31446 (P=0.53600) | -1.15708 (P=0.08980) | 0.00000 (P=0.83060) | -0.34093 (P=0.18380) | 1.08976 (P=0.85020) | 0.00617 (P=0.29660) | 0.00000 (P=0.72240) | 0.80681 (P=0.42560) |
| **Mauna Loa** | **-2.07875 (P=0.00480)** | -3.57086 (P=0.04530) | -1.30187 (P=0.08310) | -2.12701 (P=0.16770) | **-2.01551 (P=0.00830)** | **-7.51179 (P=0.00030)** | **-1.72432 (P=0.02470)** | -4.87151 (P=0.03410) |
| Kip15 | **-1.73487 (P=0.02960)** | -0.08655 (P=0.49760) | -1.15519 (P=0.12730) | -2.04057 (P=0.12600) | -1.34846 (P=0.09110) | **-3.64441 (P=0.01660)** | -1.14127 (P=0.13090) | -1.45856 (P=0.18030) |
| Kip18 | -0.61237 (P=0.37550) | 0.17185 (P=0.33490) | -0.78012 (P=0.19440) | 2.19722 (P=0.82750) | -0.61237 (P=0.38120) | 0.17185 (P=0.34350) | -0.80861 (P=0.16290) | 2.94444 (P=0.89720) |
| Kip5_37_23 | -1.01988 (P=0.19710) | **-3.07918 (P=0.01260)** | -0.56199 (P=0.39580) | 0.80363 (P=0.65170) | -1.04849 (P=0.140400 | -1.19525 (P=0.09540) | -0.84307 (P=0.09880) | -0.06549 (P=0.27620) |
| Kip33 | 0.00000 (P=1.00000) | 0.69315 (P=0.36850) | 0.00000 (P=1.00000) | 0.69315 (P=0.36800) | 0.00000 (P=1.00000) | 0.00000 (P=0.25040) | 0.00000 (P=1.00000) | 1.60944 (P=0.50410) |
| F | -0.44736 (P=0.34450) | -1.45444 (P=0.05340) | 0.44652 (P=0.69580) | -0.63952 (P=0.26870) | 0.00000 (P=0.98420) | -0.69315 (P=0.13430) | 0.00000 (P=0.77560) | 0.45758 (P=0.39000) |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Kīlauea (PMK_A) | -1.31587 (P=0.09980) | 1.09198 (P= 0.70410) | -1.27839 (P=0.10790) | 0.92756 (P=0.67950) | -1.34803 (P=0.09200) | -1.35415 (P=0.14380) | -1.38637 (P=0.06630) | 0.41029 (P=0.49300) |
| **SW Mauna Loa** | -0.94729 (P=0.18170) | -1.58664 (P=0.10930) | -0.95794 (P=0.19170) | -1.58448 (P=0.10000) | -1.56949 (P=0.05410) | **-4.56797 (P=0.00040)** | -1.24033 (P=0.11310) | -2.60030 (P=0.05560) |
| Ka'u_ Kapapala | -0.99686 (P=0.18470) | -2.17585 (P=0.04220) | -1.01063 (P=0.20520) | -1.61752 (P=0.07410) | -1.02379 (P=0.18800) | **-3.87750 (P=0.00140)** | -0.98191 (P=0.20020) | -1.73385 (P=0.08500) |
| Kona Hema | 0.00000 (P=1.00000) | NA | -0.61237 (P=0.37830) | 0.17185 (P=0.32720) | -0.61237 (P=0.38290) | 0.17185 (P=0.33650) | 0.00000 (P=0.93390) | 1.06087 (P=0.59480) |
| **Hualālai (Honoaula)** | -1.43414 (P=0.06230) | **-2.85789 (P=0.00300**) | **-1.60974 (P=0.03760)** | 0.13355 (P=0.43190) | 0.00000 (P=1.00000) | 0.62615 (P=0.507400 | -0.75445 (P=0.23520) | -1.74087 (P=0.03950) |

*First Column: in bold are the name of the volcanoes with the respective estimations for the aggregated data set.*
*Data matrix: in bold are the significant estimations.*

**Table S1: Molecular diversity indexes for COI sequences in** *T. macracantha*

| Site | n | H | Gene Diversity | S | Nucleotide diversity (average over loci) |
|---|---|---|---|---|---|
| **Ko'olau** | **25** | **17** | **0.943 +/-0.032** | **38** | **0.014 +/-0.07** |
| *Ko'olau Cramp* | 18 | 13 | 0.941 +/-0.041 | 27 | 0.014 +/-0.007 |
| *Ko'olau Day 2* | 7 | 6 | 0.952 +/-0.010 | 24 | 0.016 +/-0.010 |

*In bold the aggregated data for all the sites in the volcano.*

**Table S2: Molecular diversity indexes for Cytb sequences in** *T. macracantha*

| Site | n | H | Gene Diversity | S | Nucleotide diversity (average over loci) |
|---|---|---|---|---|---|
| **Ko'olau** | **24** | **12** | **0.8551 +/- 0.0648** | **19** | **0.008454 +/- 0.004859** |
| Ko'olau Cramp | 17 | 10 | 0.8750 +/- 0.0702 | 16 | 0.008000 +/- 0.004721 |
| Ko'olau Day 2 | 7 | 4 | 0.8095 +/- 0.1298 | 12 | 0.009933 +/- 0.006311 |

*In bold the aggregated data for all the sites in the volcano.*

**Table S3: Molecular diversity indexes for ND1 sequences in** *T. macracantha*

| Site | n | H | Gene Diversity | S | Nucleotide diversity (average over loci) |
|---|---|---|---|---|---|
| **Ko'olau** | **24** | **12** | **0.9058 +/- 0.0380** | **17** | **0.013779 +/- 0.007796** |
| *Ko'olau Cramp* | 17 | 10 | 0.9044 +/- 0.0497 | 16 | 0.013841 +/- 0.007977 |
| *Ko'olau Day 2* | 7 | 5 | 0.9048 +/- 0.1033 | 11 | 0.015126 +/- 0.009550 |

*In bold the aggregated data for all the sites in the volcano.*

**Table S4: Molecular diversity indexes for the concatenated sequences in** *T. macracantha*

| Site | n | H | Gene Diversity | S | Nucleotide diversity (average over loci) |
|---|---|---|---|---|---|
| **Ko'olau** | **21** | **21** | **1.0000 +/- 0.0147** | **65** | **0.011748 +/- 0.006077** |
| Ko'olau Cramp | 15 | 15 | 1.0000 +/- 0.0243 | 53 | 0.011795 +/- 0.006235 |
| Ko'olau Day 2 | 6 | 6 | 1.0000 +/-0.0962 | 47 | 0.012255 +/- 0.007350 |

*In bold the aggregated data for all the sites in the volcano.*

**Table S5: Neutrality tests for the Ko'olau population of *T. macracantha***

| Site | **COI** | | CytB | | **ND1** | | Concatenated | |
|---|---|---|---|---|---|---|---|---|
| | **Tajima's D** | **Fu's Fs** | **Tajima's D** | **Fu's Fs** | **Tajima's D** | **Fu's Fs** | **Tajima's D** | **Fu's Fs** |
| **Ko'olau** | **0.324 (P=0.685)** | **-3.559 (P=0.083)** | **-0.68015 (P=0.26730)** | **-2.60475 (P= 0.12090)** | **0.10426 (P=0.59440)** | **-2.06297 (P= 0.17560)** | **-0.23751 (P= 0.46460)** | **-9.16117 (P= 0.00170)** |
| Ko'olau Cramp | 0.309 (P=0.664) | -2.227 (P=0.153) | -0.67249 (P= 0.27160) | -2.39304 (P= 0.10630) | -0.02197 (P= 0.53310) | -1.72544 (P= 0.18510) | 0.20295 (P= 0.62700) | -4.89907 (P= 0.01800) |
| Ko'olau Day 2 | 0.227 (P=0.591) | 0.183 (P=0.451) | -0.04537 (P= 0.48840) | 1.66468 (P= 0.79840) | 0.78670 (P= 0.80580) | 0.34885 (P= 0.52360) | -0.88634 (P= 0.22880) | -0.14786 (P= 0.27720) |

*In bold the aggregated data for all the sites in the volcano.*

# CHAPTER 3

## Implementation of transcriptome-based Exon Capture approach reveals cryptic diversity on young species from the Hawaiian *Tetragnatha* adaptive radiation


## ABSTRACT

How adaptive radiations get started is one of the important questions related to temporal dynamics of rapid diversification events. The chronosequence of the Hawaiian Islands provide an ideal system to explore this problem. Here we examined the genetic signatures of young speciation events by applying the transcriptome-based Exon Capture approach to the study of three closely related species (*T. brevignatha, T. waikamoi* and *T. macracantha*) from Maui Nui and Big Island. The combined data showed the existence of 5 clades: *T. brevignatha* Big Island, *T. brevignatha* Maui, *T. macracantha*, *T. waikamoi* and a WaikBrevMac clade. The last one includes specimens from all three nominal species. For *T. brevignatha*, the molecular divergence between the populations of Maui and Big Island are similar to "species level" comparisons. The Big Island population shows a strong break between the Leeward and Windward populations. For *T. waikamoi* nuclear data suggests that the population break between Haleakalā and West Maui is less pronounced than the break based on mitochondrial data. In the case of *T. macracantha*, we found a separation between the Maui and Lana'i populations. Moreover, there are two very divergent lineages co-existing in Kīpahulu Valley. Lastly, the finding of the WaikBrevMac group adds a new layer of complexity where recent speciation events suggest divergent lineages with convergent (or are the ancestral) morphologies and no genetic admixture. The addition of species with different eco-morphologies will complement these findings and also inform about the genetic signatures of ecological speciation.

## INTRODUCTION

Adaptive radiations provide the unique opportunity to study a wide variety of questions in evolutionary biology (Schluter, 2000; Seehausen, 2006; Losos and Ricklefs, 2009). These episodes of rapid diversification occur at different times and geographic scales across the entire tree of life. One question to be addressed is, what processes are involved in the initiation of adaptive radiations? What makes the study of species formation interesting in the context of an adaptive radiation is that many species are been created from a single lineage, and might be considered to represent a series of replicates of the speciation events.

What are the genetic patterns of young speciation events? The most evident signal of speciation is the reduction of gene flow between the divergent populations (Feder et al., 2013). However, this is not absolute outcome and recent studies show that the diversification process could also include some degree of gene flow (Nosil, 2008). The specific relative role of introgression and reproductive barriers has not been studied extensively in the context of adaptive radiations. Some divergent lineages, such as sticklebacks, slowly separate along a gradient where there is a declining degree of gene flow between populations (Hendy et al., 2009). The other extreme consist in a situation in which geographic or other biological barriers stops gene flow (Dobzhansky, 1951; Mayr, 1963). Understanding the genetic signals left by past speciation events will provide insights to the earlier stages of diversification.

The Hawaiian Islands are an ideal system to ask questions related with the initiation of adaptive radiations. These islands are arranged on a chronosequence produced by the movement of the Pacific plate over a semi-stationary volcanic hot spot. The oldest islands are in the northwest and the youngest in the southeast (Price and Clague, 2002). Each island is a "snapshot" of the geologic development of a volcanic island. This "snapshot" idea can also be applied to the organisms that are evolving on this dynamic landscape (Roderick et al., 2012).

Some of the largest lineages from Hawai'i (Animals: *Platyini*, *Tetragnatha*, *Drosophila, Nesosydne, Laupala*; Plants: *Lobeliads*, *Cyrtandra*) represent a peak in species diversity per unit of area on the middle age islands (Maui or O'ahu), while small lineages show a constant increase in diversity towards the oldest islands (Gillespie and Baldwin, 2010). This pattern suggests that for big lineages there is a temporal change in the factors that drive their diversification, while for small lineages there is a cumulative effect over time. Another alternative could be the existence of an extinction debt due to the natural environmental reduction, because of the ageing of the island (Kuussaari et al., 2009).

Carson et al., (1990) considered the Big Island as a "Crucible of Evolution", where the combination of constant lava flow events and isolation between volcanoes fosters the formation of new species (Carson and Templeton, 1984; Vandergast et al., 2004). Other studies have also consider this island as the "snapshot" where species formation is started (Stacy et al., 2014). As the islands reach their middle age, a natural expectation is that all the isolated populations will have evolve into new allopatric species. Maui Nui represents the next "snapshot" in the diversification process. Hypothetically, species boundaries should be more distinct than what is found on younger islands. The presence of newly produced species should be also found in O'ahu

depending on the group. Finally on older islands (like Kaua'i), landslides, erosion and more importantly subsidence will deplete some of the original environments causing local extinctions. Alternatively, species could move and come into secondary contact, which might or might not produce competitive exclusion (Tapia et al., 2008; Waters, 2011).

One of the lineages showing this biodiversity trajectory are the Hawaiian *Tetragnatha* spiders (Gillespie et al., 1994; Gillespie, 2004). This lineage is represented by more than 50 species in the Hawaiian archipelago in less than 5 million years. They can by classified into two adaptive radiations. The first one corresponds to the web-building clade, which has the largest number of species (Blackledge and Gillespie, 2004). The second, and better studied, is the Spiny leg clade. This group consists of 16 species that display 4 different eco-morphologies, some displaying convergent evolution, and match the background that they live on (Gillespie, 2004). Two species with the same eco-morphology do not co-occur in the same geographical area.

If the idea about within island allopatric speciation on young volcanic islands is correct, a natural expectation would be to find many sister pairs of allopatric species on a middle age island as Maui Nui. Within the Spiny leg clade there are three species from the Green Eco-morph that live on Maui Nui (Fig. 1a). The first one is *T. waikamoi*, which has two described populations: one from Pu'u Kukui (West Maui) and the other in Upper Waikamoi (Haleakalā). The second species is *T. macracantha* which has populations on Lana'i island, Kīpahulu Valley (east Haleakalā) and a recently discovered population in the Ko'olau Rorest Reserve also on the east side of Haleakalā. Finally *T. brevignatha*, has populations in Lower Waikamoi (Haleakalā) and several volcanoes in the Big Island (Mauna Kea, Mauna Loa, Southwest slope of Mauna Loa, Hualālai and Kīlauea) (Fig. 1b).

Previous phylogenetic reconstructions show that *T. brevignatha* is a paraphyletic species. The Maui population is closer to *T. macracantha*, than to *T. brevignatha* on the Big Island. The clade consisting of *T. macracantha/T. brevignatha* is sister to a clade that includes *T. waikamoi* and two other species (*T. kamakou* (Maroon eco-morph) and *T. restricta* (Small brown eco-morph)) (Gillespie, 2004). These two species corresponds to a different eco-morph and live in sympatry with *T. waikamoi* and *T. macracantha* and it is possible that ecological speciation was involved on the divergence event.

Ecological and allopatric speciation are thought to have been involved in the diversification of Hawaiian *Tetragnatha* spiders. Here we will focus only in the role of allopatry. Which evolutionary processes explain the formation and current distribution of the three Green eco-morph species? One hypothesis is that the ancestral species was distributed on three volcanoes (Lana'i, West Maui and Haleakalā) and evolved into three allopatric species (*T. macracantha, T. waikamoi* and *T. brevignatha*, respectively). Then, *T. macracantha* and *T. waikamoi* dispersed into different regions of Haleakalā (Kīpahulu Valley / Ko'olau and Upper Waikamoi, respectively) and *T. brevignatha* colonized the Big Island. An alternative hypothesis is that the three species evolved in allopatry on different regions of Haleakalā (*T. brevignatha* in Lower Waikamoi; *T. waikamoi* in Upper Waikamoi and *T. macracantha* in Ko'olau/Kīpahulu Valley). Then, the species disperse to different volcanoes (*T. brevignatha* to Big Island; *T. waikamoi* to West Maui, and *T. macracantha* to Lana'i). These two scenarios correspond to two extremes and most likely the actual speciation and colonization dynamic corresponds to a combination of these two

elements. It is possible to consider situations where two species evolved in Haleakalā and the third one colonize the area later on. On the other hand, it is important to mention that the separation of the species pair *T. kamakou*/*T. restricta* and *T. waikamoi* was probably related with ecological speciation, but it will not be analyzed in the present study.

Our general expectation is to find the highest genetic diversity on the volcano of origin for each species and a subsample of that diversity (the result of a founder event) in the area later colonized (Nei et al., 1975).

Population genetics studies using three mitochondrial markers (COI, ND1 and Cytb) showed that there is a strong population break between Maui and Big Island populations for *T. brevignatha*. Also, the Maui population shows higher nucleotide diversity and there is a very strong signal of population expansion on the aggregate of Big Island populations. However, this last signal is possibly associated to more recent demographic events than the initial colonization of the island (Cotoras, Chapter 2). For *T. waikamoi* there is strong population structure between the two volcanoes, with only one shared haplotype amongst the three markers. The neutrality tests do not show a consistent signal of population expansion in any of the two volcanoes, so it is difficult to determine where the colonization event actually happen (Cotoras, Chapter 2). There is no complete data set for *T. macracantha*.

The data available for *T. brevignatha* and *T. waikamoi* is limited to three mitochondrial genes. Due to its fast mutation rate they will only pick up very recent demographic events. Also, as they are maternally inherited and linked, they represent a biased version of the species demographic history and correspond to a smaller effective population size than the nuclear genes (Ballard and Whitlock, 2004). However, it is possible to obtain general insights about phylogenetic structure and population separation with these data. However, in order to answer additional questions at or below the species level a large number of independent nuclear markers is required. This is important to avoid getting a biased signal related to the evolutionary history of a single gene. Today, it is possible to obtain them with the application of high throughput sequencing technologies (Wang and Nielsen, 2012).

The main goal of this project is to understand the evolutionary processes that produce three closely related species in the Hawaiian *Tetragnatha* group (two sister species and one from a sister clade) and which demographic events explain their current distribution. This information will provide insights about early stages of diversification on the context of an adaptive radiation.

## MATERIALS AND METHODS

### Collections

The spiders were collected during five field seasons (August 2010, June 2011, January 2012, June 2012 and June 2013). The sampling occurred during the day and night using a beating sheet and hand collecting. The sites visited are on State, Federal and private land. Access included 4WD, hiking and helicopter transportation. The exact collecting sites are presented Table 1. All the specimens were preliminarily identified in the field and then confirmed in the lab. While alive, they were photographed with a Nikon 3100 SLR camera with an 85 mm macro lens using the same black paper as a

background. These pictures were useful for corroborating identifications. The samples were preserved in 95% ethanol at -20°C for genomic work. As indicated above the specimens collected were identified as *T. macracantha, T. brevignatha* and *T. waikamoi*.

**Background on Exon Capture and Next Generation Sequencing**

The advent of high throughput sequencing technologies has allowed the exploration new questions and the possibility of answering old ones (Brewer et al., 2014). Next Generation Sequencing (NGS) can be performed using a wide variety of sequencing platforms and depending on the particular question some applications will be more suited than others.

The applications are as wide as the field of genetics. They include: genomic structure, population genetics, phylogenetics, functional genomics, epidemiology, etc. On arachnids these methods are starting to be applied in studies related to venom evolution, adaptation, phylogenetics and population genetics (Brewer et al., 2014). One of the biggest challenges for arachnology to enter into the genomic revolution has been the difficulty to sequence a spider genome. The first assembled genome corresponded to the African social spider *Stegodyphus mimosarum* published this year, along with a draft assembly for the mygalomorph Brazilian white-knee tarantula, *Acanthoscurria geniculate* (Sanggaard et al., 2014). The main factor that holds up the sequencing of a spider genome corresponds to its generally very high AT content. In general, for animals the assembly contig N50 length (bp) is negatively correlated with the % GC. Many spider (most of them Theridiidae) genome sequencing project correspond to outlyers that have even higher %AT (lower %GC) than other animals (Brewer et al., 2014). This situation is an analytical obstacle that confounds the genome assembly process.

Although a reference genome would be very useful, it is not completely essential for our study. There are many NGS methods that allow doing *de novo* assembly without a reference. A commonly used method for population genomics corresponds to RAD Seq (Davey et al., 2011). It is based on the digestion of the genome using restriction enzymes with infrequent cutting sites. This site-specific fragmentation provides the homologous sites that are used later for the alignment. This method requires intact DNA to start the digestion and, on the first protocols, a large amount of DNA per sample was required. This is due to the subsequent random fragmentation step and the later size selection using gel excision. Originally, this was the approach that we plan to take, however we encounter many technical problems. Most of them related with not been able to extract enough DNA for the protocol, inefficient or too efficient physical fragmentation methods and incorrect adaptor concentrations (Cotoras, *unpublished data*). For these reasons, we decided to transition into a transcriptome-based Exon Capture.

The transcriptome-based Exon Capture method corresponds to a particular case of a capture approach where a set of probes derived from a transcriptome are printed on a chip or in solution baits (Bi et al., 2012). Each sample is prepared into a genomic library for NGS sequencing (Meyer and Kircher, 2010). After indexing PCR the samples are multiplexed and used in a hybridization experiment (Hodges et al., 2009). The objective of the hybridization is to do enrichment with the sequences present in the capture array/baits and deplete the ones not present in the original target. The elution from the hybridization is later amplified and sent to a sequencing facility.

The use of chip or in solution-based methods depends on the question, magnitude of the project and budget. A chip-based experiment will have a lower capture efficiency, given the fact that the hybridization of the genomic DNA with the probes will occur on a surface instead of a tridimensional arrangement as occurs in solution. In the same way, the target size and the number of individuals that can be multiplexed on a chip is smaller than the solution approach. However, the chip technology is cheaper than the solution method and for smaller projects represents a better option (for more detailed comparison between methods see: Kiialainen et al., 2011 and Sulonen et al., 2011).

The analytical pipeline for Exon Capture is also different from the RAD Seq. data. The main difference is related with the assembly step. While in RAD Seq. it is possible to use the program Stacks (Catchen et al., 2013), the alignment of the Exon Capture data is done using the target as a reference (see below for Supplementary Material).

The downstream analysis will also depend on the question being asked. For phylogenetic reconstruction and detecting selection and functional genomics, high coverage is required in order to confidently identify genotypes. For many demographic questions high coverage is not a strict requirement anymore due to the development and implementation of methods that use genotype likelihoods (Nielsen et al., 2012). These methods are based in the use of the shape of the curve of the Site Frequency Spectrum (Nielsen et al., 2012). This is a powerful tool to make relative demographic inferences with low coverage data and improve significantly the result in cases with moderate depth of coverage.

**DNA extraction and quantification**

The Exon Capture approach was originally developed to work with highly degraded DNA from historic collections, but the protocol easier if you start with high quality DNA. For that reason we tested two commercially available methods for DNA extractions on *T. hawaiensis*. The first method was the column based Qiagen Kit and the second a salt precipitation-based 5PRIME kit. We finally decided to use the column based Qiagen Kit.

The DNA was extracted form 4 legs from the right side of each the spider. If the spider was of medium size we used all the legs and for very small ones we included the cephalothorax. We avoid as much as possible to use the abdomen, however in some cases it was necessary in order to reach the 400-500 ng required for the protocol. The double stranded DNA was measured with a Qubit® 2.0 Fluorometer (Life Technologies). We performed two replicates of each measurement and worked with the average value.

**Transcriptome Sequencing**

In order to obtain a reference transcriptome we performed a whole body extraction of RNA of a fresh frozen preserved specimen of *T. brevignatha*. The RNA was isolated following a trizol extraction protocol. The library preparation and sequencing was outsourced to Hudson Alpha (Huntsville, AL. USA).

**Positive and Negative controls**

The protocol described by Bi et al. (2012) suggests the use of positive and negative controls to determine the success of the capture experiment. Based on whole body transcriptome sequences of *T. kauaiensis* and *T. perreirai*, we performed dN/dS tests to determine which markers were evolving fast and which ones slowly (Yim et al., 2014; Brewer et al., *submitted*). We selected 16 fast evolving and 6 slow evolving transcripts. We also included the sequence for Histone 3. All sequences were BLAST against a partially annotated whole body transcriptome of *T. brevignatha*.

The *T. brevignatha* sequences were used as references for primers to amplify 500 bp in the case of the fast evolving markers and 150 bp for the slow evolving ones. The primers were designed using the web-based program Primer 3 (Koressaar and Remm, 2007; Untergrasser et al, 2012). Sequences for the primers are presented in Table 2.

To test the primers a PCR reaction mix of: 1 ul of each primer at 10 uM, 2 ul of AmpliTaq (Life Technologies) buffer 10x, 0,5 ul of $MgCl_2$ 25 mM, 11.7 ul $H_2O$, 1 ul BSA 1x, 0.2 ul DreamTaq (Thermo Scientific) and 1 ul of DNA extraction was used. The amplification profile started with 2 minutes at 95°C, followed by 40 repetitions of a cycle which started with 30 seconds at 95°C, then 30 seconds at 60°C and finally one minute at 72°C; there was an extra step of extension at 72°C for ten minutes. The PCR products were verified on an agarose/TBE gel. The annealing temperature (60°C) corresponds to a hard constraint, as it is the temperature that will be used during the qPCR as part of the Post capture validation.

For testing the efficiency of the primers that amplify fast evolving sequences we used two specimens of *T. brevignatha* from the same population in Maui (Lower Waikamoi). For the slow evolving gene primers (including Histone 3) we used two specimens of *T. brevignatha* from different islands (Kīpuka area in Big Island; and Lower Waikamoi in Maui).

We tried fast and slow evolving nuclear genes to look for nuclear markers that could be also used for Sanger sequencing for population and phylogenetic applications. The selected positive controls remained on the chip, while the selected negative control was removed. Note that for all these controls we checked for the presence of duplicates or similar sequences. Finally, we choose as a positive control the sequences Tb40119 and Tb9898 and as a negative control a fragment of Histone 3. The primer sequences correspond to Tb40119.2, Tb9898.2 and H3.1.

**Probe design (see Supplementary material)**

**COT 1 DNA library preparation (also see Supplementary material)**

The COT1 DNA corresponds to the fraction of the genome that has the lowest complexity, and re-hybridization experiments are used to determine this because the COT1 fraction is the first to re-anneal after denaturation.

Due to its low complexity it is possible that it will interfere and compete with more complex sequences during the hybridization phase. The COT1 library was used to amplify specific *Tetragnatha* COT1 DNA, which was used to block the COT1 DNA present in the genomic libraries.

For the preparation of *Tetragnatha* COT1 DNA we started with whole body DNA extraction (Qiagen Kit) of three large adults of *T. quasimodo*. This species is particularly large within the Hawaiian *Tetragnatha*. The COT1 DNA protocol corresponds to a modified version of that in Trifonov et al., 2009.

**Genomic library preparation (also see Supplementary material)**

A total of 116 libraries were prepared following the protocol described by Meyer and Kircher (2010). Seventy corresponded to *T. brevignatha*, 20 to *T. waikamoi* and 26 to *T. macracantha*. The details of the localities are presented in Table 1. The DNA quality was verified on an agarose/TBE gel and the amount calculated based on a concentration estimation measured with Qubit® 2.0 Fluorometer (Life Technologies) (2 replicates). The amount of starting material ranged from 400 to 500 ng depending on the sample.

**Hybridization Experiment (also see Supplementary material)**

A total of three hybridization experiments were performed using the same target, but different chips. The specimen composition of each experiment is as follows:

Experiment 1: 50 *T. brevignatha*
Experiment 2: 20 *T. brevignatha* + 16 *T. macracantha* + 17 *T. waikamoi*
Experiment 3: 10 *T. macracantha* + 3 *T. waikamoi* + 39 *T. anuenue*

Experiment 1 consists of the same species that we used to design the probes (*T. brevignatha*). The first experiment was also a test run for the probes. Experiment 2 included specimens from the three closely related Green ecomorph species. Finally, Experiment 3 had specimens from the two green ecomorph species (*T. macracantha* and *T. waikamoi*) and *T. anuenue*. The last one is also a member of the Spiny leg *Tetragnatha* group, but the analysis and discussion of these data will be addressed later (Cotoras, *in prep.*).

The concentrations of the libraries after indexing PCR were measured with NanoDrop (Thermo Scientific). By using bead cleaning before the measurements, the reads from the NanoDrop (Thermo Scientific) were not affected by salts or free nucleotides. In the multiplex step for each experiment we collected a total of 21 μg of DNA corresponding to equimolar amounts from all the libraries used. 20 μg went into the hybridization and 1 μg was stored as a "Pre-Capture aliquot" for positive and negative controls after the hybridization.

Experiment 1 was performed first and Experiments 2 and 3 were performed in parallel. The hybridization procedure followed the protocol described by Hodges et al. (2009) from step 29 to 61.

**Post-Capture controls (see Supplementary material)**

**Sequencing platform**

Each experiment was sequenced in one sequencing lane of an Illumina HighSeq 2500 platform with 100 Pair End reads in the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley (supported by NIH S10 Instrumentation Grants S10RR029668 and S10RR027303).

**Analytical pipeline (see Supplementary material)**

**RESULTS**

**DNA extraction and quantification**

One of the three samples failed on both methods but two worked adequately. Both methods extracted high molecular weight DNA (Fig. 2). The yield of the column based Qiagen Kit was higher in the two remaining extractions (Table 3). This method was faster and collaborators have not seen any problems in DNA quality (Smith *pers. comm.*). For these reasons we decided to use the column based Qiagen Kit for the DNA extractions of all samples.

**Transcriptome Sequencing**

The whole body transcriptome sequencing of *T. brevignatha* yielded 14,664,987 bp raw data and 1,775,169 bp after all the filters.

**Probe Design (also see Supplementary material)**

The final target size consisted of 1.7 Mega bases and corresponds to 1,826 ORFs. With a probe length of 60 bp and using a tailing density of 2 bp we were able to generate a final total of 967,487 probes, which is made up of 814,187 original probes plus 153,300 added to the terminal ends of each ORF. We added extra probes on the ends of each ORF increase the coverage. In the description of the transcriptome-based Exon Capture approach (Bi et al., 2012) reported a strong tendency for a reduction in coverage at the ends of each exon sequence.

**Positive and Negative controls**

For these positive and negative controls we tested fast and slow evolving genes. The test for fast evolving genes showed that most of the primers were not able to produce amplification products (Fig 3). The primers to amplify Tb633 and Tb2656 produced stronger amplification, but in both cases with secondary bands. Tb3601 produced weak amplification consisting of two bands. Tb4420 and Tb4476 produced only weak amplifications.

PCR optimization, different in each case, had the following effect on products of the different primer pairs: Tb633 produced a better amplification; Tb2656 and Tb3601 main band became stronger and the unspecific amplification weaker or simply

disappeared; for Tb4420 there was no amplification and Tb4476 produced a lot of unspecific amplification (data not shown).

The fact that Tb2656 and Tb3601 produced a larger fragment than expected (500 bp) could be due to the presence of intron sequences.

For the slow evolving genes, we designed two different primer pairs per fragment. In general, the amplifications were more consistent and mostly without the presence of unspecific amplification (Fig. 4). The primer pairs that produced a satisfactory amplification were: H3.1, H3.2, Tb38716.1, Tb40119.1, Tb40119.2, Tb9898.1, Tb9898.2 and Tb37737.2.

As the amplification of slow evolving genes was more consistent we selected primer pairs that we used for Positive and Negative controls from this pool. Also, slow evolving gene would likely be similar between species.

The PCR products of the slow evolving genes were sequenced (data not shown) and we chose the primer pairs based on the quality of the sequence. As indicated in the methods, we chose as positive control the sequences Tb40119 and Tb9898 and as a negative control a fragment of Histone 3. The primer pairs selected were Tb40119.2, Tb9898.2 and H3.1 respectively.

**COT 1 DNA library preparation**

The isolation of COT1 DNA was successful. The amount of fragmented DNA that was used for the denaturation/renaturation experiment was 4.31 μg. After the S1 nuclease treatment it was 1.031 μg, which corresponds to the 23.9% of the original amount.

For humans and other vertebrates the percentage of COT1 DNA present in the genome is around 10% (Singhal *pers. comm.*). However, it is known that for spiders there are many regions of repetitive (Croucher *pers. comm.*) and low complexity DNA (Brewer et al., 2014). For this reason it is not as surprising that the amount of COT1 DNA is relatively high. The library preparation for the isolated COT1 DNA was also successful, because of the application of amplification primers for the PCR reactions (Fig 5).

**Genomic library preparation**

The 116 libraries were successfully prepared and indexed. The fragment distribution of each library was checked on an agarose/TBE gel and the concentration determined using a NanoDrop.

**Hybridization Experiment: Post-Capture controls**

*Enrichment and depletion*

The effectiveness of the enrichment for the sequences present in the target was verified by measuring a shift towards less number of cycles on the enrichment curve for a captured sequence (Positive control). The opposite outcome is expected for the negative control. In Experiment 1 there was a shift of 5 cycles in both positive controls (Fig. 6a and 6b), which means that there was an enrichment of the captured sequences. Depletion

was a little more than two cycles for the Negative control (Fig. 6c). The negative control for the qPCR starts enrichment at cycle 26, which is an expected result given the sensitivity of the technique. However, it does not affect the previous interpretation of our results, as the enrichment in the positive controls starts around the cycle 12. In all the cases the replicates are similar with overlapping their curves.

In Experiment 2 the positive controls also show a shift in approximately 5 cycles (Fig. 7a and 7b). The depletion in Experiment 2 is less pronounced than in Experiment 1. It is less than a full cycle. The negative control for the qPCR shows enrichment around cycle 26 (Fig 7d). This is the same as in Experiment 1. In Experiment 2 the enrichment plots for the qPCR replicates have more variation than in Experiment 1, but it does not make any difference in the interpretation.

In Experiment 3, the positive controls also show a shift of approximately 5 cycles (Fig. 8a and 8b) and the negative control a shift of 1 cycle (Fig. 8c) Note that all the replicates, except Pre-Capture Tb40119, show overlapping curves. The negative control for the qPCR showed no enrichment (Fig 7d).

The results of the three hybridization experiments are consistent, corresponding to enrichment for positive controls of approximately 5 cycles and depletion for negative controls in 1-2 cycles. This indicates a successful capture of the targeted genomic DNA.

*Bioanalyzer*

In order to verify that the fragment distribution after hybridization and whole library amplification was suitable for sequencing we performed a Bioanalyzer analysis of each experimental outcome.

For Experiment 1 the fragment distribution ranged between 200 bp and 400 bp, with a peak in 267 bp (Fig. 9). There were no extra peaks at smaller or higher molecular weights. This implies that the size distribution is adequate for sequencing.

For Experiment 2 there is also a fragment distribution between 200 bp and 400 bp, with a peak on 245 bp (Fig 10a). However, there was a second spike at 129 bp. It corresponds to primer dimers that were produced as a consequence of the high number of PCR cycles of the Post-Capture amplification. These fragments were eliminated with a bead cleaning. After this treatment the fragment distribution was adequate for sequencing (Fig. 10b).

The fragment distribution for Experiment 3 also ranged between 200 bp and 400 bp with a peak at 244 bp (Fig. 11a). But the amount of DNA recovered was very small in comparison to Experiment 1 and 2. Correlated with the difficulty in detecting DNA after whole library amplification. As a consequence of the large number of PCR cycles the same extra peak at 129 bp was detected. It was also eliminated by an extra bead cleaning step, producing a fragment distribution amenable for sequencing (Fig. 11b).

The bead cleaning for removing primer dimers and cleaning up product for sequencing was performed at the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley.

**Analytical pipeline**

The amount of raw data per experiment was: Experiment 1 = 31 Gb, Experiment 2 = 28 Gb and Experiment 3 = 35 Gb.

***Data analysis***

The preparation of the bam files as well as the application of ANGSD for the allele frequency estimation proceeded as indicated in the methods.

*Site Frequency Spectrum*

The Site Frequency Spectrum obtained for the different species is presented in Fig. 12. In all the cases it shows an exponential decay where each SFS category is smaller than the previous one. Different cut-off levels for the filter – SNP_val were applied to the three species (data not shown). After consulting with colleagues (Lindroth *pers. comm.*) we decided to use the data set without filtering. The number of sites per species are: *T. brevignatha:* 1,224,769; *T. macracantha*: 1,912,875 and *T. waikamoi*: 1,199,495. The reason why the number in *T. macracantha* is larger than the original target (1.7 Mb) is due to the capture of intron sequences.

*T. brevignatha*

The first two Principal Components (PCA) for all the samples of *T. brevignatha* show a clear partitioning into two main groups (Fig. 13a). The separation between the two islands occurs along PCA1 (19.10% of the variation). The populations from the Big Island are differentiated mostly along the PC2 (8.69% of the variation). When plotting PC1 (19.10%) with PC3 (2.86%) most of the variation of the Big Island is present on PC1. However, there are 2 specimens from Maui that separate from the rest of the population. This was not evident on Fig. 13a. This situation is supported by plotting PC2 (8.60%) and PC3 (2.86%) (Fig. 13c). Note that in this case the Maui population is clustered with Mauna Kea.

The PCA that includes only the samples from Big Island has a greater spread in the distribution of the different specimens (Fig 14). There is a clear grouping by volcano with little structure within volcanos. All the volcanoes on the windward side are very similar to each other, although they do not overlap the space where their specimens are located. That is not true for the sites on the Leeward side (Honoaula and Kona Hema). These two sites group together and are actually separate from the rest of the specimens.

For the calculation of Fst we used a total of 550,767 sites present in all the populations. The number of shared sites is less than half of those originally present in the data set (1,224,769) because of the high divergence between populations on both islands. The values of the pair-wise comparisons are presented in Table 5. The largest pair-wise comparisons are between populations of the Big Island and Maui. These values are 2 to 3 times larger than most of the comparisons within the Big Island. Within the Big Island, samples the highest pair-wise Fst values are found between populations from the

Leeward and the Windward side of the island (Honoaula/Kīpuka: 0.182; Ka'u/Kīpuka: 0.168; Kona Hema/Laupāhoehoe: 0.169 and Kona Hema/Kīpuka: 0.219).

The highest nucleotide diversity (Table 6) is present in the Maui population (0.00459). It has more diversity than the entire Big Island combined (0.00279). Considering individual populations within the Big Island, the most diverse is Ka'u (0.00261) followed by Kona Hema (0,00252). The population with the lowest nucleotide diversity is Kīlauea (0.00190).

In terms of signals for demographic expansion (Table 6), all the populations show negative values. In particular, there is a two-fold increase in the estimation for the Big Island (-2.070) in comparison to Maui (-0.903). On the Big Island, the volcanoes with the more negative values are Mauna Loa (-1.737) and Mauna Kea (-1.426).

*T. waikamoi*

The first two Principal Components (PC1: 17.45% variance and PC2: 11.19% variance) do not show a strong separation between the two populations of *T. waikamoi* (Fig. 15a). The same pattern appears when plotting PC1 (17.45%) vs PC3 (9.74%) (Fig. 15b) or PC2 (11.19%) vs PC3 (9.74%). The Fst between both populations was 0.260.

The highest nucleotide diversity (Table 7) is present in Upper Waikamoi (0.00373). Related to signatures of population expansion, both sites present negative values that are relatively close (Table 7), making it difficult to draw any conclusion in terms of demographic processes.

*T. macracantha*

On the first two PCA (PC1: 33.93% and PC2: 6.59%) it is possible to identify 3 different groups (Fig. 16a): Lana'i, most of Kīpahulu Valley and Ko'olau plus three samples from Kīpahulu Valley (2982, 3022 and 3026), These three samples are from the locality "Kipa up" on Kīpahulu Valley. Note that even if they cluster with the other specimens from Ko'olau they are slightly separated. The group from Ko'olau and the specimens from Lana'i do not separate on the PC1. This becomes more evident by plotting the PC1 (33.93%) and PC3 (3.38%). On this case Lana'i and the Ko'olau appear cluster together and the three specimens from Kīpahulu separate along the PC3 (Fig. 16b). Finally, plotting the PC2 (6.59%) and PC3 (3.38%), shows that the differentiation between Maui and Lana'i occurs mostly in the PC2, while within Maui most of the differences are explained by PC3.

We were interested to see in more detail the relative distribution of the specimens from Lana'i, Ko'olau and the three specimens from Kīpahulu that cluster next to the Ko'olau specimens. They do not differentiate much along the PC1 (Fig. 16a). Given that the PCA is influenced by extreme values, we decided to do another PCA including only those specimens. In this more detailed analysis the specimens from Ko'olau appear separated from the ones from Lana'i and Kīpahulu. The plot of PC1 (21.68%) vs PC2 (11.51%) (Fig. 17a) is very similar to PC1 (21.68%) vs PC3 (10.68%) (Fig. 17b). With the PC2 (11.51%) and PC3 (10.68%) it is possible to see that the variability of the Ko'olau samples is larger than the ones from the other sites (Fig. 17c).

For the Fst calculations we used the groups identified by the first two components of the variance (PC1+PC2=40.52%). It included a total of 1,624,513 sites. The highest values of Fst are between Ko'olau or Lana'i with Kīpahulu (Table 8). They also correspond to the highest values of Fst reported on this study (approximately 0.5). The comparison between Ko'olau and Lana'i includes the highest values for the other species.

The population with the highest nucleotide diversity is Ko'olau (0.00392) followed by Lana'i (0.00296) (Table 9). In terms of signatures of demographic processes all the populations had values for Tajima's D closer to -2.0, so it is difficult to make statements related to the directionality of colonization (Table 9).

*All the species combined*

In order to analyze all the species together we had to repeat the mapping step and produce new bam files. The first two Principal Component (PC1: 40.70% and PC2: 6.61%) show the presence of four well-defined groups (Fig. 18a). *T. brevignatha* is split into two groups: Maui and Big Island. On this PCA it is not possible to see the break between Leeward and Windward populations. The third recognizable group includes most of the *T. macracantha* specimens. In particular the populations from Lana'i, Ko'olau and the three specimens from Kīpahulu Valley that cluster with Ko'olau. The fourth group includes *T. waikamoi*, most of the Kīpahulu population of *T. macracantha* and a bit less clustered, but still close, the two specimens from a "divergent lineage" of Maui *T. brevignatha*. Note that the group with the Maui population of *T. brevignatha* includes one specimen of *T. waikamoi* from Upper Waikamoi and the group with most of *T. macracantha* has one specimen from the Pu'u Kukui population of *T. waikamoi*.

The plotting of the PC1 (40.70%) and PC3 (5.65%) (Fig. 18b) is very similar to PC1 vs PC2. Actually, it looks like PC1 vs PC2 but inverted. Finally, the PC2 and PC3 (Fig. 18c) show a very similar pattern the distributions present in PC1 vs PC2. The only difference is that the group with the Big Island population of *T. brevignatha* appears closer to the group that includes specimens from the three nominal species.

To have a better idea of the relative distribution of the specimens present in the group with representatives of the three nominal species we did a PCA with only these samples. The first two PC (PC1: 24.54% and PC2: 3.91%) (Fig. 19a) show a clear separation of the two populations of *T. waikamoi*. The Upper Waikamoi population of *T. waikamoi* is not very differentiated on the PC1 from the Kīpahulu population of *T. macracantha*. One of the two specimens of *T. brevignatha* lies between these two groups. The plot of PC1 (24.54%) and PC3 (3.47%) (Fig. 19b) does not show any new information. Finally, in the plot of PC2 (3.91%) vs PC3 (3.47%) (Fig. 19c) the separation of the two populations of *T. waikamoi* is less evident, but the grouping of *T. macracantha* still present.

*Admixture detection*

To investigate the presence of hybridization we used the program ngsADMIX (Skotte et al., 2013). This program is designed to find admixture proportions of NGS data using genotype likelihoods as input. We included all the specimens in the analysis and

tested for different numbers of ancestral populations from k=2 until k=10. We did 10 replicates per k value with different random seeds.

Regular methods of model selection (AIC or BIC) cannot be used here, as the models are not nested. We implemented the Evanno method (Evanno et al., 2005) to select the ideal value of k. As a result of this implementation, we found that k=2 was the optimal clustering (Fig 20a). However, after discussing with the author of the program (Skotte *per. comm.*) and others (Crawford and Nielsen *per. comm.*), they indicate that ngsADMIX was not developed to identify an optimal number of clusters. They indicate that once independent runs have problems reaching convergence it was indicative that the k values used were too large. However, it is not possible to select among lower k values, where there is consistent convergences. For our data the values of k where this situation occurs are from k=2 up to k=5.

For k=2 (Fig. 21a) the populations of *T. macracantha* (Kīpahulu Valley, Ko'olau and Lana'i) and *T. waikamoi* (Upper Waikamoi and Pu'u Kukui) are grouped together (red). The other group (green) corresponds to the Big Island populations of *T. brevignatha*. All the individuals of the Maui population of *T. brevignatha* (Lower Waikamoi) present apparent levels of admixture. However, note that when the admixture proportions are 50:50 it is possible that this is an artifact of a small k value. Also, note that there are two specimens that are completely red. They correspond to the previous identified divergent lineages of *T. brevignatha* (Fig. 18a).

For k=3 (Fig. 21b) most of the specimens from the Maui population of *T. brevignatha* form a new group. The two specimens from the divergent lineage remain as part of the group that includes *T. waikamoi* and *T. macracantha*. Also, one specimen from Lower Waikamoi (*T. waikamoi*) is included in the same group as the other Maui *T. brevignatha*. This specimen was the same one that clustered with the Maui population of *T. brevignatha* in the PCA (Fig. 18a). It is also possible to identify low levels of admixture (<20%) in some specimens, but this is an artifact as it disappeared in the next grouping (k=4).

When 4 ancestral populations are selected (k=4) (Fig. 21c), the group consisting of mostly *T. macracantha* is recovered (orange). Note that some of the specimens from Kīpahulu Valley remain in the red grouping with *T. waikamoi* and the two divergent specimens of *T. brevignatha*. For k=5 (Fig. 21d), the populations from the older volcanoes of the Big Island (*T. brevignatha*) form a new group. In particular the specimens from Mauna Loa (Kīpuka) are unequivocally assigned to this new group, while the ones from Mauna Kea (Laupāhoehoe) suggest admixture. One specimen in the Pu'u Kukui population appears to belong to the same group as most of the *T. macracantha* with > 80% probability. This specimen corresponds to the same that clusters with this species on the PCA (Fig. 18a). In k=6 (Fig. 21e) the two Leeward populations of *T. brevignatha* on the Big Island form a separate group. Note that Ka'u (southwest population) appears as an admixed population.

The separation between the Haleakalā (Upper Waikamoi) and West Maui (Pu'u Kukui) populations for *T. waikamoi* is only seen at k=7 (Fig. 21f). Then, at k=8 (Fig. 21g) the *T. brevignatha* populations of Pu'u Maka'ala (Kīlauea volcano) and Ka'u form a separate group. Note that since k=6 Ka'u appeared as an admixed population. In k=9 (Fig. 21h) six specimens from the Kīpuka population separate to a different group.

Finally, in k=10 (Fig. 21i) some admixture elements appear in the Upper Waikamoi population and one specimen from the Lower Waikamoi is distinct in a new group.

It is important to emphasize that the clearest convergence of the analysis occurs in the k values from 2 to 5. For highest values of k, the simulation with the highest likelihood value, but in these cases more than one likelihood value appears as a result of the different repetitions. These results show that up to k=8, every new group that appears by increasing the value of k corresponds to a geographically isolated population. In terms of convergence, k=4 presents more variability than k=5, which is unexpected. However, the author of the program mentioned that she has seen this before and it is not something to be concerned about (Skotte *per. comm.*).

These sequential grouping schemes can be visualized as slices on a dendrogram (Fig. 22). However, this is not a phylogenetic reconstruction as groups that cluster together can do so because of population size or factors other than common ancestry (Skotte and Nielsen *pers. comm.*)

*Phylogenetic Network*

The phylogenetic relationships among the different populations are shown based on a phylogenetic network generated by the SplitsTree 4.13.1 (Huson and Bryant, 2006). 116 terminals were included and a total of 8,397 sites were used. Note that this is a significantly smaller number than the sites used for the other low coverage applications. The main reason for this is the filters applied for this particular application.

The phylogenetic network consists in 359 splits events and there are 5 clearly recognizable groups (Fig. 23). The colors represent the nominal description of each species (see legend Fig. 23).

At the bottom of the network are all *T. brevignatha* from the Big Island. There is a clear partition into two groups. On the left side are the specimens from the two oldest windward volcanoes (Mauna Loa and Mauna Kea). On the right side are the specimens from Leeward and the young Kīlauea. Within this clade it is possible to see most of the specimens from Honoaula (Hualālai) on a separate branch as well as the ones from Pu'u Maka'ala (Kīlauea). The specimens from Ka'u and Kona Hema appear from the base of this clade.

Another very distinctive group is the Maui population of *T. brevignatha*. All the specimens were collected at the same site and there is no clade structure within. As part of this group it is also possible to identify one specimen from *T. waikamoi*.

The other big group consists of most of the specimens of *T. macracantha*. Within this group there is a very clear break between the specimens from Maui (Ko'olau and a few from Kīpahulu Valley) and Lana'i. Inside the Lana'i clade there is one specimen from Pu'u Kukui of *T. waikamoi*.

The fourth well-defined group includes all the specimens from the Pu'u Kukui population of *T. waikamoi*. Finally, there is a clade (WaikMacBrev) that includes specimens initially described as different species. Note that this clade shares a common stem with the group of *T. waikamoi* from Pu'u Kukui. Inside WaikMacBrev there is no a clear structure related to the original nominal species. Different nominal species are intercalated between each other. The branch length of each terminal after the last split corresponds to individual apomorphies.

## DISCUSSION

*Implementation of Exon Capture on a spider genome*

The present work is the first to use transcriptome-based Exon Capture in arachnids. The approach successfully produced the data to address our questions. Particularities of the spider genome and the resources available made it necessary to use some modifications to overcome technical complications. These issues will be discussed and considered for future experiments.

The first technical complication was due to the absence of commercial COT1 DNA for arachnids. It is available for humans (Life Technologies), mouse (Life Technologies), salmon (Applied Genetics Laboratories), etc, but not for spiders. We therefore had to prepare our own library for the genus using *T. quasimodo*. Besides the time investment, it had the risk of introducing errors in the experimental procedure. A mistake in the calculation of the re-annealing times could produce more complex DNA or overlook the COT1 in the genome. The first issue has the risk of blocking sequences that are part of the target, while the second is to have a less efficient blocking of the low complexity sequences present in the genomic libraries.

The low complexity DNA by definition corresponds to "simple" sequences. These sequences are more likely to evolve by random chance than regulatory or protein coding regions. For this reason the use of closely related species COT1 DNA should not be a problem. As to how far it might be possible to apply COT1 DNA from other species is something that has not yet been explored extensively. Sonal Singhal, UC Berkeley, has successfully used a mix of homemade lizard COT1 DNA and commercial chicken COT1 DNA (*pers. comm.*). In the same way, Ke Bi, UC Berkeley, has combined homemade COT1 DNA for chipmunks with commercial COT1 DNA for mouse (*pers. comm.*). These and other experiences are encouraging with regards the possibility of using COT1 DNA across species. However, the jump from vertebrates to invertebrates has yet to be done. Ways to indirectly test this would be very useful for streamlining the experimental protocol.

There are also two issues of the library preparation that could be improved in future experiments. The first one is related to the sonication of the DNA. It was necessary to use long cycles (3.5 minutes continuous) and manually change the water with ice from the sonicator machine. This was very time consuming and due to the random sizes of the ice reduces the reproducibility of the obtained fragment distributions. Also, these prolonged fragmentation cycles affected the material integrity of the fragmentation tubes produced by the manufacturer (Diagenode). For this reason we had to use other tubes that were not as consistent in transmitting the ultrasound waves.

The implementation of an automatic water cooler system (Diagenode) and the reduction of the continuous ultrasound exposure to 30 seconds per cycle (manufacturer tubes do not degrade), improves enormously the reproducibility of the fragments obtained.

The problem with high variation in the obtained fragment sizes is mostly related to the samples with smaller fragment sizes. The smaller fragments will outcompete large

fragments during the hybridization. For this reason it is possible that the final result of the capture is biased towards samples with smaller fragments.

Also related to small fragment sizes is the issue of the ratio used for the bead cleaning. Our implementation (1.8x) followed the recommendations of Meyer and Kircher (2010). However, other researchers have found that by using smaller ratios (1.3x) it is possible to eliminate smaller fragments without compromising the library (Smith and Bi, *pers. comm.*). This is useful, as very small fragments, even if sequenced, cannot be used for the final analysis. This at the end only reduces the coverage of the fragments of interest.

Another technical aspect to consider for future experiments is the type of capture used. We realized a chip based capture, which costs approximately $703 per unit (reference prize from October 20013, Agilent Technologies). For each chip we multiplex approximately 50 individuals. So, the total cost was around $2,100. The alternative would have been to do an in solution capture. Currently the company MYbaits offers in solution kits for 96 reactions at $8,640 and $13,440 for 192 reactions (MYbaits-1 (20,000 for Maximum number of bait sequences)). These prices include the probe design, which is one of the reasons why it is much more expensive.

In general in solution captures are more efficient than a chip, which gives more specificity and higher coverage. They also allow a larger target size, 50 Mb as a maximum (NimbleGen, Roche) in comparison to, for example, 1,7 Mb that we used in our chip-based hybridization experiment. However, this is highly dependent on the tiling density. This is particularly beneficial in the case of studies that look for signatures of genome wide selection. The problem is that for smaller projects (less than 150 samples) it is too expensive. Only for larger projects does it become cost effective.

For our question the use of a chip was ideal for two main reasons. The first is that the number of specimens was right at the margin of the cost efficiency for in solution capture. The second and more important is the fact that this was the first implementation of an Exon Capture for spiders, so it was necessary to do test experiments before engaging in a larger number of reactions. Only once we corroborate that the capture worked in the first experiment did we moved forward to acquiring two new chips. The selection of an in solution capture format for future experiments will depend on the size of the project and the target size captured.

The final technical aspect to consider corresponds to the reduced capture efficiency of the Experiment 2 and 3. The first indication of reduced capture efficiency corresponded to the high number of PCR cycles for whole library amplification that were required. Then, it was confirmed by the high number of duplicates removed by PICARD and consequently low coverage of these experiments with respect to Experiment 1 (Table 4). We have three alternatives explanations for this result.

The first issue is related to the fact that exon capture was not as efficient on species that differed from the taxon used for probe design (*T. brevignatha*). This can result by high sequence divergence in nuclear coding genes targeted for capture. These organisms live on islands and may undergo periodic bottlenecks associated with colonization events. Also, they have a relatively short generation time (approximately 1 year). Lastly, the fact that they are part of an adaptive radiation could also have an effect if the coding genes under strong selection. For all these reasons it is possible that the exon sequences captured could have shown excessive divergence.

This explanation requires several assumptions. Chipmunk data from the same lab shows that for species with more than 20 MY of divergence there are not differences in efficiency (Bi *pers. comm.*). Also, the fact that the efficiency (measured as coverage) of the *T. brevignatha* libraries was very similar to the other two species on Experiment 2. This suggests an "Experiment wide" effect, instead of something related to a particular species.

Our second issue concerns COT1 DNA library preparation. In particular, it is possible that we included not only the COT1 portion, but also some coding sequences. The experiment with the lowest capture efficiency was Experiment 3, which includes *T. anuenue*. This species, not discussed here, is the closest relative to the species utilized for the COT1 DNA library preparation (*T. quasimodo*). Thus, it is possible that the COT1 DNA may have blocked part of the genomic library. This explanation, leaves in question why the effect was also strong in Experiment 2 (*T. brevignatha, T. waikamoi* and *T. macracantha*) with respect to Experiment 1. Another possibility is the existence of manufacturing defects on the chip. This could be tested only by repeating the experiment with a new chip.

Of all the technical aspects discussed for this method, the last one is the most concerning. Although it is difficult to know what factor or combination of factors account for the variability in the capture efficiency, it does provided a technical basis on which to build for future applications.

*Cryptic diversity after a speciation event*

In the SplitTree reconstruction of 116 samples it is possible to recognize 5 clades: the Big Island population of *T. brevignatha*, Maui population of *T. brevignatha*, most of the specimens of *T. macracantha* (Lana'i, Ko'olau and three specimens from Kīpahulu Valley), Upper Waikamoi population of *T. waikamoi* and a clade that we informally named WaikMacBrev. This last one consists of the Pu'u Kukui population of *T. waikamoi*, most of the Kīpahulu Valley specimens of *T. macracantha* and the two very "divergent" specimens of *T. brevignatha*. To this arrangement we include one specimen of the Pu'u Kukui population of *T. waikamoi* grouping with *T. macracantha* and one from Upper Waikamoi grouping with the Maui population of *T. brevignatha*.

A similar grouping pattern is found in the PCA analysis and on k=4 of the ngsADMIX reconstruction. The main difference is seen in *T. waikamoi*, which only separates its populations when there is a PCA exclusive for WaikMacBrev and at k=7 for the ngsADMIX analysis. Clades of Upper Waikamoi *T. waikamoi* and WaikMacBrev share a longer branch in the SplitsTree reconstruction, than with any other clade. For *T. brevignatha* on the Big Island and at k=5 for the ngsADMIX analysis shows the break between the old volcanoes on Windward (Mauna Loa and Mauna Kea) from the rest of the populations. Note that at k=5 there is a change in the slope of the increase in the average likelihoods of the runs (Fig. 20b). This has been suggested by Evanno et al. (2005) as an indication of a way to detect the "optimal" number of ancestral populations. This break between old windward volcanoes and the rest of the sites is not detectable on the PCA, but it is on the SplitsTree analysis (Fig. 23).

This general congruence between analytic methods is expected given they all used the same data set, however it is remarkable that the very different approaches used to cluster specimens all converge on the same groupings.

For *T. brevignatha* we found a very strong genetic break between populations on different islands. That is shown in the PCA analysis (PC1 vs PC2) and very high Fst values (between 0.274 and 0.351) in comparison of the Big Island with Maui. In terms of nucleotide diversity, Maui is more diverse than the entire Big Island data combined (Table 6). The genetic difference between these two islands is as large as what is found in the SplitsTree reconstruction between distinct or nominal species. Finally, the populations of both islands present negative values for Tajima's D, but the one on The Big Island is two times larger than the one on Maui (Table 6). This suggests a strong signal of population expansion. These are the same patterns found in the three mitochondrial markers (Cotoras, Chapter 2), and reinforces the scenario of an older population on Maui and a colonization event to the Big Island followed by a population expansion.

The estimation of Tajima's D was obtained by conducting a global Tajima's D across all whole, captured fragments. This estimation has to be considered as a coarse estimation. A global Tajima's D combines genes under different selection regimens. Thus, this estimation is influenced by coding regions undergoing directional selection (negative Tajima's D) and disruptive selection (positive Tajima's D) (Eytan and Hellberg 2010). This is an important aspect to consider as this transcriptome-based Exon Capture uses coding regions as the homologous sites for the capture. However, demography has a homogeneous effect across the whole genome and which is what we hope estimate by performing a global Tajima's D estimation. As a first approximation it is a good estimator. For detecting signals of natural selection it is necessary to have average estimations for the whole genome before examining particular genes. This is necessary in order to have a base line for the genome wide heterozygosity produced by demography. Our global Tajima's D approximates this genome wide estimation, which reveals information on demography.

In terms of the diversity within the Maui population, our analysis shows clearly the presence of two highly divergent specimens. It can be visualized when plotting the PC3 (2.86%) (Fig. 13b and c). These are the exact same specimens that were recognized as highly divergent using mitochondrial markers (Cotoras, Chapter 2). The fact that they are also recognizable as "outliers" in the Maui population in this approach indicates that there is also strong differentiation in the nuclear genome. So, they are not only mitochondrial divergent lineages but nuclear as well.

Closer examination of the Big Island population of *T. brevignatha* shows a well defined by volcano structure in PCA space (Fig. 14). Using nuclear markers each volcano is recognizable as a separate unit, but they are all next to each other in PCA space. The exceptions to this are the sites on the Leeward side of the island. Specifically, Honoaula (Hualālai) and Kona Hema (Southwest Mauna Loa). These two sites are clustered together and there is a clear separation from the other sites/volcanoes. This partition between the Leeward side and the rest of the volcanoes/sites was also detected with mitochondrial markers (Cotoras, Chapter 2). However, the fine partition between volcanoes did not appear in any of our previous analysis with mitochondrial DNA. In the ngsADMIX analysis the first break detected (k=5) within the Big Island is between the

older windward volcanoes (Mauna Kea and Mauna Loa) and the rest of the sites. Then at k=6, Kona Hema and Honoaula form a separate group. Ka'u appears as an admixed population between this group of Leeward sites and Kīlauea (Pu'u Maka'ala). Genetically, Kona Hema is closer to Honoaula even when the first locality is closer in distance to Ka'u. However, between Ka'u and Kona Hema there is a biogeographic barrier for rainforest. It consists on the southwest rift zone of Mauna Loa.

Within the Big Island the highest nucleotide diversity is found in Ka'u (0.00261) and Kona Hema (0.00252), while the lowest is on Kīlauea (0.00190). This pattern conflicts with our finding using mitochondrial markers (Cotoras, Chapter 2). Based on mitochondrial markers the most diverse population was actually Kīlauea, while Kona Hema was rather low and Ka'u, depending on the marker, was average. The result from the NGS data should be considered with caution as the nucleotide diversity is estimated indirectly from Theta Pi (Theta Pi / number of sites). The populations from Kona Hema and Ka'u are the ones with lower number of sites, so it is possible that their Nucleotide diversity is inflated.

Finally with regard to the population dynamics within the Big Island, it was possible to find a significant signature of population expansion for the mitochondrial COI and ND1 data on the Hualālai population (Fu's Fs and Tajima's D, respectively). However, for the nuclear markers there is no signal of population expansion on the Leeward sites. Indeed, the volcano with the most negative value is Mauna Loa. Again, this result should be considered only preliminary.

In summary, for *T. brevignatha* the nuclear as well as mitochondrial data supports colonization from Maui. Within the Big Island there is a population break between the Leeward and the Windward sites. Whole exome estimations for Tajima's D suggest a population expansion on the Big Island.

In the case of *T. waikamoi* none of the first three principal components of the variance separate the two populations in a clear manner (Fig. 15). This was unexpected as for the three mitochondrial markers there is a very strong separation. Only one haplotype is shared among these three genes (Cotoras, Chapter 2). The estimated Fst is relatively high (0.260), however it could be due to the overdispersion of the samples. The values for Tajima's D are relatively similar for both volcanoes. Thus, it is difficult to interpret the source population.

In the combined PCA of all specimens (Fig. 18a), most specimens of *T. waikamoi* appear as a single cluster, which is congruent with the lack of separation detected in the PCA of this species alone (Fig. 15). However, in the individual PCA for WaikMacBrev, the separation by volcano for *T. waikamoi* becomes very evident (Fig. 19). Another feature of this PCA is that the *T. waikamoi* PCA (Fig. 15) has all specimens from Upper Waikamoi (Haleakalā) clustered in three very tight groups. They are so tight, that in practice they are stacked on top of each other. As the PCA are driven by the extreme values it is possible that this situation is simply an artifact of the analysis. This idea is supported by the fact that in the PCA with specimens from the other species this pattern does not occur. In the ngsADMIX analysis the two volcanoes separate as different populations only at k=7 (Fig. 21f).

Because the populations from the two volcanoes are distinguishable on the PCA along with other species and also in the ngsADMIX analysis, we consider this evidence that there is a genetic break between both volcanoes. However, this genetic break must be

rather recent as the signal in the exome is not very strong. This contrasts with the apparent genetic separation that is found in the mitochondrial genome (Cotoras, Chapter 2). This discordance suggests that the separation between the two populations is too recent to have left nuclear signalture, but old enough to evolve different mitochondrial haplotypes at each site.

In the case of *T. macracantha*, the PC1 (33.93%) in the PCA shows a clear separation between two groups. One that includes the populations of Lana'i, Ko'olau and three specimens from Kīpahulu Valley and another that has most of the specimens from Kīpahulu Valley. The first group clusters as an independent clade on the SplitsTree reconstruction, while the second (most of Kīpahulu Valley) is part of WaikMacBrev. In the SplitsTree it is possible to see the differentiation between the specimens from Lana'i and Ko'olau. Three of the four specimens from Lana'i were collected at the beginning of the 90's (RG24, RG25 and RG26) and the last one in June 2013 (3277). Interestingly the specimen collected in 2013 clusters with two of the old ones, while RG25 appears more separated. This could suggest the presence of some kind of variation within Lana'i. However, it is difficult to be certain due to the low sampling.

For the Fst estimation we group together the Ko'olau population with the specimens from Kīpahulu Valley that appear next to it on the PCA analysis. The Fst values of Ko'olau or Lana'i compared to Kīpahulu Valley are the highest recorded in this study (> 0.5). These levels of divergence are only found in between species comparisons and between very divergent species. Given the high divergence within Kīpahulu Valley we did not do a between islands comparison mixing in the same group from all Maui sites. Neither Tajima's D nor the degree of nucleotide diversity provide conclusive evidence for the directionality of the colonization. Note that the nucleotide diversity of the Kīpahulu population is lower than the other two sites.

The fact that specimens from three different nominal species cluster together is unexpected, especially because their separation as a distinct group is very clear. Initially we suspected that it could be an artifact of specimens with low coverage grouping together. However, that is not the case. After checking the coverage of the different specimens present in WaikMacBrev we did not find a clear pattern (Table 10). This situation adds a new element to consider related to the study of newly evolved species. So far, two species of the same ecomorph have not been found to occur in sympatry. However, our data shows the existence of genetically distinct lineages of the green ecomorph present in the same area. In other words, the morphological diversity of the other species is contained within the morphological diversity of WaikMacBrev.

Something similar happens with the specimens of *T. waikamoi* that cluster with the Maui population of *T. brevignatha* and *T. macracantha*. The fact that on SplitsTree *T. waikamoi* from Upper Waikamoi has a common stem with WaikMacBrev, suggests that its separation into clades is more recent than any other nominal species. It might be even more recent than the separation between Maui and Big Island populations of *T. brevignatha*.

A deeper morphological study of these species will be required to determine whether there are any characters that unite them or whether this is cryptic diversity only detectable with molecular methods. Biologically it adds a new layer of complexity as a recent speciation event with divergent lineages with convergent morphologies. These

divergent lineages show no admixture with other species of the same ecomorph (ngsADMIX analysis).

On the other hand, the clade where these three nominal species are present has two other species corresponding to different ecomorphologies (*T. restricta*: small brown and *T. kamakou*: maroon). What would happen if we include specimens of these species to this data set? *T. restricta* is present on Maui and the Big Island and it would be interesting to know if it would show the same degree of genetic differentiation as *T. brevignatha.* In the case of *T. kamakou* it has a wider distribution including the two volcanoes of Maui, Lana'i and Moloka'i. How structured are these populations? Moreover, what would be the effect of including these taxa in a phylogenetic analysis with other green ectomorph species?

To understand the genomics that accompany ecological speciation the addition of these two species is essential, as they co-exist in sympatry with the three green ecomorph, and are sister to *T. waikamoi*. Which genes are driving the change in ecomorphology? Is it possible to find signatures of selection in the exome? Finally, would the WaikMacBrev include specimens from these other two species? If that were the case, we would be in the presence of a "chameleon" clade with a very high phenotypic plasticity. Its better understanding could give insights about the allopatric and sympatric process of speciation within an adaptive radiation

## CONCLUSION

The adaptive radiations present in the chronosequence of the Hawaiian Islands provide excellent systems to study different aspects of the temporal dynamic of diversification. Here we examined genetic signatures of young speciation events by focusing on three closely related species (*T. brevignatha, T. waikamoi* and *T. macracantha*) from the middle-aged island group of Maui Nui. To determine which demographic and evolutionary processes explain their current distribution we used a transcriptome-based Exon Capture approach.

By combining all the data, different analysis (SplitsTree, PCA and ngsADMIX) showed the existence of 5 well defined clades: *T. brevignatha* Big Island, *T. brevignatha* Maui, *T. macracantha*, *T. waikamoi* and WaikBrevMac. The last one includes specimens from the three nominal species. There are also specimens from *T. waikamoi* included on the clusters of other species, in particular *T. macracantha* and the Maui population of *T. brevignatha*. This is an unexpected result, which adds more complexity to the understanding of young speciation events.

Although morphologically identical, the molecular divergence between the *T. brevignatha* populations of Maui and Big Island are similar to "species level" comparisons. Also, there are two strongly divergent specimens present from the Maui population, which are included in the WaikBrevMac clade. For the Big Island population, there is a strong break between Leeward and Windward. On the other hand, for *T. waikamoi*, the nuclear data suggests that the population break between Haleakalā and West Maui is less pronounced than it appears with mitochondrial data. This could be an indication of a recent population split, but old enough for the evolution of different mitochondrial haplotypes. Finally, we found a separation between the Maui (Ko'olau and

some specimens from Kīpahulu Valley) and Lana'i populations of *T. macracantha*. More interestingly, there are two very divergent lineages co-existing in Kīpahulu Valley.

The future addition of two more species present in the same clade (*T. restricta* and *T. kamakou*) is critical in order to generate insights on the genetic bases of recent speciation events. in this case, as the species correspond to two different eco-morphologies, it will add information about the genomics of ecological speciation. This, in combination of the present study focused on allopatric speciation, will provide a more complete picture of the earlier stages of population differentiation and species formation.

## ACKNOWLEDGEMENTS

The authors would like to acknowledge a large number of people and institutions that collaborated at different stages of this research. The fieldwork in Hawai'i was supported by Laura Arnold, Timothy Bailey (HALE), David Benítez (HAVO), Katie Champlin (Limahuli Botanical garden), James Friday (UH Mānoa), Emory Griffin-Noyes (Limahuli Botanical Garden), Faith Inman-Narahari (UH Mānoa), Darcey Iwashita (UH Mānoa), Raina Kaholoaa (HALE), Susan Kennedy (UC Berkeley), Jessie Knowlton (Michigan Tech), Rick Lapoint (U Arizona), Scott Laursen (UH Mānoa), Karl Magnacca (O'ahu Army Natural Resources Program), Elizabeth Morrill (SFSU), Patrick O'Grady (UC Berkeley), Rita Pregana (HAVO), Donald Price (UH Hilo), David Rankin (U Vermont), William Roderick (Stanford), Andrew Rominger (UC Berkeley), Karen Uy (UH Hilo), Erin Wilson (UC Riverside) and the Kīpuka team.

The permit processing and access to different reserves and private land was possible thanks to Steve Bergfeld (DOFAW Big Island), Pat Bily (TNC) Maui, Tabetha Block (HETF), Shalan Crysdale (TNC Big Island), Lance DaSilva (DOFAW Maui), Dean Danae (Kahoma Ranch), Charmian Dang (NAR), Melissa Dean (HETF), Betsy Gagne (NAR), Elizabeth Gordon (HALE), Lisa Hadway (DOFAW Big Island), Paula Hartzell (Lana'i Resorts, LLC), Greg Hendrickson (Kealakekua Ranch), Mel Johansen (TNC Big Island), Pomaikaʻi Kaniaupio-Crozier (Maui Land and Pinneapple), Cynthia King (DLNR), Peter Landon (NAR Maui), Rhonda Loh (HAVO), Russell Kallstrom (TNC Moloka'i), Joey Mello (DOFAW Big Island), Ed Misaki (TNC Moloka'i), Elliot Parsons (Pu'u Wa'awa'a HETF), Lani Petrie (Kapapala Ranch), Shawn Saito (Parker Ranch), Joe Ward (Maui Land and Pinneapple) and Kawika Winter (Limahuli Botanical Garden).

We also appreciate the advices on lab and analytical work of Ke Bi (UC Berkeley), Michael Brewer (East Carolina University), Jacob Crawford (UC Berkeley), Peter Croucher (Illumina), Emiliano Méndez (Universidad de Magallanes), Rasmus Nielsen (UC Berkeley), Tyler Linderoth (UC Berkeley), Sonal Singhal (UC Berkeley), Line Skotte (U Copenhagen), Lydia Smith (UC Berkeley), LindLab, EvoLab. Barker DNA Sequencing Facility (UC Berkeley) and Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley (supported by NIH S10 Instrumentation Grants S10RR029668 and S10RR027303). Jonathan Price (UH Hilo) for discussions about the age of the wet forest on Leeward Big Island and Amy Vandergast (USGS) for advices at the beginning of the project. George Roderick (UC Berkeley) and Charles Marshall (UC Berkeley) also contributed with constructive commentds on earlier versions of the manuscript.

Darko Cotoras' PhD program has been founded by a Fulbright/CONICYT

fellowship and a researcher positon on a NSF Dimensions of Biodiversity Project. The fieldwork on Hawai'i was funded by Graduate Research Allocation Committee (Integrative Biology dept. UC Berkeley), Summer Research Grant (Integrative Biology dept. UC Berkeley), Walker Grant (Essig Museum of Entomology), Sigma Xi grant and Research Grant Graduate Division UC Berkeley.

## LITERATURE CITED

Ballard J.W.O. and Whitlock M.C. 2004. The incomplete natural history of mitochondria. Molecular Ecology 13: 729–744

Bandelt H.J. and Dress A.W.M..1992. A canonical decomposition theory for metrics on a finite set. Advances in Mathematics, 92:47–105

Bi K., Vanderpool D., Singhal S., Linderoth T., Moritz C. and Good J.M. 2012. Transcriptome-based exon capture enables highly cost-effective comparative genomic data collection at moderate evolutionary scales. BMC Genomics 13: 403

Blackledge T.A. and Gillespie R.G. 2004. Convergent evolution of behavior in a adaptive radiation of Hawaiian web-building spiders. Proceedings of the National Academy of Sciences USA 101: 16228–16233

Brewer M.S., Cotoras D.D., Croucher P.J.P. and Gillespie R.G. 2014. New sequencing technologies, the development of genomics tools, and their applications in evolutionary arachnology. The Journal of Arachnology 42: 1–15

Carson H.L. and Templeton A.R. 1984. Genetic revolutions in relation to speciation phenomena: the Founding of new populations. Annual Review of Ecology and Systematics. 15: 97-131

Carson H.L., Lockwood J.P. and Craddock E.M. 1990. Extinction and recolonization of local populations on a growing shield volcano. Proceedings of the National Academy of Sciences USA 87: 7055–7057

Catchen J., Hohenlohe P., Bassham S., Amores A. and Cresko W. 2013. Stacks: an analysis tool set for population genomics. Molecular Ecology. 22: 3124–3140

Croucher P.J.P., Brewer M.S., Winchell C.J., Oxford G.S. and Gillespie R.G. 2013. *De novo* characterization of the gene-rich transcriptomes of two color-polymorphic spiders, *Theridion grallator* and *T. californicum* (Araneae: Theridiidae), with special reference to pigment genes. BMC Genomics 14: 862

Davey J.W, Hohenlohe P.A., Etter P.D., Boone J.Q., Catchen J.M. and Blaxter M.L. 2011. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. Nature Reviews Genetics 12: 499-510

Dobzhansky T. 1951. Genetics and the Origin of Species 3rd edn Columbia Univ. Press, New York

Earl D.A. 2011. Structure harvester v0.6.1. Available at http:// taylor0.biology.ucla.edu/struct_harvest/. Accessed April, 2011.

Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology. 14: 2611–2620.

Eytan R.I. and Hellberg M.E. 2010. Nuclear and mitochondrial sequence data reveal and conceal different demographic histories and population genetic processes in Caribbean reef fishes. Evolution 64: 3380–3397

Feder J.L., Flaxman S.M., Egan S.P., Comeault A.A. and Nosil P. 2013. Geographic Mode of Speciation and Genomic Divergence. Annu. Rev. Ecol. Evol. Syst. 44:73–97

Fumagalli M., Vieira F.G., Linderoth T. and Nielsen R. 2014. ngsTools: methods for population genetics analyses from next-generation sequencing data. Bioinformatics Applications 30: 1486-1487

Fumagalli M., Vieira F.G., Linderoth T. and Nielsen R. 2014. ngsTools: methods for population genetics analyses from next-generation sequencing data. Bioinformatics. 30: 1486-1487

Gillespie R.G. 2004. Community assembly through adaptive radiation in Hawaiian spiders. Science. 303: 356-359

Gillespie R.G., Croom H.B. and Palumbi S.R. 1994. Multiple origins of a spider radiation in Hawaii. Proc. Nati. Acad. Sci. USA 91: 2290-2294

Gillespie, R.G. and Baldwin, B.G. 2010. Island Biogeography of remote archipelagos: Interplay between ecological and evolutionary processes. In: The Theory of Island Biogeography at 40: impacts and prospects. Losos JB, Ricklefs RE (eds). Princeton University Press.

Grabherr M.G., Haas B.J., Yassour M., Levin J.Z., Thompson D.A., Amit I., Adiconis X., Fan L., Raychowdhury R., Zeng Q., Chen Z., Mauceli E., Hacohen N., Gnirke A., Rhind N., di Palma F., Birren B.W., Nusbaum C., Lindblad-Toh K., Friedman N., Regev A. 2011. Full-length transcriptome assembly from RNA-seq data without a reference genome. Nat. Biotechnol. 29: 644-652

Hendry A.P., Bolnick D. I., Berner D. and Peichel C.L. 2009. Along the speciation continuum in sticklebacks. Journal of Fish Biology 75: 2000–2036

Hodges E., Rooks M., Xuan Z., Bhattacharjee A., Gordon D.B., Brizuela L., McCombie W.R. and Hannon G.J. 2009. Hybrid selection of discrete genomic intervals on custom-designed microarrays for massively parallel sequencing. Nature Protocols. 4: 960-974

Huson D.H. and Bryant D. 2006. Application of phylogenetic networks in evolutionary studies, Molecular Biology and Evolution. 23: 254-267

Huson D.H., Auch A.F., Qi J. and Schuster S.C. 2007. MEGAN analysis of metagenomic data. Genome Research. 17: 377-386

Kiialainen A., Karlberg O., Ahlford A., Sigurdsson S., Lindblad-Toh K. and Syvänen A.C. 2011. Performance of Microarray and Liquid Based Capture Methods for Target Enrichment for Massively Parallel Sequencing and SNP Discovery. PLoS ONE 6: e16486

Koressaar T. and Remm M. 2007. Enhancements and modifications of primer design program Primer3. Bioinformatics 23:1289-1291

Korneliussen T.S., Moltke I., Albrechtsen A. and Nielsen R. 2013. Calculation of Tajima's *D* and other neutrality test statistics from low depth next-generation sequencing data. BMC Bioinformatics 14: 289

Kuussaari M., Bommarco R., Heikkinen R.K., Helm A., Krauss J., Lindborg R., Öckinger E., Pärtel M., Pino J., Rodà F., Stefanescu C., Teder T., Zobel M. and Steffan-Dewenter I. 2009. Extinction debt: a challenge for biodiversity conservation. Trends in Ecology and Evolution 24: 564- 571

Li H., Handsaker B., Wysoker A., Fennell T., Ruan J., Homer N., Marth G., Abecasis G., Durbin R. and 1000 Genome Project Data Processing Subgroup. 2009. The Sequence alignment/map (SAM) format and SAMtools. Bioinformatics. 25: 2078-2079

Losos J.B. and Ricklefs R.E. 2009. Adaptation and diversification on islands. Nature 457: 830-836

Martin S.H., Dasmahapatra K.K., Nadeau N.J., Salazar C., Walters J.R., Simpson F., Blaxter M., Manica A., Mallet J. and Jiggins C.D. 2013. Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies. Genome Research. 23: 1817–1828

Mayr E. 1963. Animal Species and Evolution. Belknap Press, Cambridge, Massachusetts

Meyer M. and Kircher M. 2010. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harb Protoc.

Nei M., Maruyama T. and Chakraborty R. 1975. The bottleneck effect and genetic variability in populations. *Evolution* 29: 1–10.

Nielsen R., Korneliussen T., Albrechtsen A., Li Y. and Wang J. 2012. SNP Calling, Genotype Calling, and Sample Allele Frequency Estimation from New-Generation Sequencing Data. PLoS ONE 7: e37558

Nosil P. 2008. Speciation with gene flow could be common. Molecular Ecology. 17: 2103–2106

Parra G., Bradnam K. and Korf I. 2007. CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. Bioinformatics 23: 1061-1067.

Price J.P. and Clague D.A. 2002. How old is the Hawaiian biota? Geology and phylogeny suggest recent divergence. Proceedings of the Royal Society of London B 269: 2429–2435

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945–959.

R Development Core Team. 2008. R: a language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. Available at: http://www.R-project.org.

Rius M. and Darling J.A. 2014. How important is intraspecific genetic admixture to the success of colonising populations? TREE 29: 233-242

Roderick G.K., Croucher P.J.P., Vandergast A.G. and Gillespie R.G. 2012. Species differentiation on a dynamic Landscape: shifts in metapopulation genetic structure using the chronology of the Hawaiian Archipelago. Evolutionary Biology 39: 192–206

Sanggaard K.W., Bechsgaard J.S., Fang X., Duan J., Dyrlund T.F., Gupta V., Jiang X., Cheng L., Fan D., Feng Y., Han L., Huang Z., Wu Z., Liao L., Settepani V., Thøgersen I.B., Vanthournout B., Wang T., Zhu Y., Funch P., Enghild J.J., Schauser L., Andersen S.U., Villesen P., Schierup M.H., Bilde T. and Wang J. 2014. Spider genomes provide insight into composition and evolution of venom and silk. Nature Comunications. 5: 3765

Schluter D. 2000. The Ecology of Adaptive Radiation. Oxford University Press, Oxford.

Seehausen O. 2004. Hybridization and adaptive radiation. Trends Ecol. Evol. 19: 198–207

Seehausen O. 2006. African cichlid fish: a model system in adaptive radiation research. Proc. R. Soc. B 273: 1987–1998

Simpson J.T., Wong K., Jackman S.D., Schein J.E., Jones S.J. and Birol I. 2009. ABySS: a parallel assembler for short read sequence data. Genome Research. 19: 1117-1123

Skotte L., Korneliussen T.S. and Albrechtsen A. 2013. Estimating individual admixture proportions from next generation sequencing data. Genetics. 195: 693-702

Stacy E.A., Johansen J.B., Sakishima T., Price D.K. and Pillon Y. 2014. Incipient

radiation within the dominant Hawaiian tree *Metrosideros polymorpha*. Heredity. 113: 334–342

Sulonen A.M., Ellonen P., Almusa H., Lepistö M., Eldfors S., Hannula S., Miettinen T., Tyynismaa H., Salo P., Heckman C., Joensuu H., Raivio T., Suomalainen A. and Saarela J. 2011. Comparison of solution-based exome capture methods for next generation sequencing. Genome Biology. 12: R94

Tapia D.H., Troncoso A.K., Vargas R.R., Olivares-Donoso R. and Niemeyer H.M. 2008. Experimental evidence for competitive exclusion of *Myzus persicae nicotianae* by *Myzus persicae s.s.* (Hemiptera: Aphididae) on sweet pepper, *Capsicum annuum* (Solanaceae). Eur. J. Entomol. 105: 643-648

Tatusov R.L., Fedorova N.D., Jackson J.D., Jacobs A.R., Kiryutin B., Koonin E.V., Krylov D.M., Mazumder R., Mekhedov S.L., Nikolskaya A.N., Rao B.S., Smirnov S., Sverdlov A.V., Vasudevan S., Wolf Y.I., Yin J.J. and Natale D.A. 2003. The COG database: an updated version includes eukaryotes. BMC Bioinformatics 4: 41

Trifonov V.A., Vorobieva N.N. and Rens W. 2009. Chapter 9 FISH With and Without COT1 DNA. In: T. Liehr (ed.) Fluorescence In Situ Hybridization (FISH) – Application Guide. Springer-Verlag Berlin Heidelberg.

Untergrasser A., Cutcutache I., Koressaar T., Ye J., Faircloth B.C., Remm M. and Rozen S.G. 2012. Primer3 - new capabilities and interfaces. Nucleic Acids Research 40: e115

Vandergast A. G., Gillespie R. G. and G. K. Roderick. 2004. Influence of volcanic activity on the population genetic structure of Hawaiian Tetragnatha spiders: fragmentation, rapid population growth and the potential for accelerated evolution. Molecular Ecology. 13: 1729-1743

Wang Y. and Nielsen R. 2012. Estimating population divergence time and phylogeny from single-nucleotide polymorphisms data with outgroup ascertainment bias. Molecular Ecology 21: 974–986

Waters J.M. 2011. Competitive exclusion: phylogeography's 'elephant in the room'? Mol Ecol. 20: 4388-4394

Yim K.M., Brewer M., Miller C.T. and Gillespie R.G. 2014. Comparative transcriptomics of maturity-associated color change in Hawaiian spiders. Journal of Heredity 105: 771–781

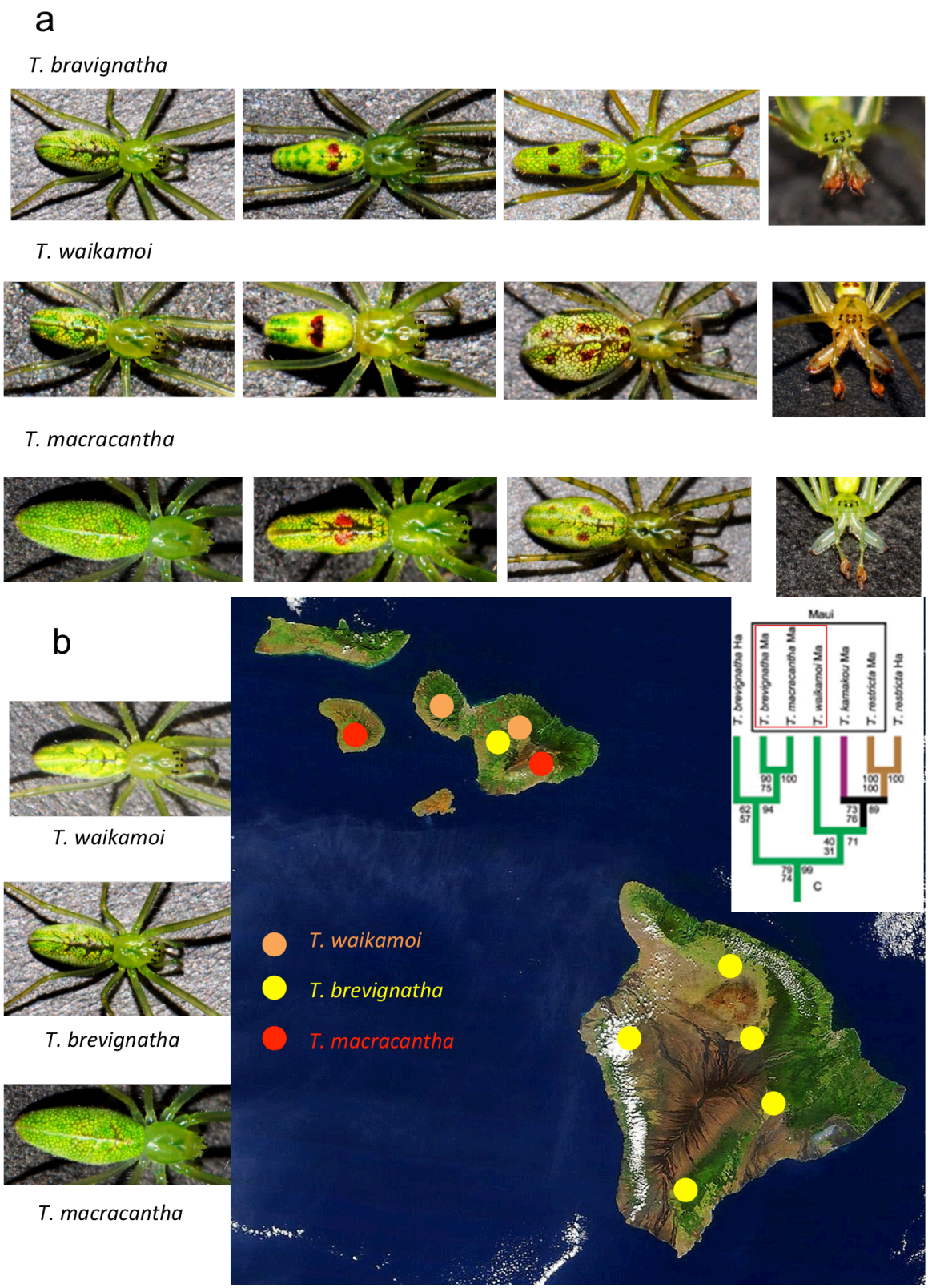



**Figure 1: Study system.** (a) The three green ecomorph *Tetragnatha* spiders present in Maui Nui have a shared color polymorphism. The most frequent morph is the plain green

followed by the “red heart”. The third variant is different on each species. Note that the chelicera on *T. brevignatha* are significantly shorter than on the other species. (b) The three species have populations on different areas of Haleakalā and a second population on a different volcano. In the case of *T. brevignatha* there are several populations within the Big Island. The phylogeny was modified from Gillespie, 2004.

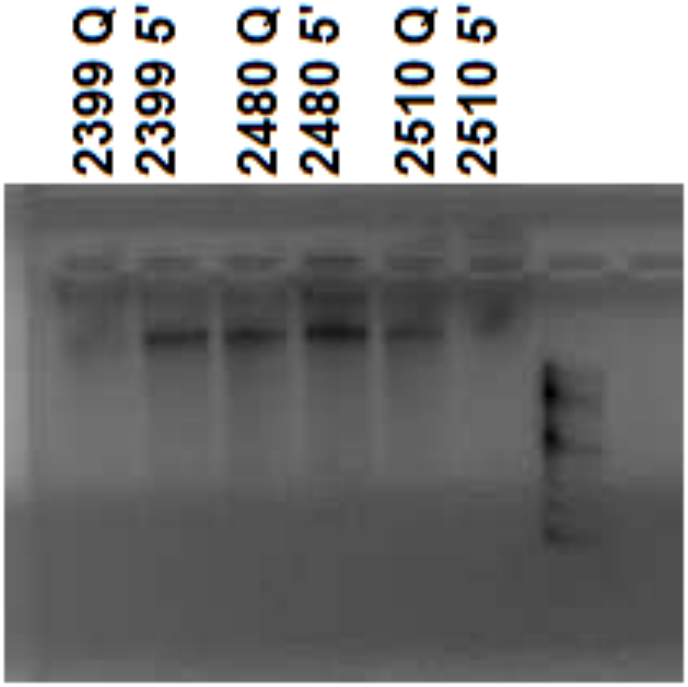



**Figure 2: DNA quality on different DNA extraction methods.** The numbers correspond to different specimens of *T. hawaiiensis*. "Q" refers to the Qiagen kit and "5'" to the 5PRIME kit. The ladder corresponds to a 100bp.

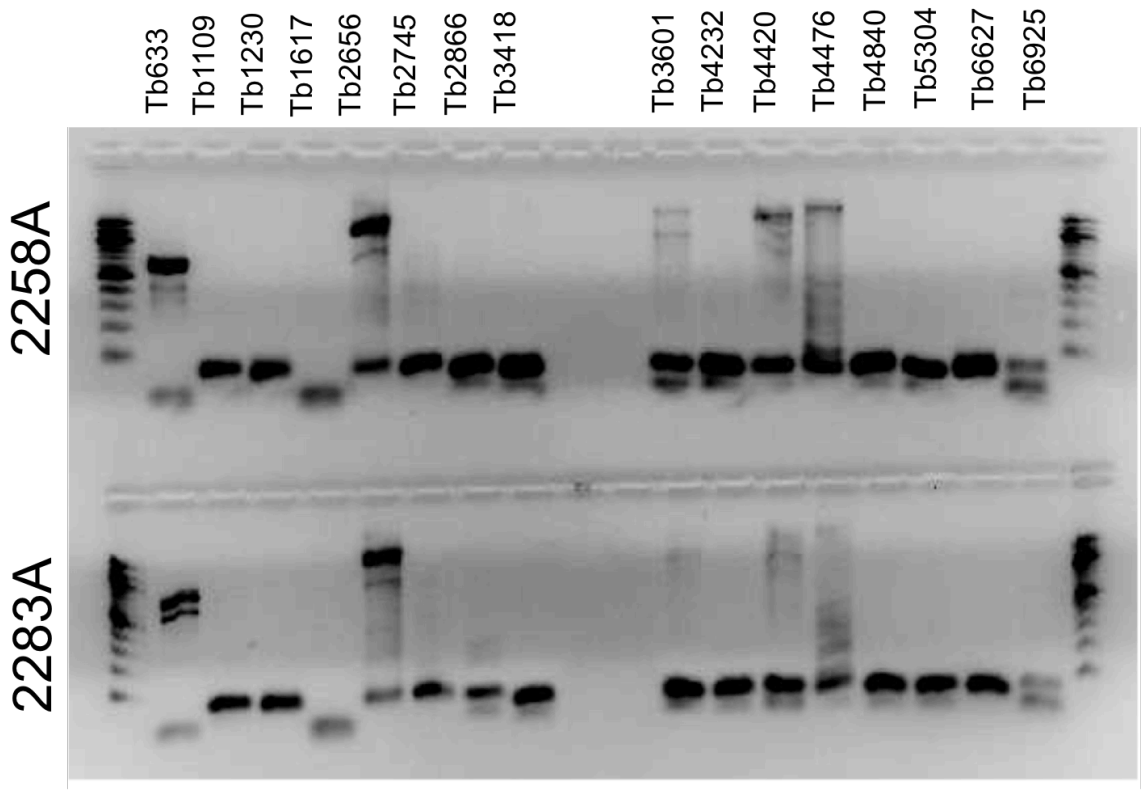



**Figure 3: Test primers for Positive and Negative controls (fast evolving contigs):** The two tested specimens are from Maui. The ladder corresponds to a 100bp.

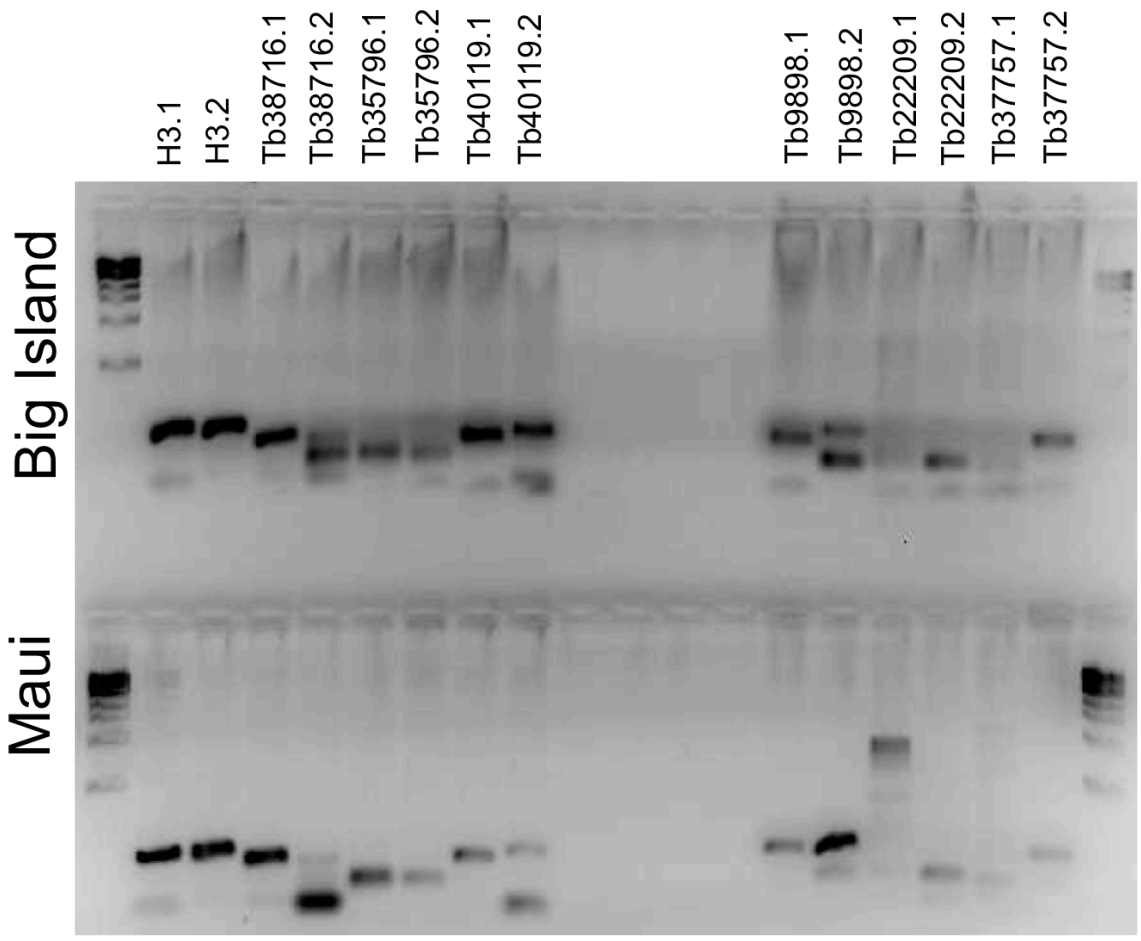



**Figure 4: Test primers for Positive and Negative controls (slow evolving contigs):** The ladder corresponds to a 100bp.

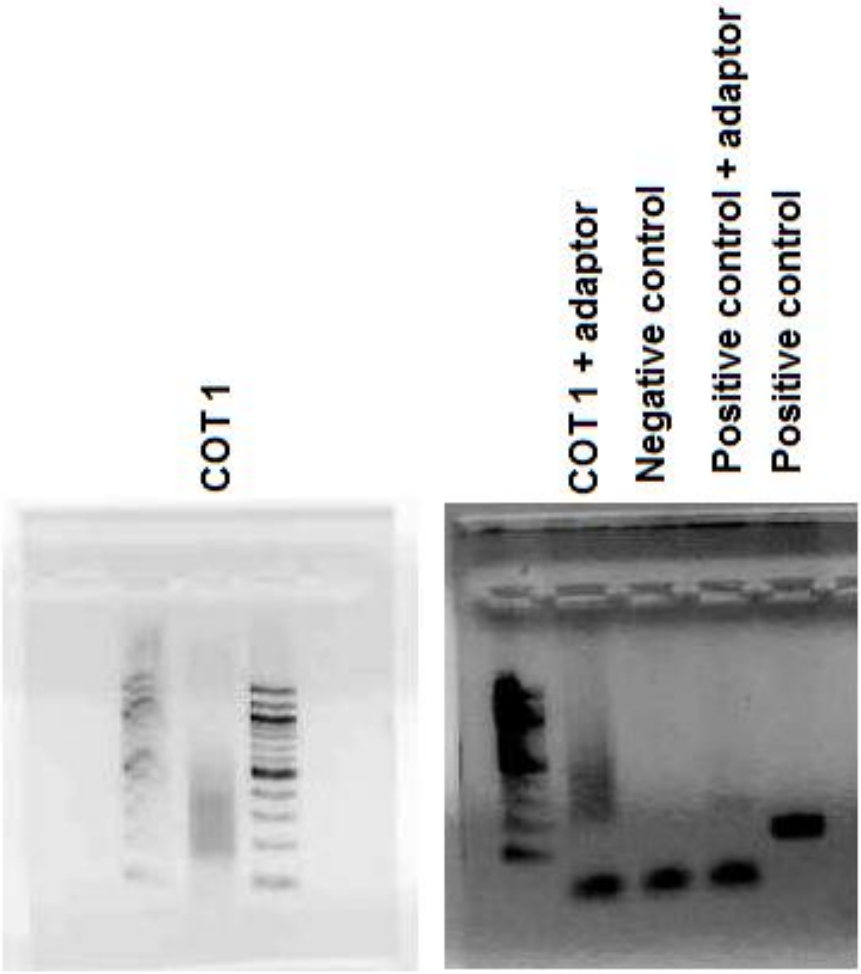



**Figure 5: Isolation of COT1 DNA.** *Right gel.* Fragmented COT1 DNA. *Left gel.* COT1 DNA library after adaptor ligation. Note the shift towards a heavier molecular weight. The negative control does not show anything, indicating that there is no contamination. The positive control after ligation is heavier than before ligation, indicating a success in the adaptor ligation. The band in the positive control after ligation is more faint than the one before ligation.

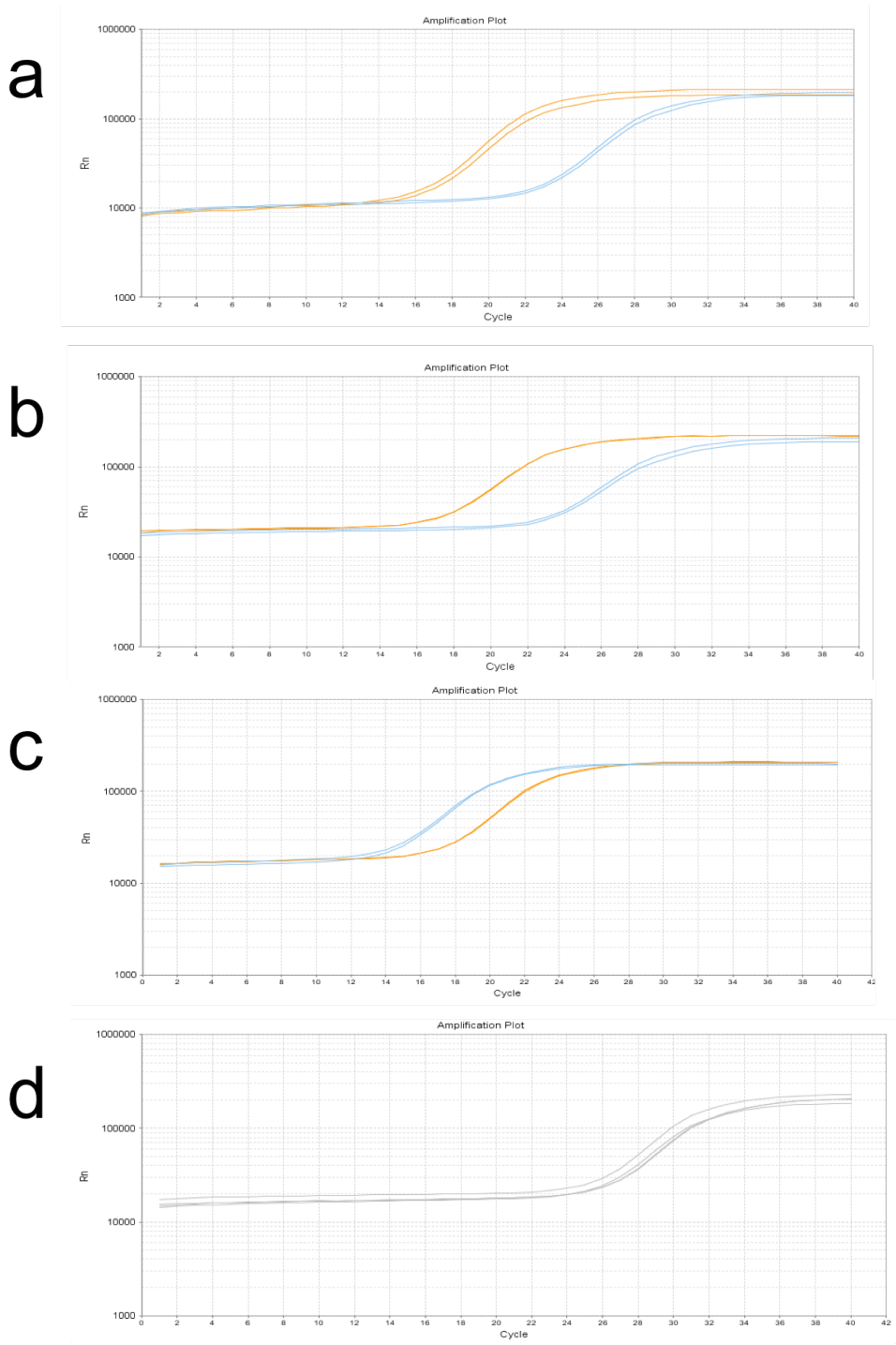



**Figure 6: qPCR controls Experiment 1.** (a) Positive control (Tb9898), (b) Positive control (Tb40119), (c) Negative control (H3), (d) qPCR negative control. Orange: Post-capture, Blue: Pre-capture.

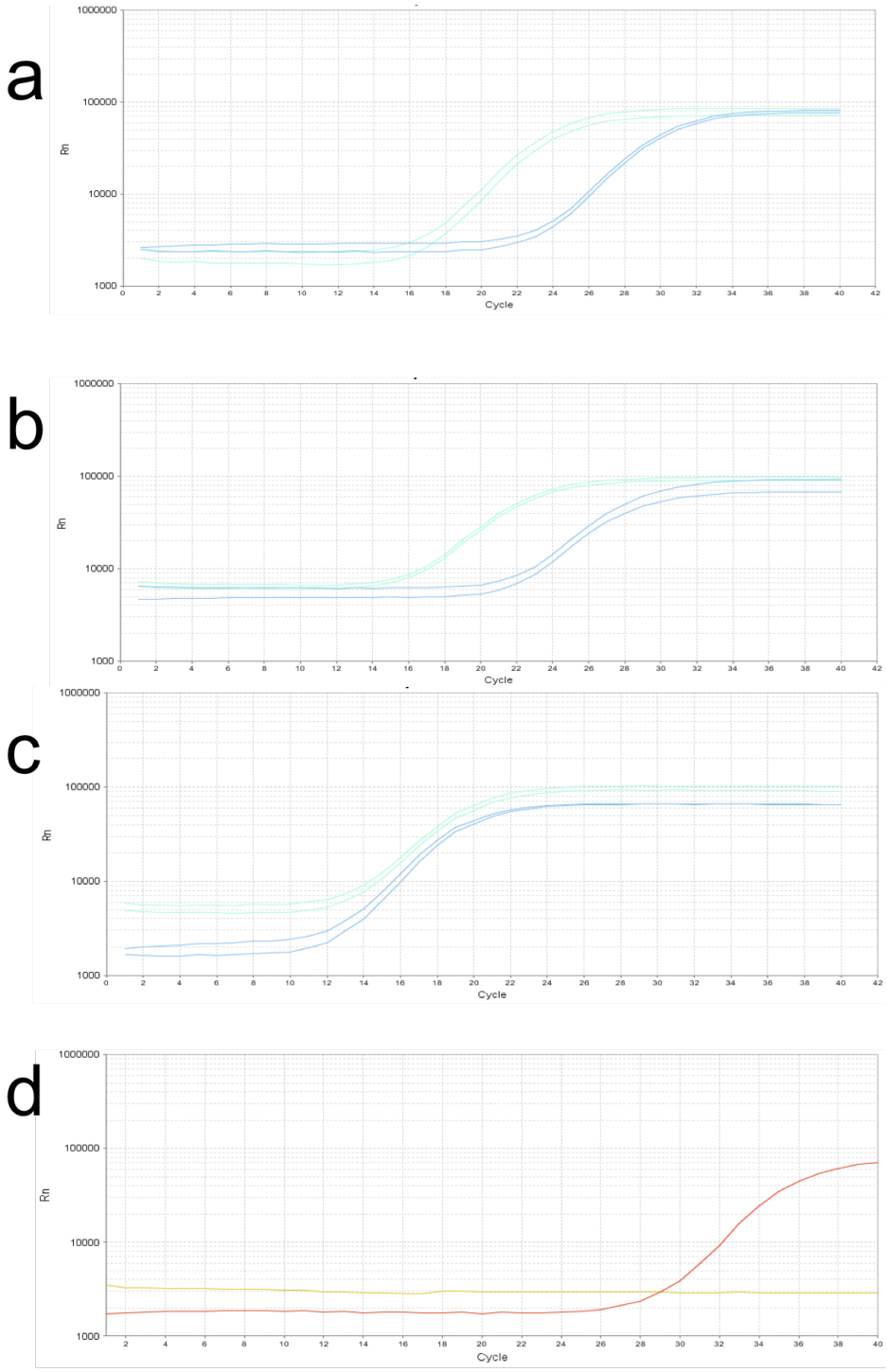



**Figure 7: qPCR controls Experiment 2.** (a) Positive control (Tb9898), (b) Positive control (Tb40119), (c) Negative control (H3), (d) qPCR negative control Experiment 2 and 3. Light blue: Post-capture, Blue: Pre-capture, Red: qPCR negative control Experiment 2, Orange: qPCR negative control Experiment 3.

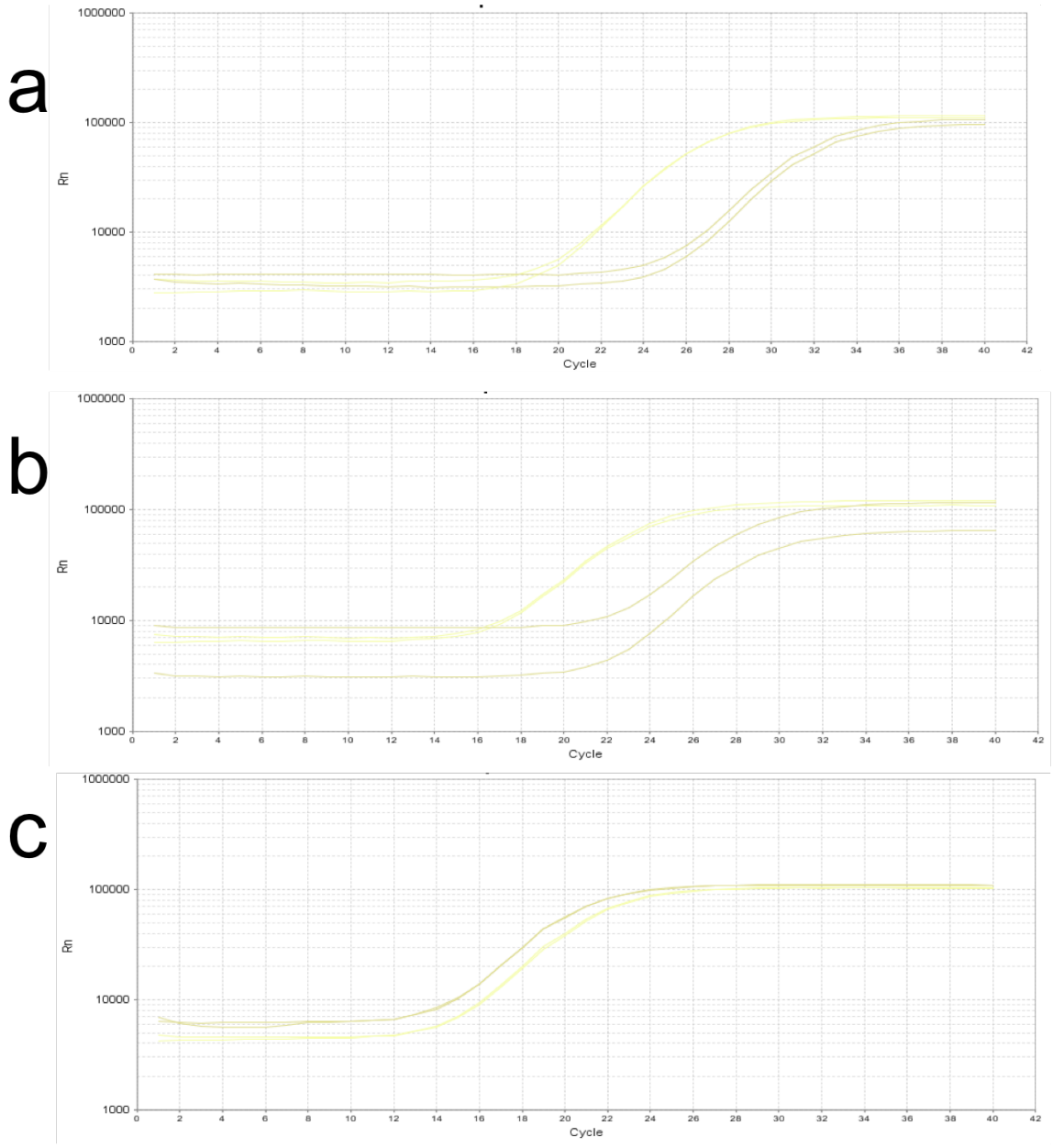



**Figure 8: qPCR controls Experiment 3.** (a) Positive control (Tb9898), (b) Positive control (Tb40119), (c) Negative control (H3). Yellow: Post-capture, Orange: Pre-capture.

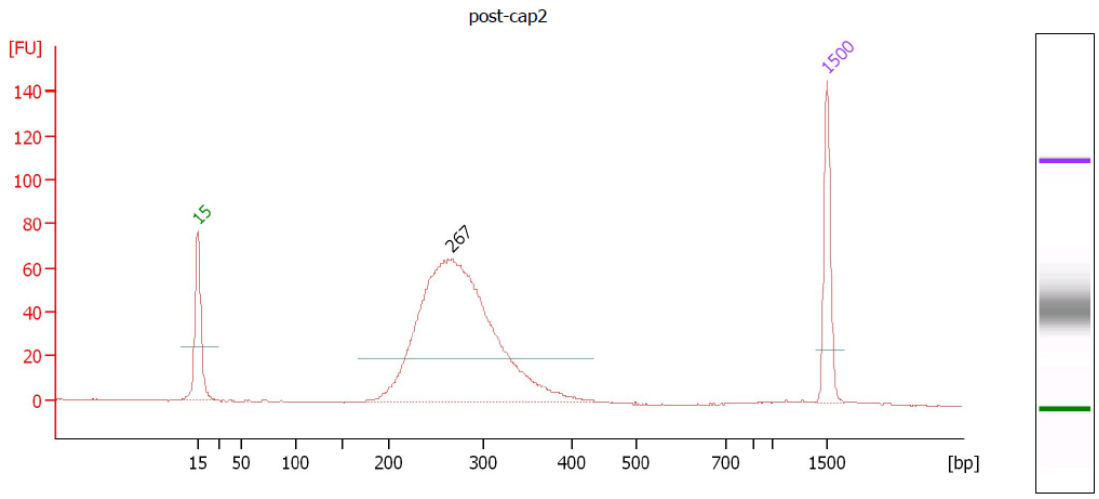



**Figure 9: Bioanalyzer read of the fragment distribution of the whole library after amplification (Experiment 1).**

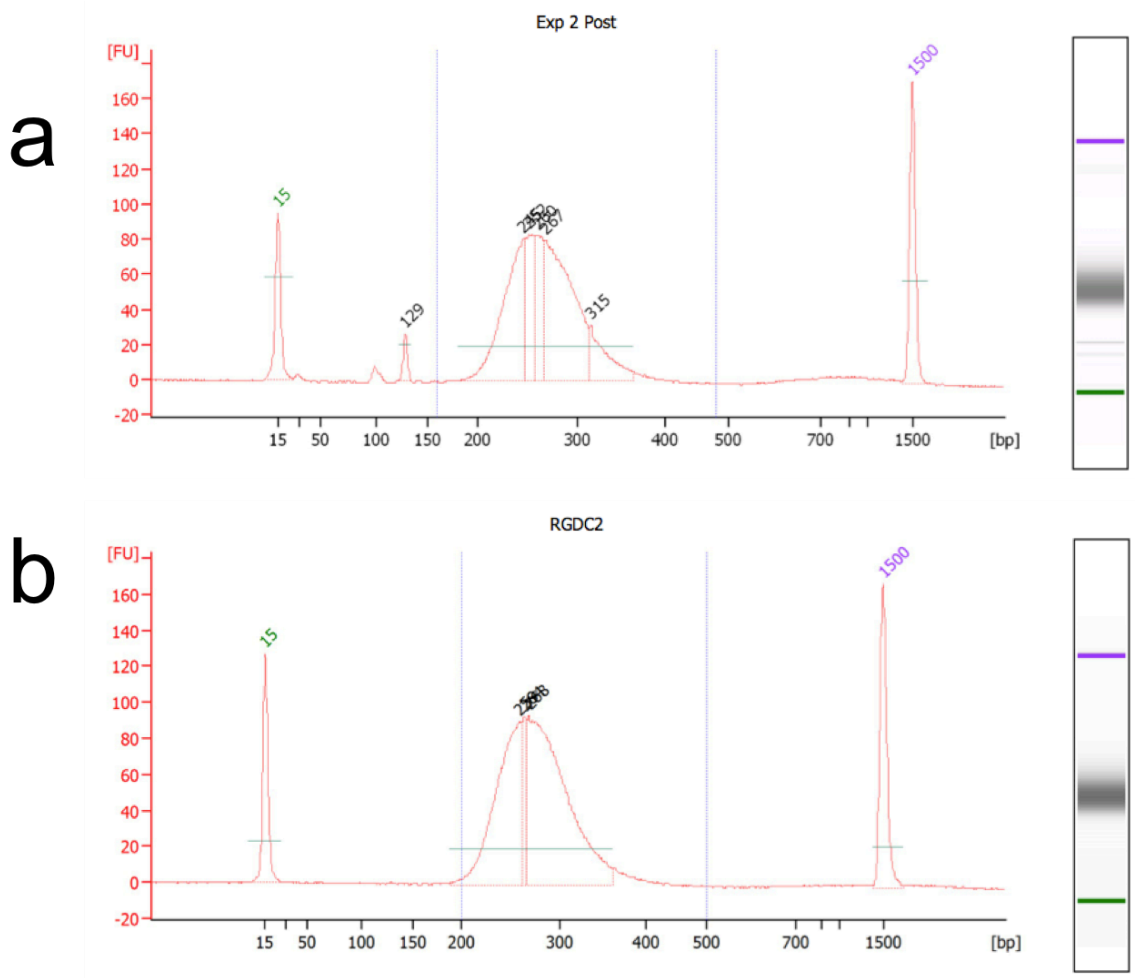



**Figure 10: Bioanalyzer read of the fragment distribution of the whole library after amplification (Experiment 2).** (a) Before beads clean up, note the extra spike at 129 bp. (b) After beads clean up, the spike at 129 bp was removed.

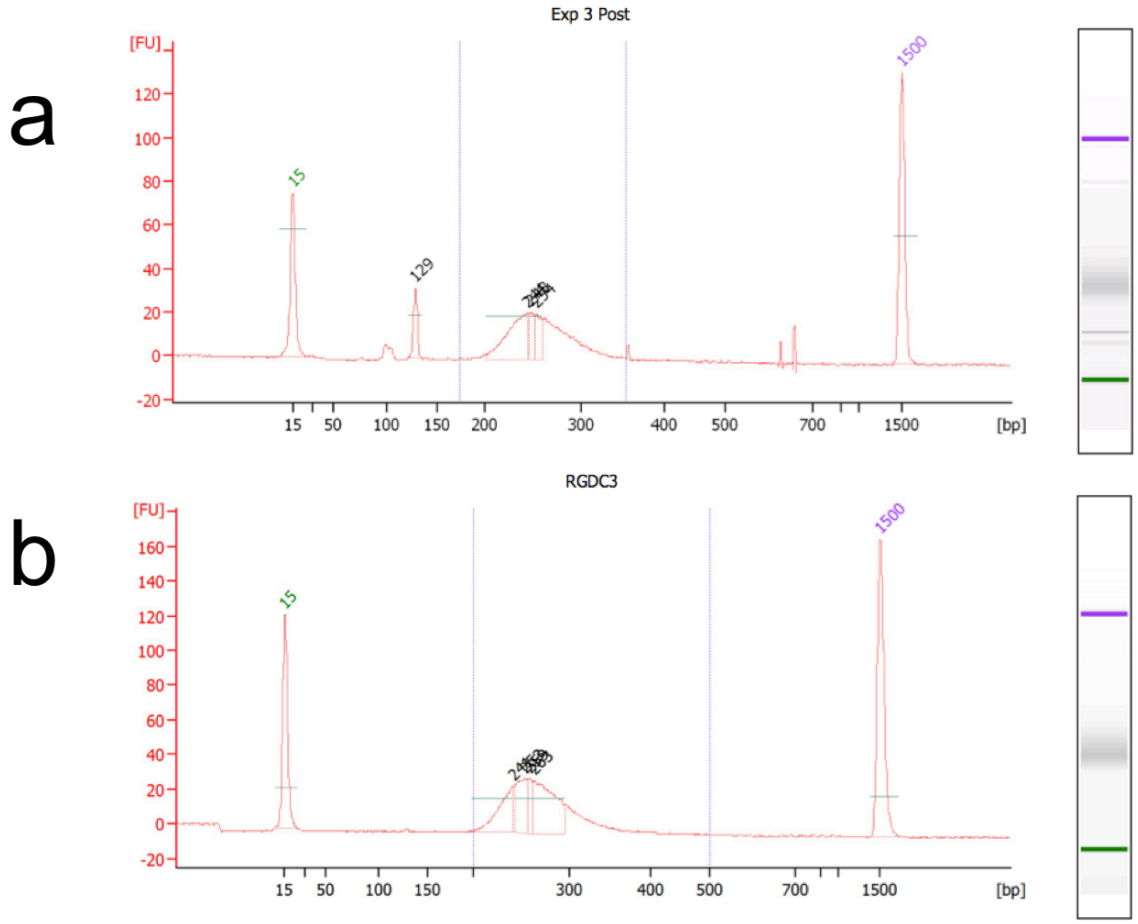



**Figure 11: Bioanalyzer read of the fragment of the whole library after amplification (Experiment 3).** (a) Before beads clean up, note the extra spike at 129 bp. (b) After beads clean up, the spike at 129 bp was removed.

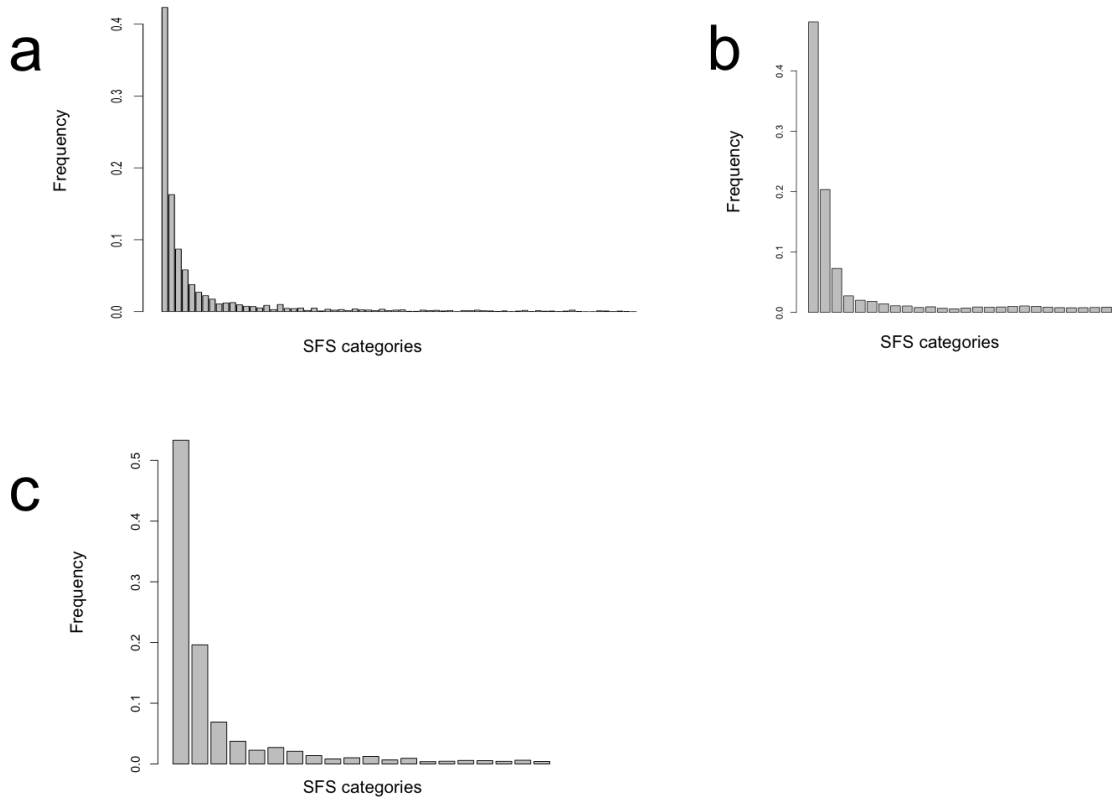



**Figure 12: Site Frequency Spectrum.** (a) *T. brevignatha*, (b) *T. macracantha* and (c) *T. waikamoi.*

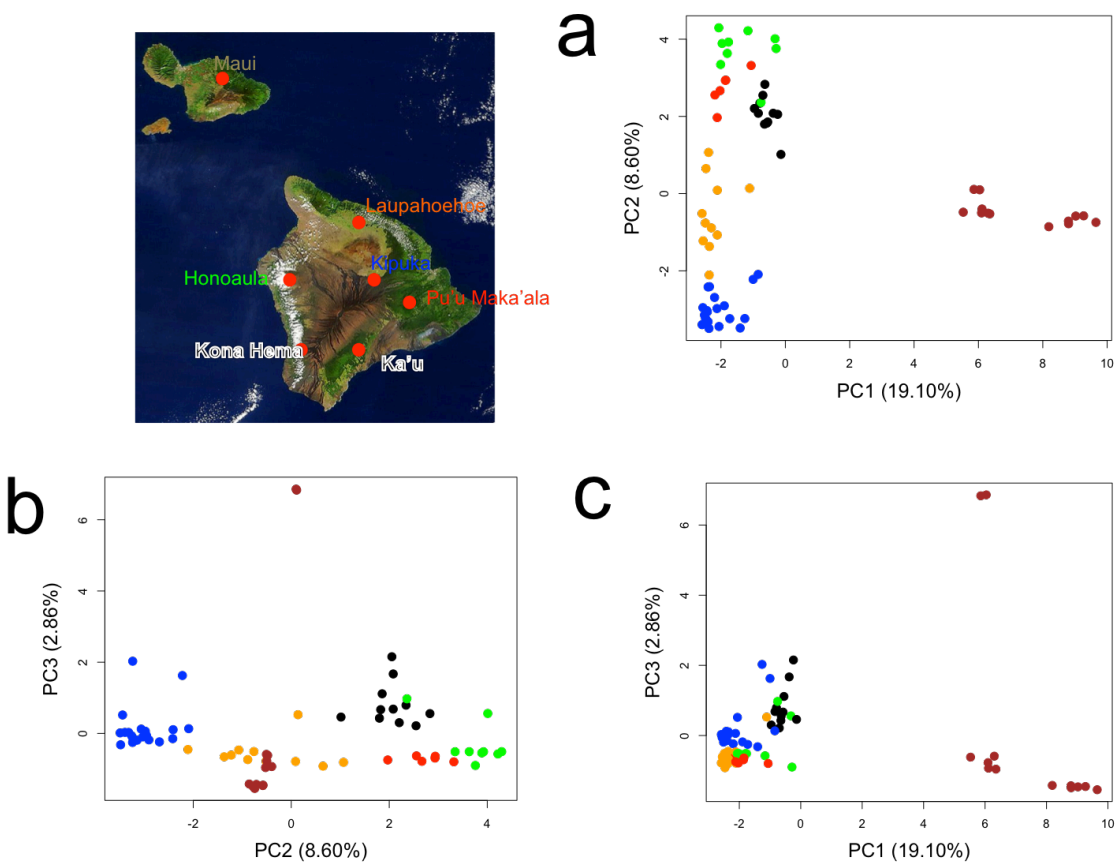



**Figure 13: Principal Component Analysis *T. brevignatha*.** (a) PC1 vs PC2, (b) PC1 vs PC3 and (c) PC2 vs PC3. The insert corresponds to the sampling localities. There is correlation between the colors in the map and in the graph.

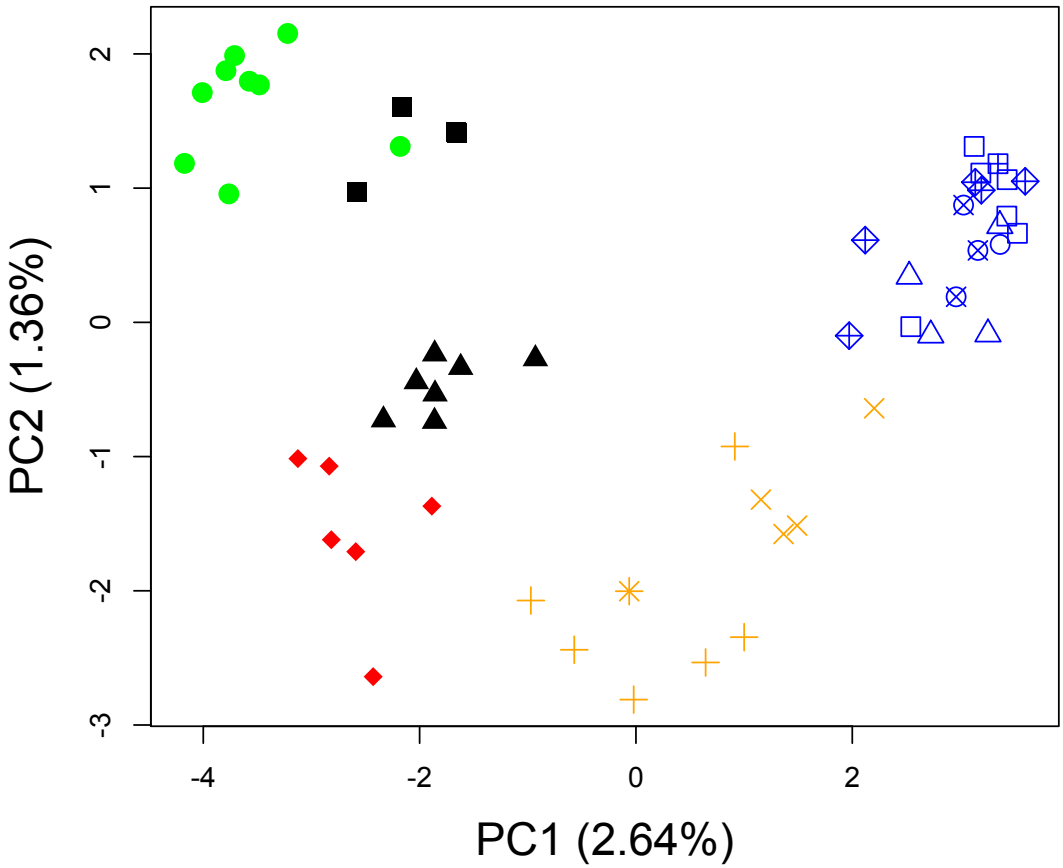



**Figure 14: Principal Component Analysis of the Big Island populations of *T. brevignatha*.** The colors represent the general sampling localities. They follow the same colors as in Fig. 13. The shape of the point correspond to the specific locality: Laupāhoehoe Hight: cross; Laupāhoehoe Maula: X; Laupāhoehoe HETF: asterisk; Kīpuka 23: hollow circle; Kīpuka 18: hollow triangle; Kīpuka 5: hollow square; Kīpuka 15: circle with X; Kīpuka 37: sequeary with cross; Forest: rhombus wih cross; Pu'u Maka'ala, Army Road: rhombus; Kona Hema: square; Ka'u: triangle; and Honoaula: circle.

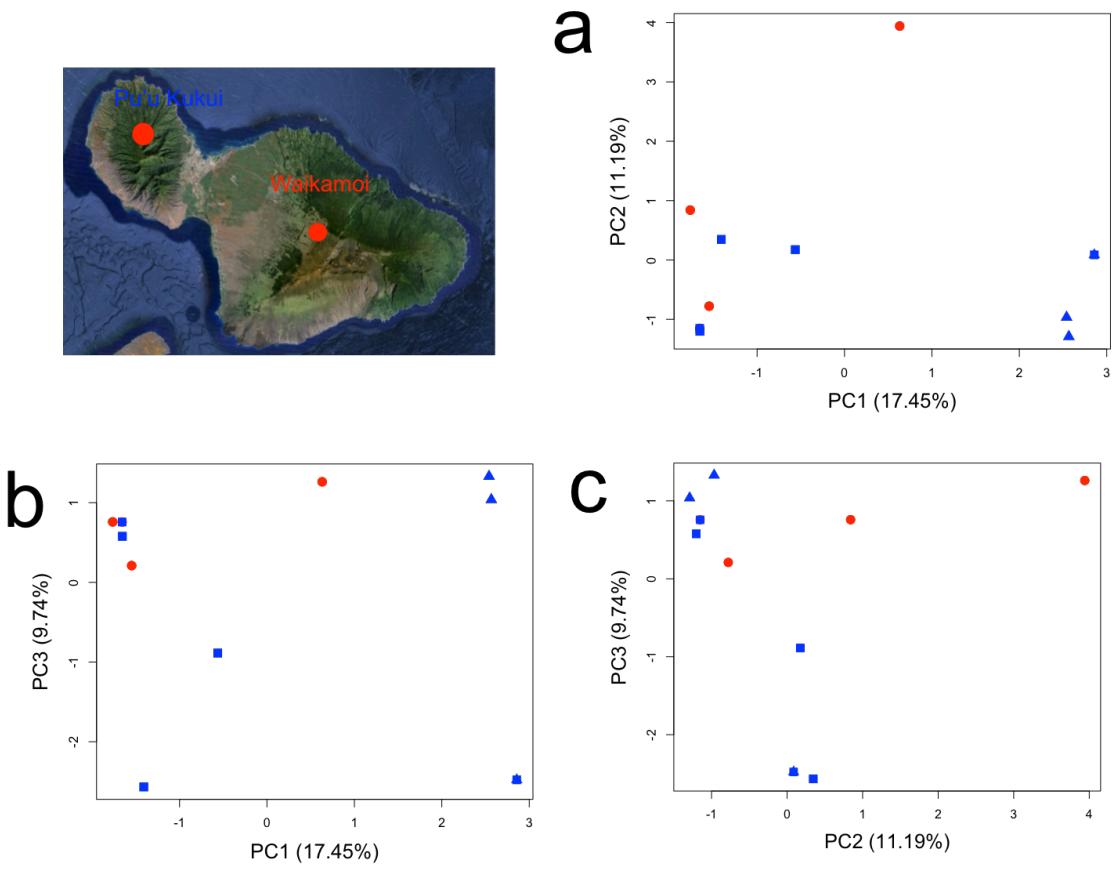



**Figure 15: Principal Component Analysis *T. waikamoi*.** (a) PC1 vs PC2, (b) PC1 vs PC3 and (c) PC2 vs PC3. The insert corresponds to the sampling localities. There is correlation between the colors in the map and in the graph. The shape of the point correspond to the specific locality: Upper Waikamoi: circle; Pu'u Kukui, upper: squeare; and Pu'u Kukui, lower: circle.

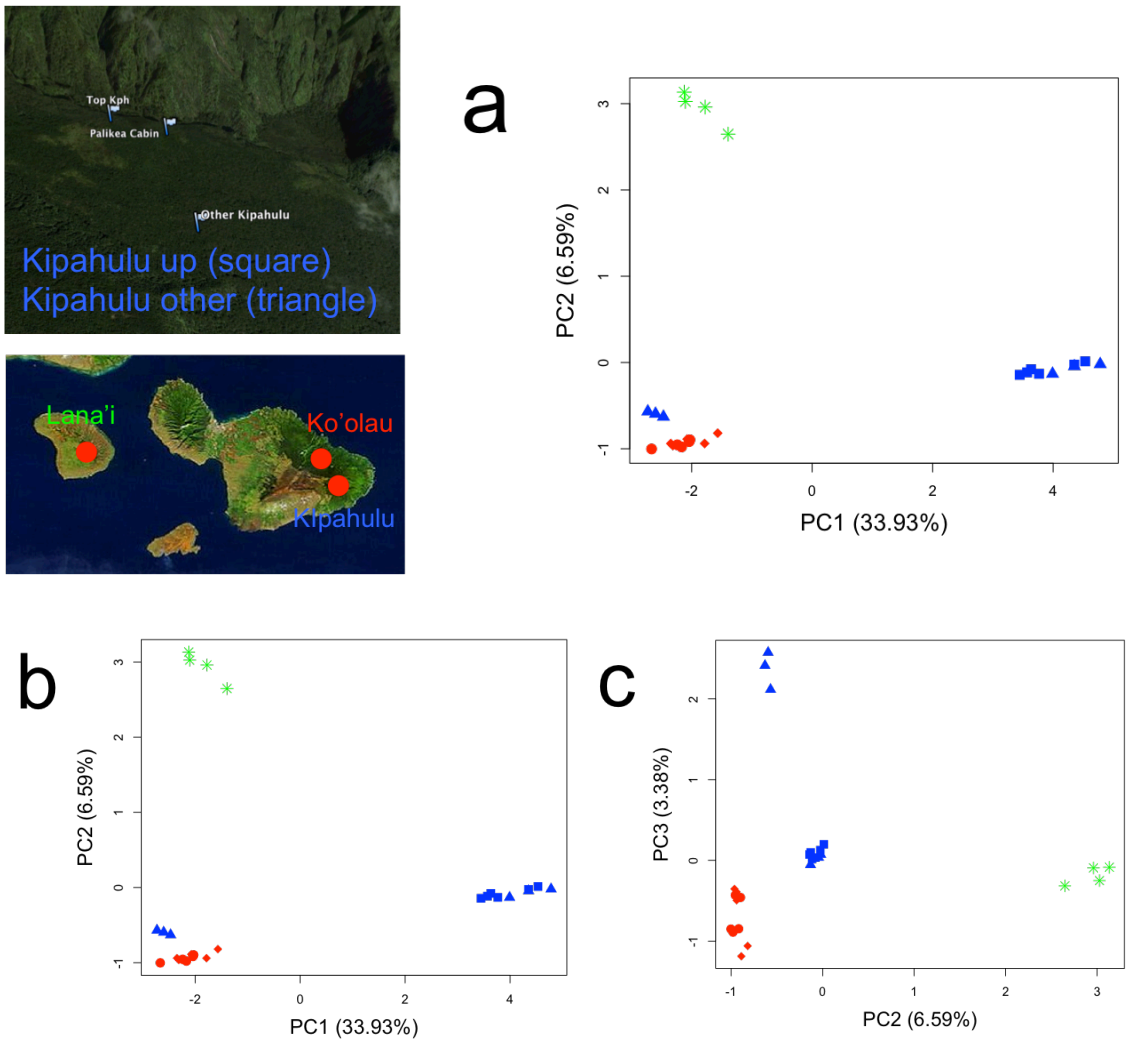
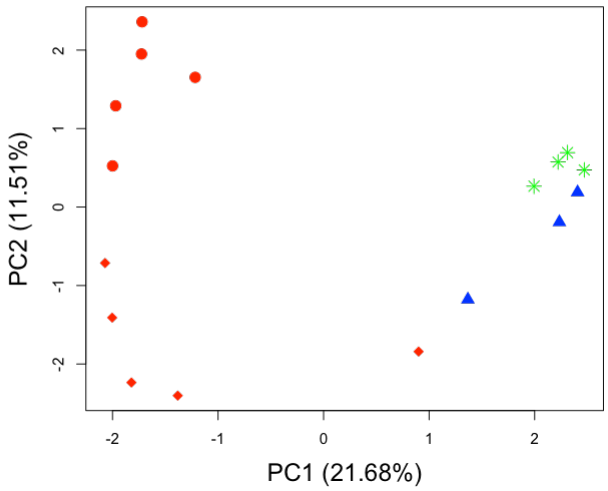
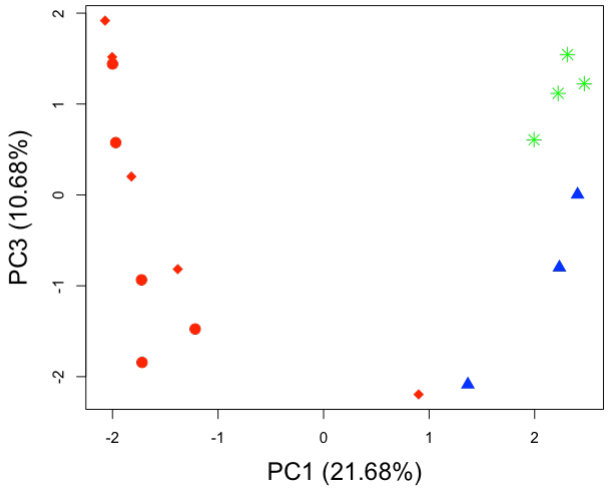



**Figure 16: Principal Component Analysis *T. macracantha*.** (a) PC1 vs PC2, (b) PC1 vs PC3 and (c) PC2 vs PC3. The insert corresponds to the sampling localities. The distance between "Kipa up" and "Kipa other" is 950 meters. There is correlation between the colors in the map and in the graph. The shape of the point correspond to the specific locality: Ko'olau, near Cramp: circle; Ko'olau, Day 2: rhombus; Kīpahulu Valley, up: triangle; Kīpahulu Valley, other: square; and Lana'i: asterisk.

a


PC2 (11.51%)
PC1 (21.68%)
-2
-1
0
1
2

b


PC3 (10.68%)
PC1 (21.68%)
-2
-1
0
1
2

c

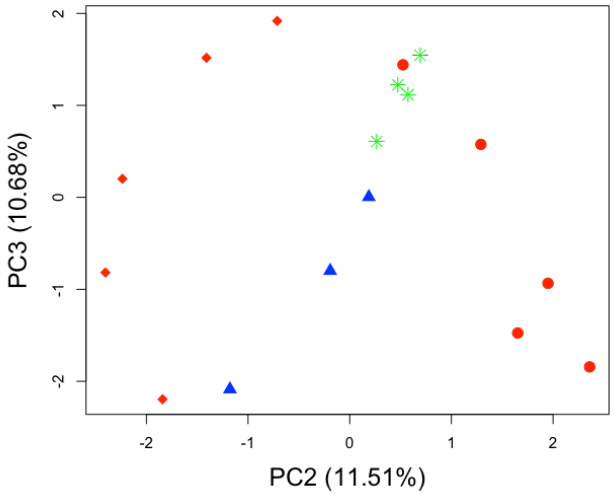
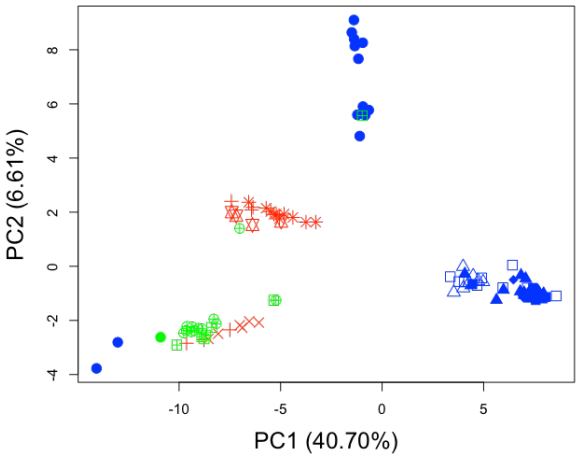
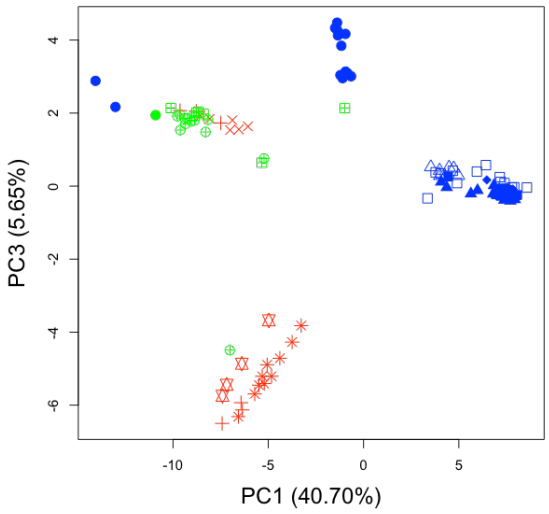

PC3 (10.68%)
PC2 (11.51%)
-2
-1
0
1
2

**Figure 17: Principal Component Analysis *T. macracantha,* all the specimens except the ones from Kīpahulu Valley that differenciate on PC1 in Fig. 16.** (a) PC1 vs PC2, (b) PC1 vs PC3 and (c) PC2 vs PC3. The colors represent the general locality: Red: Ko'olau; Green: Lana'i; Blue: Kīpahulu Valley. The shape of the point correspond to the specific locality: Ko'olau, near Cramp: circle; Ko'olau, Day 2: rhombus; Kīpahulu Valley, up: triangle; and Lana'i: asterisk.

a




b




c

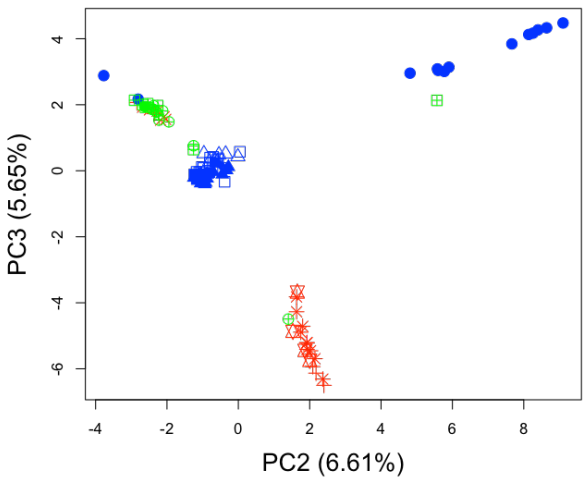
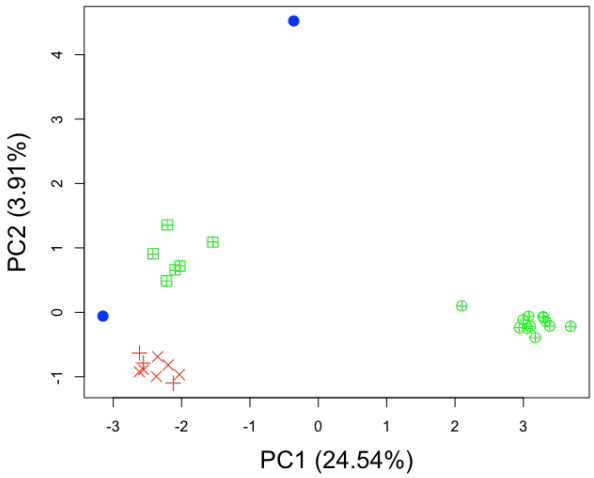
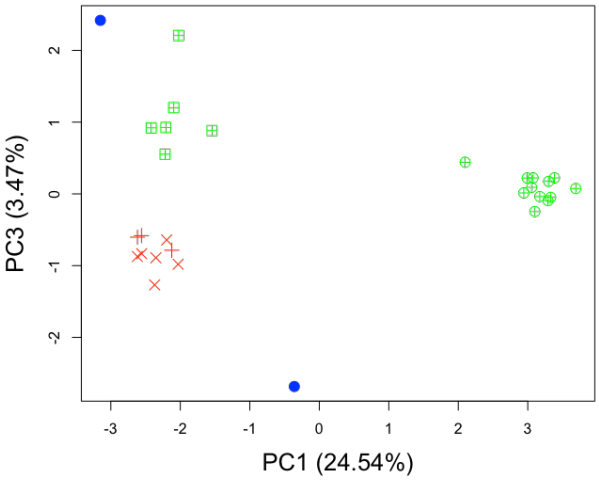


**Figure 18: Principal Component Analysis of all the specimens.** (a) PC1 vs PC2, (b) PC1 vs PC3 and (c) PC2 vs PC3. The colors represent the nominal species: Blue: *T. brevignatha*; Red: *T. macracantha*; and Green: *T. waikamoi*. The shape of the point correspond to the specific locality: Lower Waikamoi: circle; Mauna Kea: squeare; Mauna Loa: triangle; Kīlauea: rhombus; Ka'u: hollow triangle; Kona Hema and Honoaula: hollow squeare; Kīpahulu Valley, up: cross; Kīpahulu Valley, other: X; Ko'olau: asterisk; Lana'i: Triangle star; Upper Waikamoi: squeare with cross; and Pu'u Kukui: circle with cross.

a




b




c

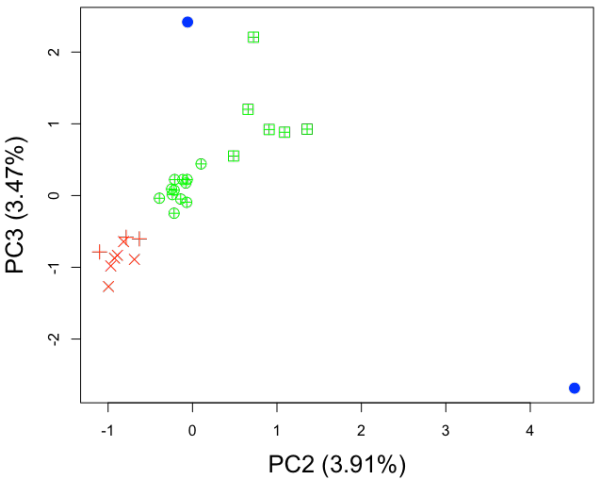


**Figure 19: Principal Component Analysis of WaikMacBrev.** (a) PC1 vs PC2, (b) PC1 vs PC3 and (c) PC2 vs PC3. The colors represent the nominal species: Blue: *T. brevignatha*; Red: *T. macracantha*; and Green: *T. waikamoi*. The shape of the point correspond to the specific locality: Lower Waikamoi: circle; Kīpahulu Valley, up: cross; Kīpahulu Valley, other: X; Upper Waikamoi: squeare with cross; and Pu'u Kukui: circle with cross.

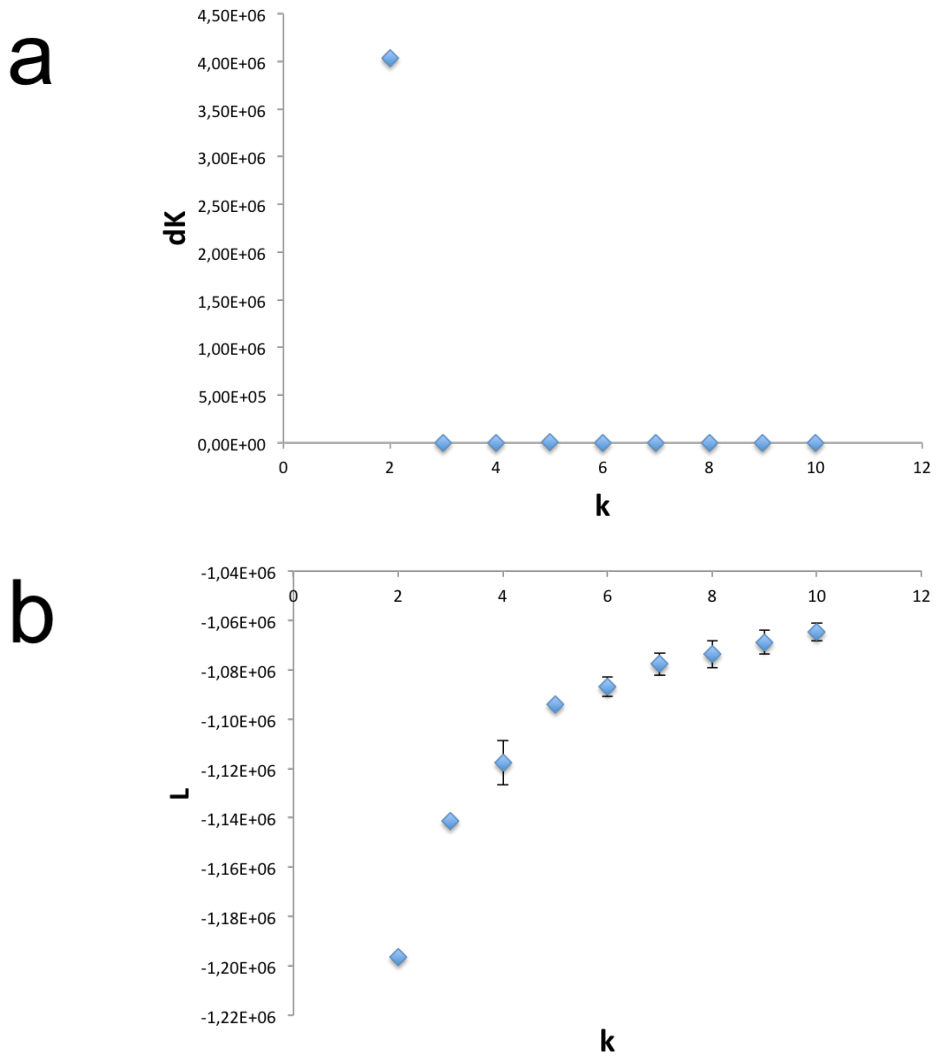
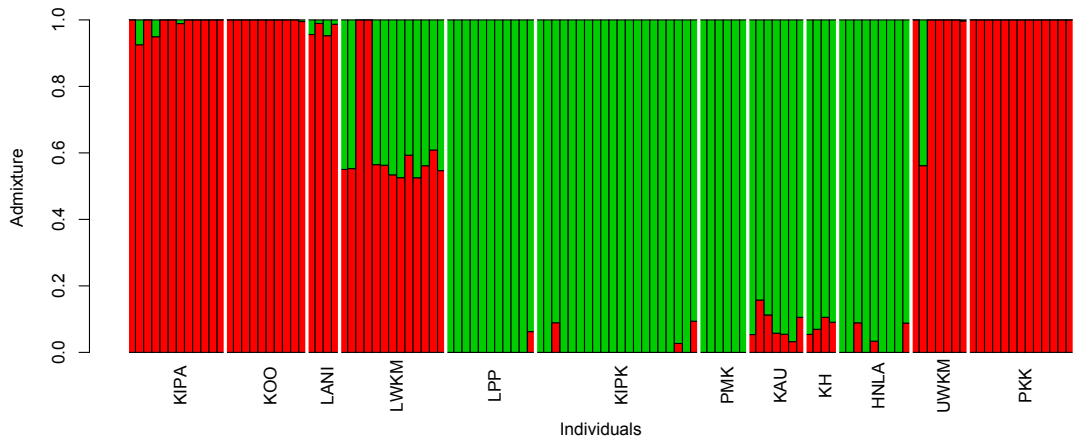
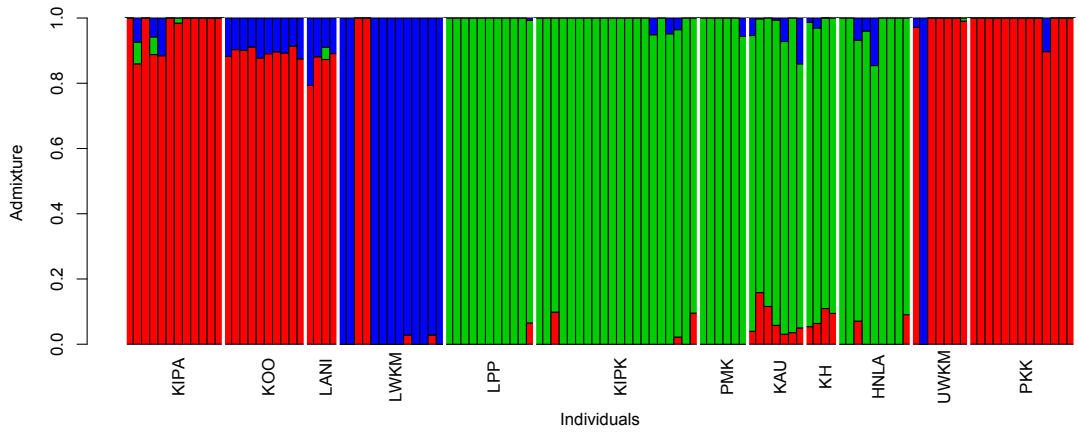



**Figure 20: Implementation of Evanno method to choose optimal k value.** (a) Δk vs k. See Evanno et al., 2005 for definition of Δk. (b) Average Likelihood (10 independent replicates) vs k. The graph does not include k=1.

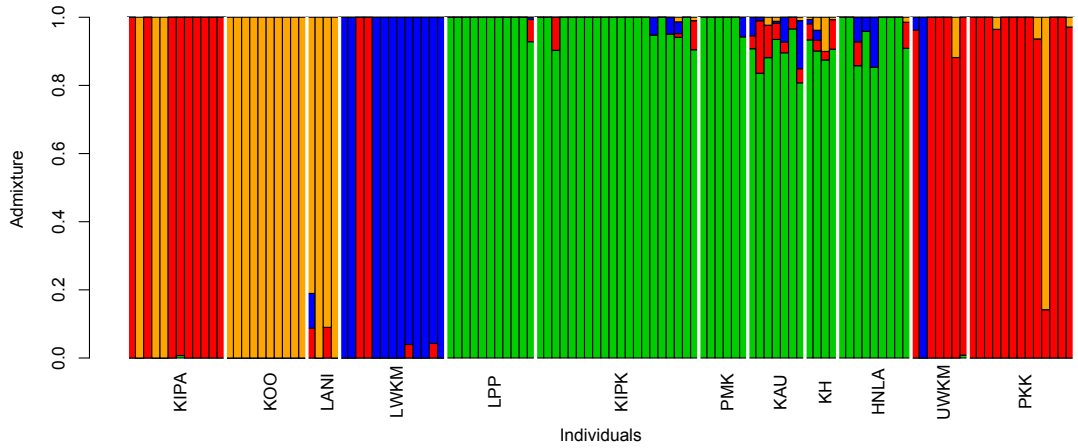
a


Admixture
0.0
0.2
0.4
0.6
0.8
1.0
KIPA
KOO
LANI
LWKM
LPP
KIPK
PMK
KAU
KH
HNLA
UWKM
PKK
Individuals

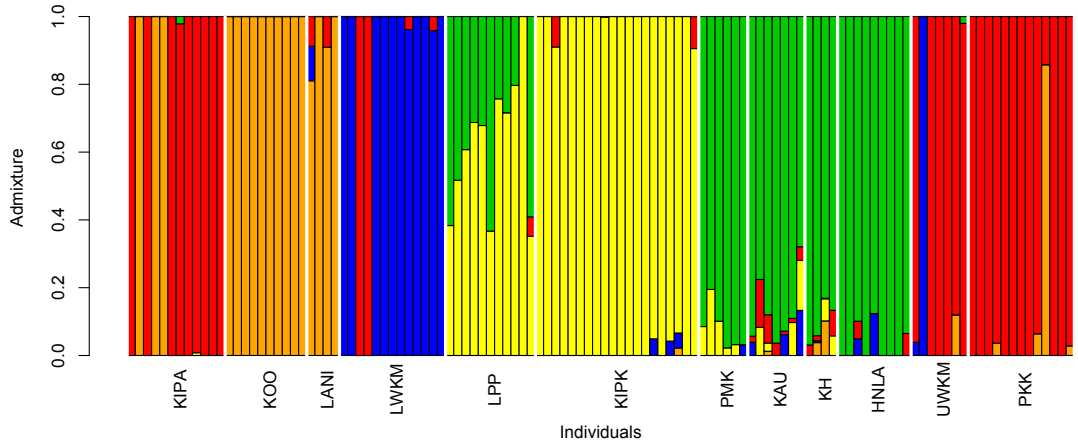
b


Admixture
0.0
0.2
0.4
0.6
0.8
1.0
KIPA
KOO
LANI
LWKM
LPP
KIPK
PMK
KAU
KH
HNLA
UWKM
PKK
Individuals

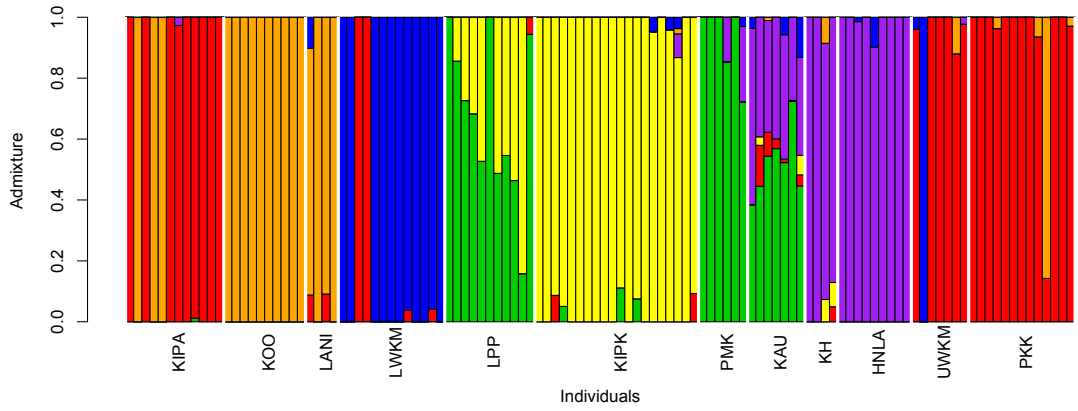
c


Admixture
0.0
0.2
0.4
0.6
0.8
1.0
KIPA
KOO
LANI
LWKM
LPP
KIPK
PMK
KAU
KH
HNLA
UWKM
PKK
Individuals

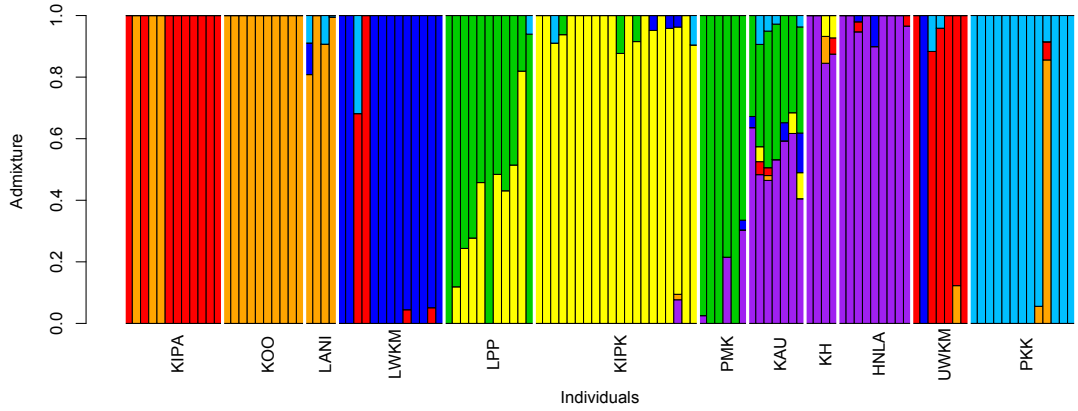
d


Admixture
0.0
0.2
0.4
0.6
0.8
1.0
KIPA
KOO
LANI
LWKM
LPP
KIPK
PMK
KAU
KH
HNLA
UWKM
PKK
Individuals

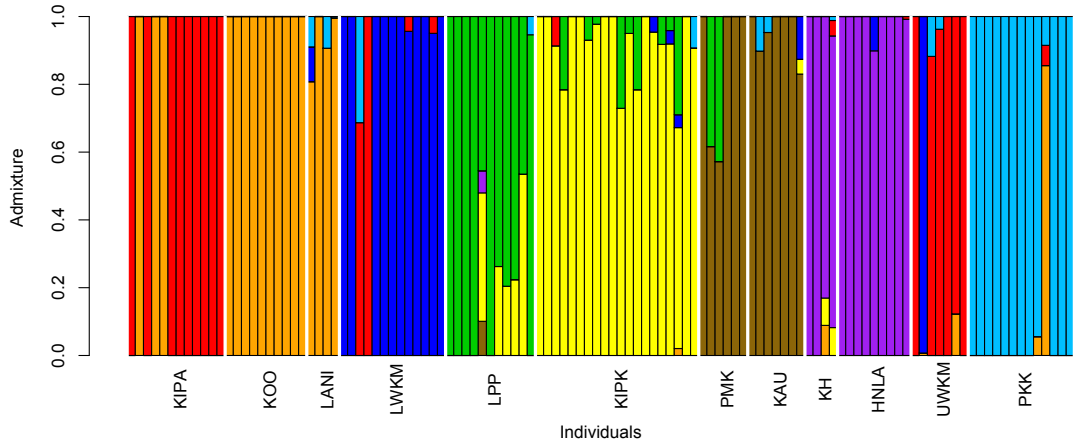
e


Admixture
0.0
0.2
0.4
0.6
0.8
1.0
KIPA
KOO
LANI
LWKM
LPP
KIPK
PMK
KAU
KH
HNLA
UWKM
PKK
Individuals

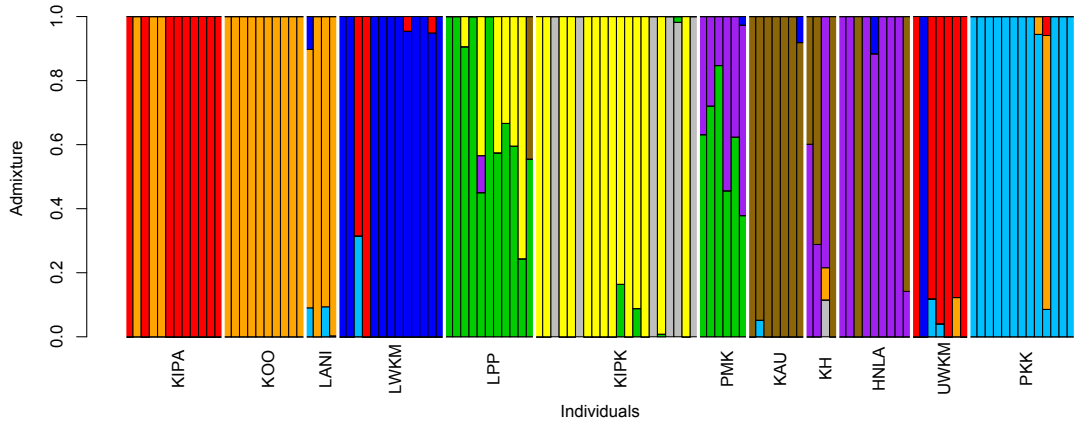
f


Admixture
0.0
0.2
0.4
0.6
0.8
1.0
KIPA
KOO
LANI
LWKM
LPP
KIPK
PMK
KAU
KH
HNLA
UWKM
PKK
Individuals

g


Admixture
0.0
0.2
0.4
0.6
0.8
1.0
KIPA
KOO
LANI
LWKM
LPP
KIPK
PMK
KAU
KH
HNLA
UWKM
PKK
Individuals

h


Admixture
0.0
0.2
0.4
0.6
0.8
1.0
KIPA
KOO
LANI
LWKM
LPP
KIPK
PMK
KAU
KH
HNLA
UWKM
PKK
Individuals

i

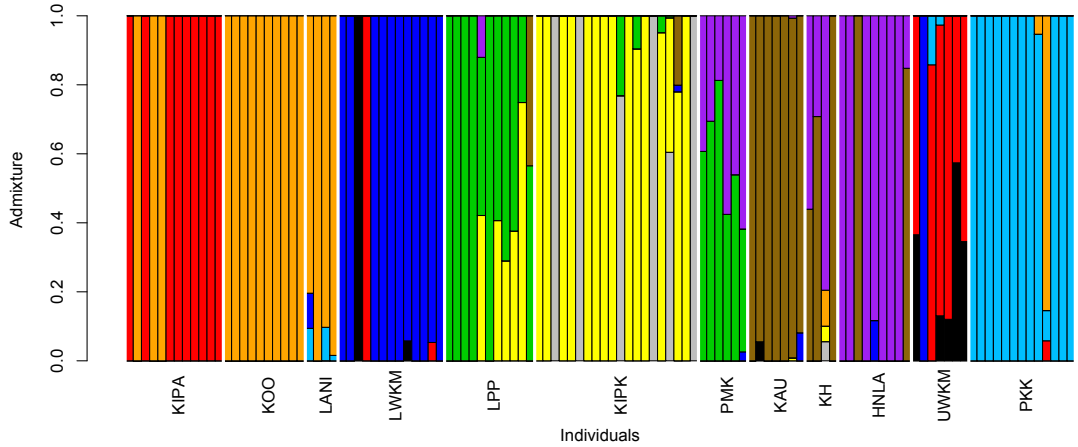

Admixture
0.0
0.2
0.4
0.6
0.8
1.0
KIPA
KOO
LANI
LWKM
LPP
KIPK
PMK
KAU
KH
HNLA
UWKM
PKK
Individuals

**Figure 21: ngsADMIX runs.** (a) k=2, (b) k=3, (c) k=4, (d) k=5, (e) k=6, (f) k=7, (g) k=8, (h) k=9 and (i) k=10. The specimens are separated by localities. The localities correspond to the following nominal species: *T. macracantha*: KIPA (Kīpahulu Valley), KOO (Ko'olau) and LANI (Lana'i); *T. brevignatha*: LWKM (Lower Waikamoi), LPP (Laupāhoehoe), KIPK (Kīpuka), PMK (Pu'u Maka'ala), KAU (Ka'u), KH (Kona Hema) and HNLA (Honoaula); *T. waikamoi*: UWKM (Upper Waikamoi) and PKK (Pu'u Kukui).

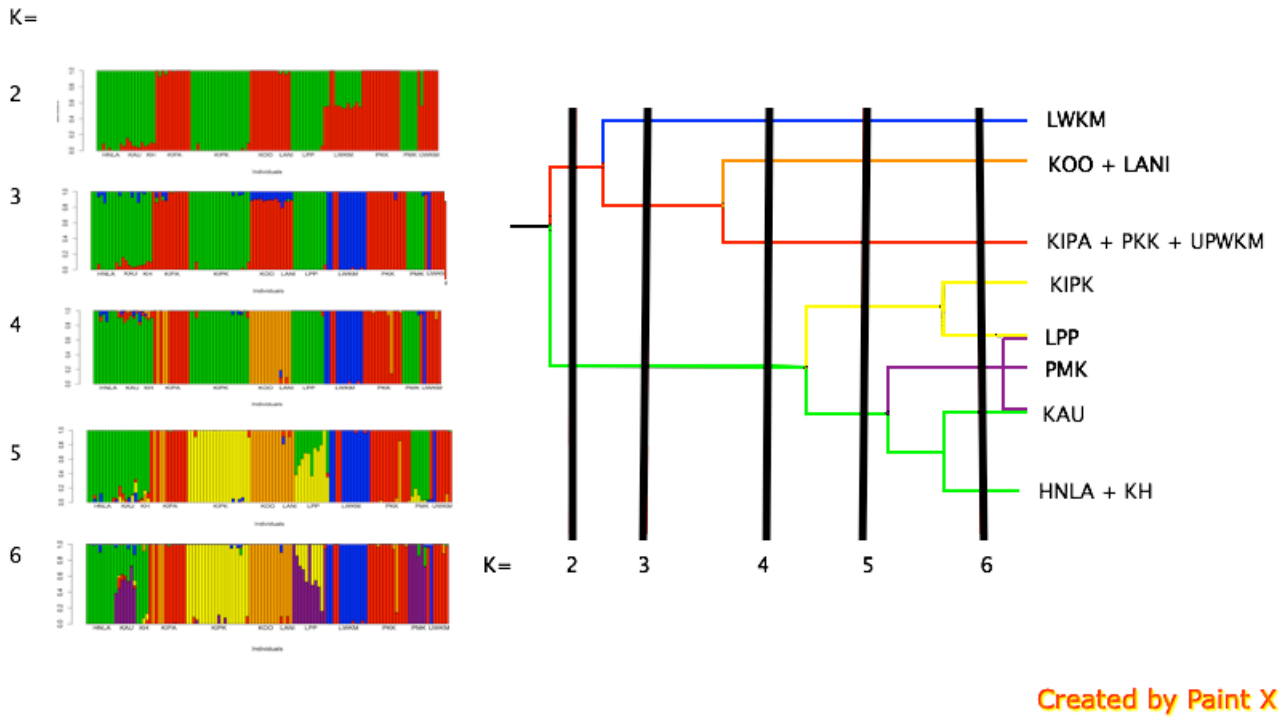



**Figure 22: Dengrogram representation of k=2 to k=6 of ngsADMIX.** The present diagram corresponds to a cartoon to understand the succession of different values of k as sectional cuts on a dendrogram. It is not a phylogenetic reconstruction. Note in k=6 the presence of admixture populations is represented with different branches merging in the terminal.

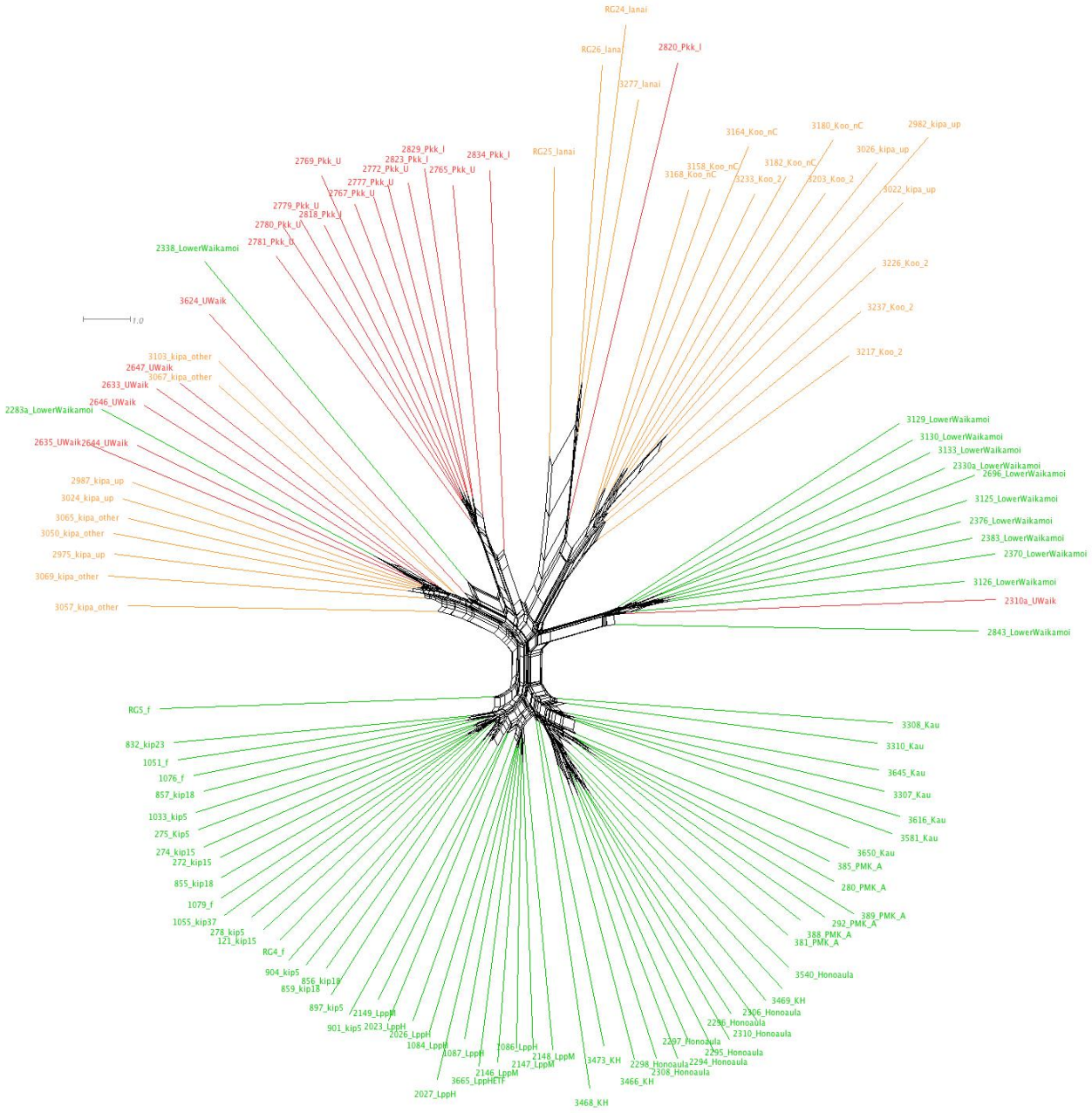



**Figure 23: Phylogenetic network of all the specimens.** The present diagram was created with SplitsTree 4,13,1. The colors represent the original nominal species classification. Green: *T. brevignatha*; Orange: *T. macracantha*; Red: *T. waikamoi.*

**Table 1: Collecting sites**

| Island | Volcano | Site | GPS coordinate | Number of specimens | Specimens |
|---|---|---|---|---|---|
| *T. waikamoi* | | | | | |
| Maui | Haleakalā | Upper Waikamoi TNC | N 20°46'41.60'' W 156°13'43.21'' | 7 | 2310A, 2633, 2635, 2644, 2646, 2647, 3624 |
| | West Maui | Pu'u Kukui upper | N 20°54'52.04" W 156°35'31.84" | 8 | 2765, 2767, 2769, 2772, 2777, 2779, 2780, 2781 |
| | | Pu'u Kukui lower | N 20°54'56.73" W 156°35'35.51" | 5 | 2818, 2820, 2823, 2829, 2834 |
| *T. brevignatha* | | | | | |
| Maui | Haleakalā | Lower Waikamoi TNC | N 20°48'24.22'' W 156°15'18.02'' | 13 | 2283A, 2330A, 2338, 2370, 2376, 2382, 2696, 2843, 3125, 3126, 3129, 3130, 3133 |
| Big Island | Hualālai | Honoaula FR, Makahi St. entry | N 19°43'07.96'' W 155°56'57.34'' | 9 | 2294, 2295, 2296, 2297, 2298, 2306, 2308, 2310, 3540 |
| | Mauna Kea | Laupāhoehoe HETF, HIPPNET | N 19°55'50.43" W 155°17'20.48" | 1 | 3665 |
| | | Laupāhoehoe HETF, High | N 19°54'57.62" W 155°18'22.91" | 6 | 1084, 1086, 1087, 2023, 2026, 2027 |
| | | Laupāhoehoe HETF, Maulua trail. Plot 32, Transect 31 | N 19°53'54.53" W 155°18'46.35" | 4 | 2146, 2147, 2148, 2149 |
| | Mauna Loa | Kīpuka 15 | N 19°40'18.15" W 155°20'18.36" | 5 | 121, 272, 274, 275, 278 |
| | | Kīpuka 18 | N 19°40'27.03" W 155°19'59.45" | 4 | 855, 856, 857, 859 |
| | | Kīpuka 5+37+23 | N 19°39'50.19" | 6 | 832, 897, 901, 904, 1033, 1055 |

| | | | | | |
|---|---|---|---|---|---|
| | | | W 155°21’09.72” | | |
| | | Forest (f2) | N 19°39’57.39” W 155°21’09.91” | 5 | 1051, 1076, 1079, RG4, RG5 |
| | | Ka’u FR, Kapapala access | N 19°20’41.82” W 155°28’03.14” | 7 | 3307, 3308, 3310, 3581, 3616, 3645, 3650 |
| | | Kona Hema TNC | N 19°12’49.41” W 155°49’44.48” | 4 | 3466, 3468, 3469, 3473 |
| | Kīlauea | Pu’u Maka’ala NAR, Army Road | N 19°33'03.36" W 155°13'51.54" | 6 | 280, 292, 381, 385, 388, 389 |
| *T. macracantha* | | | | | |
| Maui | Haleakalā | Ko’olau FR, near Cramp | N 20°45'39.70" W 156°08'34.57" | 5 | 3158, 3164, 3168, 3180, 3182 |
| | | Ko’olau FR, Day 2 | N 20°45'26.85" W 156°08'35.56" | 5 | 3203, 3217, 3226, 3233, 3237 |
| | | Kīpahulu Valley, up | N 20°43'11.06" W 156°05'15.74" | 6 | 2975, 2982, 2987, 3022, 3024, 3026 |
| | | Kīpahulu Valley, other | N 20°42'48.83" W 156°06'12.36" | 6 | 3050, 3057, 3065, 3067, 3069, 3103 |
| Lana’i | Lana’i | Munro trail | N 20°48'28.37" W 156°52'03.96" | 4 | 3277, RG24, RG25, RG26 |

TNC: The Nature Conservancy; NAR: Natural Area Reserve; FR: Forest; HETF: Hawai’i Experimental Tropical Forest; HAVO: Hawai’i Volcanoes National Park

**Table 2: Primer sequences for Positive and Negative controls**

| Gene | Primer name | Sequence |
|---|---|---|
| *Fast Evolving contigs* | | |
| Tb1617 | Tb1617F | TGCCTCTCTTCGAAGGCTTG |
| | Tb1617R | CTTTGTGCTTTCCGGTCAGG |
| Tb6627 | Tb6627F | CGTCATGGACCTGGAGAAGG |
| | Tb6627R | CCACCTCCGTCTTCACCTTG |
| Tb2745 | Tb2745F | CGGACCCTCAGGACAAGATG |
| | Tb2745R | ACGATAATGGGCACGGACAC |
| Tb4420 | Tb4420F | CGCTGGATAAAATTGCGTTTG |
| | Tb4420R | CATTGGATCTTCACTTTTCCATCC |
| Tb4232 | Tb4232F | GGGGAGACAGTACGCTGGTG |
| | Tb4232R | GATTTCACCACTGGCGCTTC |
| Tb4476 | Tb4476F | CGAACCTGAGATGGGCAAAG |
| | Tb4476R | CCAAGGCTAAGCAGGTGGTG |
| Tb4840 | Tb4840F | CCCTCATGAAGCTGGATTCG |
| | Tb4840R | CCATATTGCACCCATCTTTGC |
| Tb6925 | Tb6925F | AAGGTCGGGGAGTGCTTCTC |
| | Tb6925R | TGATGTTGCCCAGCATCTTG |
| Tb5304 | Tb5304F | CGACGTTCTGGACCTCAACC |
| | Tb5304R | GATATTGAGCGGGCAACTCG |
| Tb1230 | Tb1230F | GCTGTTAGCCGGTGGTCAAG |
| | Tb1230R | CACCTTCTCCACCACCTTGG |
| Tb633 | Tb633F | TGAGCTTCCTCGCTCGAAAC |
| | Tb633R | GATGGCGTCGAAGTCGTAGG |
| Tb2656 | Tb2656F | TGGCGATGATGATTGTGGAG |
| | Tb2656R | CCAAAATGGCAAGAGCACTG |
| Tb3418 | Tb3418F | ACCTCCAGTGCCAAGCTCAG |
| | Tb3418R | CCTCTGAGATGCCGTGGAAC |
| Tb2866 | Tb2866F | GTTGCCCATCGAAAAAGCTG |
| | Tb2866R | GTGCCCAGCGAGGTATTCTG |
| Tb1109 | Tb1109F | ATACACCATTTCGCGGCAAC |
| | Tb1109R | CTCCGGTGGTCGGATAAGTG |
| Tb3601 | Tb3601F | CCGAGTGCTCAAAGGTTTCG |
| | Tb3601R | GTCACCAATGCCTCCCATTC |
| *Slow evolving contigs* | | |
| H3 | **H3.1_L** | **GAAGCAGTTGGCAACCAAGG** |
| | **H3.1_R** | **GAAGCTCGGTGGACTTCTGG** |
| | H3.2_L | GAAGCAGTTGGCAACCAAGG |
| | H3.2_R | GGATCAGAAGCTCGGTGGAC |
| Tb35796 | Tb35796.1_L | TGAAAGTCGTGTCCCTGTTCC |
| | Tb35796.1_R | TGGCTCCTTTGACAAGACCAG |
| | Tb35796.2_L | TGAAAGTCGTGTCCCTGTTCC |
| | Tb35796.2_R | TGTCCTTCATTGGCTCCTTTG |

| | | |
|---|---|---|
| Tb38716 | Tb38716.1_L | GTCGCATGCAGATTGTGAGC |
| | Tb38716.1_R | CACATTGGGTGCAAGCAAAC |
| | Tb38716.2_L | GCATGCAGATTGTGAGCAGTG |
| | Tb38716.2_R | CATTGGGTGCAAGCAAACTG |
| Tb22209 | Tb22209.1_L | TCGACAGGGACTCCACCATC |
| | Tb22209.1_R | TGGGGCAGTCGTAATTCACC |
| | Tb22209.2_L | ACAGGGACTCCACCATCTGC |
| | Tb22209.2_R | TGGGGCAGTCGTAATTCACC |
| Tb9898 | Tb9898.1_L | GTGGCAAGCCTTTGAAGGTG |
| | Tb9898.1_R | CAGCACACATGAAGCGATCC |
| | **Tb9898.2_L** | **CGGAAAAGAATGTGCTGTGC** |
| | **Tb9898.2_R** | **CAGCACACATGAAGCGATCC** |
| Tb40119 | Tb40119.1_L | AAGGACTTGCCGAAGTGACG |
| | Tb40119.1_R | TGAGCTGTGGGGGTAAGAGG |
| | **Tb40119.2_L** | **ACGGGACTCTCCTGATGTCG** |
| | **Tb40119.2_R** | **TGAGTGTGAGGGCACCTGAG** |
| Tb37737 | Tb37737.1_L | TGGAGACGTCATGCAAATCG |
| | Tb37737.1_R | GGGACTTCTCCGTGGGATTC |
| | Tb37737.2_L | GCATGAATCCCACGGAGAAG |
| | Tb37737.2_R | TCTTCCAGCCGGTCGTAGTC |

In bold the selected primer pairs (Tb40119.2, Tb9898.2 and H3.1)

**Table 3: Final DNA amounts on different DNA extraction methods: Qiagen (columns) vs 5 PRIME (salt precipitation)**

| Sample | Qiagen (µg) | 5PRIME (µg) |
|---|---|---|
| 2399 | 0.308 | 0.276 |
| 2480 | 1.42 | 0.868 |
| 2510 | Failed | Failed |

**Table 4: Coverage per library**

| | Coverage (x) |
|---|---|
| EXP1 | dups_removed |
| index01 | 7,30567 |
| index02 | 11,62380 |
| index03 | 11,48140 |
| index04 | 10,95560 |
| index05 | 13,67200 |
| index06 | 12,24930 |
| index07 | 9,74584 |
| index08 | 9,06246 |
| index09 | 9,61714 |
| index10 | 12,19960 |

| | |
|---|---|
| index11 | 12,98540 |
| index12 | 13,47230 |
| index13 | 13,63770 |
| index14 | 12,74520 |
| index15 | 11,89760 |
| index16 | 14,30320 |
| index17 | 13,04030 |
| index18 | 14,00630 |
| index19 | 13,38660 |
| index20 | 12,22890 |
| index21 | 12,04270 |
| index22 | 12,96560 |
| index23 | 12,21530 |
| index24 | 14,00330 |
| index25 | 12,93530 |
| index26 | 11,34770 |
| index27 | 12,37040 |
| index28 | 12,45970 |
| index29 | 11,20550 |
| index30 | 11,26590 |
| index31 | 12,47830 |
| index32 | 13,01430 |
| index33 | 12,59290 |
| index34 | 10,37610 |
| index35 | 10,82270 |
| index36 | 12,39090 |
| index37 | 10,23870 |
| index38 | 10,10330 |
| index39 | 11,74630 |
| index40 | 12,72920 |
| index41 | 15,41670 |
| index42 | 10,19100 |
| index43 | 13,24150 |
| index44 | 14,69870 |
| index45 | 13,11610 |
| index46 | 13,18450 |
| index47 | 11,45860 |
| index48 | 10,89920 |
| index49 | 14,31270 |
| index50 | 12,92100 |

EXP2

| | |
|---|---|
| index01 | 3,71541 |

| | |
|---|---|
| index02 | 5,11161 |
| index03 | 5,14204 |
| index04 | 5,54592 |
| index05 | 5,69213 |
| index06 | 3,91954 |
| index07 | 4,50947 |
| index08 | 4,14700 |
| index09 | 3,87366 |
| index10 | 3,52750 |
| index11 | 3,32319 |
| index12 | 5,22478 |
| index13 | 3,51048 |
| index14 | 5,11311 |
| index15 | 4,43346 |
| index16 | 4,66858 |
| index17 | 5,20497 |
| index18 | 4,94554 |
| index19 | 5,76792 |
| index20 | 4,42241 |
| index21 | 3,37431 |
| index22 | 3,21609 |
| index23 | 2,52368 |
| index24 | 2,83486 |
| index25 | 2,88852 |
| index26 | 2,92590 |
| index27 | 2,00359 |
| index28 | 3,09761 |
| index29 | 2,24848 |
| index30 | 2,31810 |
| index31 | 2,12646 |
| index32 | 2,95443 |
| index33 | 3,55412 |
| index34 | 2,12118 |
| index35 | 3,31478 |
| index36 | 3,31478 |
| index37 | 3,20859 |
| index38 | 3,71182 |
| index39 | 3,71182 |
| index40 | 3,19103 |
| index41 | 3,47370 |
| index42 | 4,23431 |
| index43 | 3,97969 |
| index44 | 4,94897 |

| | |
|---|---|
| index45 | 3,72533 |
| index46 | 4,18738 |
| index47 | 3,79081 |
| index48 | 4,04179 |
| index49 | 4,19151 |
| index50 | 3,34810 |
| index51 | 3,05711 |
| index52 | 3,24466 |
| index53 | 3,92396 |

EXP3

| | |
|---|---|
| index01 | 1,89935 |
| index02 | 2,74401 |
| index03 | 2,14778 |
| index04 | 2,28612 |
| index05 | 1,65833 |
| index06 | 2,04559 |
| index07 | 2,16062 |
| index08 | 1,56869 |
| index09 | 1,83661 |
| index10 | 2,90754 |
| index 50 | 1,95604 |
| index51 | 1,99886 |
| index52 | 2,09183 |
| index11 | 1,66693 |
| index12 | 1,95914 |
| index13 | 1,86667 |
| index14 | 1,84903 |
| index15 | 1,59515 |
| index16 | 1,77717 |
| index17 | 2,01338 |
| index18 | 1,77935 |
| index19 | 1,61564 |
| index20 | 2,12558 |
| index21 | 2,39638 |
| index22 | 2,25387 |
| index23 | 1,46423 |
| index24 | 1,97836 |
| index25 | 1,75109 |
| index26 | 1,71362 |
| index27 | 1,96016 |
| index28 | 1,63715 |
| index29 | 1,95613 |

| | |
|---|---|
| index30 | 1,90952 |
| index31 | 1,93803 |
| index32 | 1,78989 |
| index33 | 2,30363 |
| index34 | 1,95274 |
| index35 | 1,731050 |
| index36 | 1,667610 |
| index37 | 1,751360 |
| index38 | 1,847480 |
| index39 | 1,792950 |
| index40 | 1,522420 |
| index41 | 1,784730 |
| index42 | 1,925910 |
| index43 | 1,905510 |
| index44 | 1,755880 |
| index45 | 1,547590 |
| index46 | 1,722190 |
| index47 | 1,658500 |
| index48 | 1,380510 |
| index49 | 2,050130 |

Red = *T. brevignatha*
Blue = *T. macracantha*
Green = *T. waikamoi*
Purple = *T. anuenue*

**Table 5: Pair-wise Fst for *T. brevignatha***

| | Honoaula | LPP | Kīpuka | PMK | Ka'u | KH | Maui |
|---|---|---|---|---|---|---|---|
| Honoaula | - | | | | | | |
| LPP | 0.122 | - | | | | | |
| Kīpuka | **0.182** | 0.064 | - | | | | |
| PMK | 0.082 | 0.081 | 0.158 | - | | | |
| Ka'u | 0.104 | 0.113 | **0.168** | 0.080 | - | | |
| KH | 0.112 | **0.169** | **0.219** | 0.142 | 0.118 | - | |
| Maui | **0.324** | **0.324** | **0.351** | **0.310** | **0.275** | **0.287** | - |

Globa Fst: 0.206. In bold the highest comparisons.
LPP: Laupāhoehoe; PMK: Pu'u Maka'ala; KH: Kona Hema

**Table 6: Nucleotide diversity and neutrality tests for *T. brevignatha* (full values)**

| | Tajima's D | Theta Pi | Number of sites | Pi (ThetaPi/Sites) |
|---|---|---|---|---|
| Big Island | -2.070457 | 3498.810904 | 1255696 | 0,0027863519 |
| Hualālai | -1.062081 | 2090.127826 | 1062469 | 0,0019672365 |
| Ka'u | -1.199406 | 1702.226838 | 652630 | 0,0026082571 |
| KonaHema | -1.141110 | 1335.309455 | 530047 | 0,0025192284 |
| Kīlauea | -1.067374 | 1898.976772 | 999719 | 0,0018995105 |
| MaunaLoa | -1.737084 | 2706.706961 | 1177933 | 0,0022978446 |
| MaunaKea | -1.426162 | 2311.944077 | 1095717 | 0,0021099828 |
| Maui | -0.902656 | 5248.234677 | 1143531 | 0,0045894993 |

**Table 7: Nucleotide diversity and neutrality tests for *T. waikamoi* (full values)**

| | Tajima's D | Theta Pi | Number of sites | Pi (ThetaPi/Sites) |
|---|---|---|---|---|
| Pu'u Kukui | -1.909211 | 3896.725719 | 1355519 | 0,0028747113 |
| U_Waikamoi | -1.572736 | 4316.126711 | 1158552 | 0,0037254493 |

**Table 8: Pair-wise Fst for *T macracantha* (full values)**

| | Kīpahulu Valley | Ko'olau* | Lana'i |
|---|---|---|---|
| Kīpahulu Valley | - | | |
| Ko'olau* | 0.528445851578741 | - | |
| Lana'i | 0.572279008218038 | 0.269045810730168 | - |

*The Ko'olau population includes the three specimens from Kīpahulu Valley that cluser with them in PC1 vs PC2.

Global Fst is:0.474284052751551

**Table 9: Nucleotide diversity and neutrality tests for *T. macracantha* (full values)**

| | Tajima's D | Theta Pi | Number of sites | Pi (ThetaPi/Sites) |
|---|---|---|---|---|
| Kīpahulu Valley | -1.999114 | 4856.689301 | 2043612 | 0,002376522 |
| Ko'olau | -2.014889 | 7424.613747 | 1893573 | 0,003920955 |
| Lana'i | -1.893574 | 5252.011775 | 1774228 | 0,002960167 |

**Table 10: Coverage of the specimens present in WaikMacBrev clade and *T. waikamoi* clustering with other species**

| Specimen | Coverage (x) |
|---|---|
| **WaikMacBrev clade** | |
| Exp3_index09 | 1,83661 |
| Exp2_index27 | 2,00359 |
| Exp2_index31 | 2,12646 |
| Exp2_index29 | 2,24848 |
| Exp2_index30 | 2,31810 |
| Exp2_index23 | 2,52368 |
| Exp2_index25 | 2,88852 |
| Exp2_index32 | 2,95443 |
| Exp2_index28 | 3,09761 |
| Exp2_index40 | 3,19103 |
| Exp2_index37 | 3,20859 |
| Exp2_index21 | 3,37431 |
| Exp2_index41 | 3,47370 |
| Exp2_index38 | 3,71182 |
| Exp2_index39 | 3,71182 |
| Exp1_index50 | 12,92100 |
| Exp1_index43 | 13,24150 |
| ***T. waikamoi* clustering with other species** | |
| Exp3_index10 | 2,90754 |
| Exp2_index51 | 3,05711 |

Red: *T. brevignatha*
Blue: *T. macracantha*
Green: *T. waikamoi*

## SUPPLEMENTARY MATERIAL

### Probe Design

The Probe design used a modified version of the pipeline and the scripts deposited in the MVZ Github site (https://github.com/MVZSEQ). Detailed instructions for each script can be found on that site. We follow the next steps:

1.- *Quality control and sequence assembly:* The raw reads were trimmed based on per-read quality scores, primer artifacts were removed, and reads were filtered for length using the FASTX Toolkit (Hannon, hannonlab.cshl.edu/fastx_toolkit/). All sites in a read occurring after a quality score of 20 or less were removed and all reads less than 30 bp in length afterwards were deleted. The first nine bases of each read were removed to eliminate primer/sequencing artifacts. The quality of processed reads was assessed using FASTQC (FastQC, bioinformatics.babraham.ac.uk/projects/fastqc/). The individually processed paired-end read files were resynchronized using a custom script. Each transcriptome was assembled using the Trinity pipeline (Grabherr at al., 2011) requiring a minimum contig length of 100 bases and minimum kmer coverage of 2.

2.- *Contaminant removal:* The resulting contig files were filtered for mitochondrial (BLASTX), ribosomal (BLASTN), and contaminant sequences as determined by BLASTX and MEGAN4 analyses of the deeply sequenced *T. grallator* somatic transcriptome (Croucher et al., 2013) (BLASTN) using custom BLAST databases and searched (BLASTX) against the NCBI non-redundant (nr) protein database.

3.- *Visualization of BLASTx result*: Using MEGAN (Hudson et al., 2007) we classified each read into the lowest taxonomic unit. For some genes there is good evidence that they belong to a specific low taxonomic unit, so they were included on that group. Others were classified into a more generic node, because there is no evidence that they go into lowest levels. We were as stringent as we could in order to not add contaminants. All the sequences with hits to non-metazoan or nematodes were removed.

4.- *ORF prediction:* Putative open reading frames (ORFs) were predicted from the contigs remaining after contaminant removal using Trinity's transcripts_to_best_scoring_ORFs.pl script, minimum length 50 amino acids.

5.- *Transcriptome coverage estimation:* It was assessed by comparing each assembly to a standard set of 248 core eukaryotic genes (CEGs) using the program Cegma (Parra et al., 2007). These CEGs were derived from the eukaryotic orthologous groups (KOGs), a subset of the Cluster of Orthologous Groups (COGs) of Proteins database (Tatusov et al., 2003).

6.- *Select gene families ("comps" is the denomination of gene families from TRINITY) with single representative in order to avoid paralogs*: The following line of BASH code generates a list without sequences. It uses as an input the result of TRANSDECODER.

```
>grep "^>" Name_nrBLAST_out-ex.fasta.transdecoder.cds | cut -d"_"
-f1 | sort | uniq -u > Name_cds_uniq_comps.txt
```

7.- *Change line break by an underscore*: This is necessary for the downstream analysis.

```
>sed 's/$/_/g' Name_cds_uniq_comps.txt > Name_cds_uniq_comps2.txt
```

8.- *Estimate the number of ORFs:* We used the following BASH command:

```
>wc -l Name_cds_uniq_comps2.txt
```

9.- *Create FASTA file with no paralogs*: Search in the fasta file (<Name_nrBLAST_out-ex.fasta.transdecoder.cds>) for all the sequences that are singleton families (<Name_cds_uniq_comps.txt>; this is the output file of step 6). This will create a file that has all the information of the sequences present in the list (<Name_cds_NoParalogs.fasta>).

```
>cat Name_cds_uniq_comps.txt | while read line; do grep -A 1
$line Name_nrBLAST_out-ex.fasta.transdecoder.cds_nent.fasta; done
> Name_cds_NoParalogs.fasta
```

10.- *GC filer:* This filter takes out sequences with %GC below 30 and above 70. This program is available upon request and it is not in the MVZ Github site. Script used: GC_percentage_calc.py

11.- *kmer count:* The kmers were counted using the JELLYFISH program implemented on TRINITY (Grabherr et al., 2011). We used a kmer size of 15 (-m 15)

12.- *Write the tab delimited kmer count file:* We created that file using the JELLYFISH program implemented on TRINITY (Grabherr et al., 2011).

13.- *Indexed file*: This file is created using the program: indexer_15mers.pl.

14.- *Symbol changes*: For downstream analysis we did some notation changes, which make more convenient the use of the different files.

```
>sed 's/>/>chr_/g' Name_cds_NoParalogs_goodGC.fasta | sed s'/
/_/g' > Name_cds_NoParalogs_goodGC_4ProbeDesign.fasta
```

15.- *Create the target region file*: It uses as an input the file <Name_cds_NoParalogs_goodGC_4ProbeDesign.fasta> in the program target_file_maker.py, which is available upon request.

16.- *Array design*. Using the script array_design_modified.pl the probes were design. We used a probe length of 60 bp (-p 60) and a tailing density of 2 bp (-t 2). We did not

include any flanking region (-f 0), because for spiders the Intron-Exon boundaries are not known. For the same reason the tailing density was so high.

17.- Preparation of the array by Agilent (Santa Clara, CA. USA).

Another important thing to consider during the array design is to remove mitchondrial genome to avoid over representation during the capture experiment. The mitochondrial genome has so many copies and their sequences could completely out compete other sequences to capture. Only one mitochondrial gene is included in the chip in order to have a control to test for empirical error rates if it is needed. Alternatively a sex-linked gene could be used for this purpose. The only caveat is that it will only be possible to use on the samples where this sex chromosome is present. In general as a way to check for empirical error rate, it is necessary to use a haploid marker.

**COT 1 DNA library preparation**

We started with 13 μg. of high molecular weight DNA. The DNA quality was verified on an agarose/TBE gel and the amount calculated based on a concentration estimation measured with Qubit® 2.0 Fluorometer (Life Technologies).

The DNA was fragmented using a Bioruptor® Standard (Diagenode) set in the "High" mode and following a fragmentation cycle of 3.5 minutes on, change half of the cold water and then 3.5 minutes on again. It was repeated 3 times or until reach a fragment distribution between 100 bp and 500 bp, centered around 250 bp. The fragment distribution was checked on an agarose/TBE gel. After fragmentation the remaining amount was 4.31 μg. Then, we did an ethanol precipitation to purify the DNA.

The DNA denaturation was done for 5 to 10 minutes at 95°C. Then, 12 volumes of SSC were added and the reannealing was done at 60°C for 22 minutes. This amount of time was calculated based on the formula described in Trifonov et al., 2009. Immediately after the 22 minutes we proceed to the S1 nuclease (Thermo Scientific) hydrolysis, which removed the single stranded DNA. All the reagents involved in this enzymatic digestion and free nucleotides were removed with another ethanol precipitation.

The isolated COT1 DNA is immortalized by a regular library preparation for NGS (Meyer and Kircher, 2010). We used custom designed COT1 adaptors (cot-1 Forward adapter: AGCTATCAGACTCGGACTACTGATGCAGTG; cot-1 Reverse adapter: CACTGCATCAGT). For the indexing PCR step we used the same primer for the reverse and forward direction (cot 1 primer: AGCTATCAGACTCGGACTACTG). These sequences were later removed by an infrequent restriction enzyme (BstI) (New England Bio Labs), which has a restriction site on the COT1 DNA primers. We used 2 units (2U) of enzyme per each μg of COT1 DNA. After this enzymatic digestion we proceed to clean the product by a beads clean up protocol as described on Meyer and Kircher, 2010.

A final amount of 50 μg of COT1 DNA was required for each hybridization experiment (3 in total). More than 100 PCR reactions were required to reach this quantity. Note, that the PCR product had to be purified with a beads clean up before use it. We did 15 cycles of the indexing PCR protocol described on Meyer and Kircher, 2010.

The amount of template that we found optimizes the amount of PCR product and how much of the original COT1 DNA library we were able to use per reaction was 20 ng.

Higher numbers of cycles lead to problems due to the formation of secondary structures and daisy chaining. These high molecular weight structures were not digested by the restriction enzyme, which removed the indexing primers. Initially we also tried using PCR product as a template for a new PCR reaction, however it was found not to be a good approach because after 2 or 3 iterations there was a substantial accumulation of high molecular weight DNA.

The final 50 µg of COT 1 DNA had to be concentrated on a SpeedVac (CentriVap DNA Concentrator, LABCONCO) in order to reach the volume required in the hybridization experiment (Hodges et al., 2009)

**Genomic library preparation**

A total of 116 libraries were prepared following the protocol described on Meyer and Kircher, 2010. 70 corresponded to *T. brevignatha*, 20 to *T. waikamoi* and 26 to *T. macracantha*. The details of the localities are on Table 1. The DNA quality was verified on an agarose/TBE gel and the amount calculated based on a concentration estimation measured with Qubit® 2.0 Fluorometer (Life Technologies) (2 replicates). The amount of starting material ranged from 400 to 500 ng depending on the sample.

The DNA was fragmented using a Bioruptor® Standard (Diagenode) set in the "High" mode following a fragmentation cycle of 3.5 minutes on, change on half of the cold water and then 3.5 minutes on again. It was repeated 3 times or until it reached a fragment distribution between 100 bp and 500 bp, centered around 250 bp. The fragment distribution was checked on an agarose/TBE gel. The first 50 samples were prepared using cold water and ice in the Bioruptor® Standard. For the following 66 we used a water cooler system (Diagenode). The fragmentation protocol remained the same. Right after fragmentation we performed a beads clean up in order to concentrate the samples and eliminate the small fragments of DNA.

The only difference with respect to the original protocol is that we used Sera-Mag magnetic beads (GE Healthcare Life Sciences) instead of AMPure® XP beads (Agencourt). However, it does not affect the result of the protocol (Smith *pers. comm.*). For every batch of libraries prepared we used a 150 bp positive control (fragment of H3) to check the success of the adaptor ligation, before proceed to indexing PCR. We also carried a negative control (buffer EBT) to control for contaminations.

For the indexing PCR of the first 50 samples we tested for 8, 9, 10 and 11 cycles of amplification. As we saw an increase in the concentration, we showed that at 11 cycles had not reached plateau. We chose 8 cycles to reduce the potential number of PCR duplicates. For the other 66 samples we performed 10 cycles in order obtain a better yield. In the cases where there was not enough PCR product for the multiplexing step we repeated the library amplification step. The results of the indexing PCR were verified on an agarose/TBE gel in order to observe the shift in the fragment distribution (126 bp more heavy). The final elution of the beads was cleaned up after indexing PCR was performed with water instead of buffer EBT. The rationale for this step was that, after multiplexing, the samples have to be concentrated on a SpeedVac machine (CentriVap

DNA Concentrator, LABCONCO) and the use of EBT buffer will increase the salt concentration, which might affect the hybridization.

**Hybridization Experiment**

The Post-Capture elution was concentrated using a SpeedVac (CentriVap DNA Concentrator, LABCONCO) until approximate 245 ul. We used an aliquot to test the number of cycles (16 and 18) for the whole library amplification. As there was an increase in concentration between 16 and 18 cycles, we decided to amplify the rest of the Post-Capture with 16 cycles as we could be sure that number of cycles was below PCR plateau. The Master mix consisted of: 31.1 ul of concentrated Post-Capture elution, 10 ul of Phusion buffer (5x), 1 ul of IS5 primer (10 μM), 1 ul IS6 primer (10 μM), 0.5 ul Phusion Hot Start High-Fidelity DNA Polymerase (2 U/μL) and dNTPs (25 mM each).

For Experiments 2 and 3 we initially tested for a smaller number of cycles in order to reduce the possibility of PCR duplicates. However we were not able to obtain any DNA in either 14 or 16 cycles. For Experiment 2 we ended up pooling the PCR products of 2 reactions of 20 cycles and 2 reactions of 21 cycles. While for Experiment 3 we pooled the PCR product of 2 reaction of 21 cycles and 2 reactions of 23 cycles.

The number of extra cycles that we performed for each sample was guided by the concentration measured by Qubit® 2.0 Fluorometer (Life Technologies) that we obtained for each case. The main problem with increasing the number of amplification cycles was the inevitable production of PCR duplicates, which decreases the sequencing depth of the genes of interest. The bioinformatics treatment that we used for the removal of PCR duplicates is explained in the section "Data preparation".

**Post-Capture controls**

*Enrichment and depletion*

For the qPCR validation, all the Pre and Post-Capture aliquots were diluted to a final concentration of 2 ng/ul. The master mix corresponded to: 2 ul of DNA (2 ng/ul), 6 ul of water, 10 ul of qPCR reaction mix, 1 ul Forward primer and 1 ul Reverse primer.

For each experiment we had two positive and one negative control. Those controls were tested in duplicates using as a template aliquots from: Pre-Capture, Post-Capture and qPCR Negative (no template). Reading the amplification plot we determined the shift in the number of cycles associated with the enrichment (Positive control) and depletion (Negative control).

A reduction in the number of cycles required to reach plateau in the Post-Capture for the positive control sequences is indicative of a success in the enrichment. In the same way, an increase in the number of cycles required to reach a plateau in the Post-Capture for the negative control is indicative of the success in the depletion of the non-specific sequences. It is not expected an absolute depletion of the non-captured sequences as it is known that the chip based approach captures a lot of unspecific fragments.

*Bioanalyzer*

In order to ensure that the correct size distribution was sent to the sequencing facility, it was estimated with a Bioanalyzer (Agilent Technologies) for each experiment. We performed two replicates and one negative.

**Analytical pipeline**

***Data preparation***

All the scripts used on the steps 1 to 6 for data preparation are on the MVZSEQ github web site (https://github.com/MVZSEQ/Exon-capture).

1.- *Merging left and right files*: Note that it is necessary to change the names into the same format as the genomic DNA reads. Script used: 1-pre-cleanup.pl.

2.- *Remove contaminants, duplicates and low complexity sequences*: The "libinfo file" is a tab-separated file with the equivalences of the bar codes and name of the specimens (it expects a header):

```
INDEX      SPECIMEN
IndexXXX   specimenXXX
indexYYY   specimenYYY
```

The insert size (-e) was 185. The contaminant file (-c) consisted in a human and bacterial genomic data set to screen for contaminants.

The output files are: 1, 2, u, duplicates, contaminants and low complexity. Script used: 2-scrubReads_Nov12.pl.

3.- *kmeric assembly:* First, all the files 1, 2 and u for each library were combined with the respective 1, 2 and u files from the other libraries within the same species. It produced three combined files (1, 2 and u) for each species.

Then, we run ABySS (Simpson et al., 2009) to create four different assemblies with a different kmer size each (k=23, 32, 45, 64). A sliding window approach was used to create the different kmers sizes and ABySS only saves the unique kmers into a library. After that, the program produces De Bruijn graphs with all the unique kmers. These graphs were used as a reference to define the contigs (these contigs have all the information of the polymorphism). All the individual reads of each library were mapped to these contigs, the script used was: 3-generateAssemblies_LOCAL.pl.

The result is a single FASTA file for each kmer size. These files are concatenated into a unique file.

5.- *Merge kmeric assemblies:* As a way to remove redundancy, the script 5-finalAssembly.pl takes information from all the kmer sizes to produce a new file with the best contigs.

6.- *Pseudo Reciprocal BLAST:* In order to find assemblies derived from targets we performed a Pseudo Reciprocal BLAST of the contigs produced in the Merged kmeric assembly against the sequences of the original target. This could be considered as an *in silico* repetition of the hybridization experiment. All the unspecific contigs will not be able to find a hit in the original target, which will eliminate them. It is important to mention that this step corresponds to a BLAST and not to an assembly, because the assembly algorithm would have problems with the intronic flaking regions next to each exon (100 bp approximate). The script used for this is: 6-pseudoReciprocalBlastToMakeFinalTargets.pl.

The output is the <intarget.fasta> file, which is equivalent to the original set of target sequences, but including the intronic flanking regions. This FASTA file will be used as a reference (-ref) for downstream analysis and as we do not have an ancestral reference to polarize the characters, it will be also use as an ancestral file (-anc). It means that the estimations of Site Frequency Spectrum (SFS) will correspond to folded SFS.

7.- *Mapping of each library:* For each library we mapped the reads present in the files 1, 2 and u. The average insert size of the libraries (-i) was 185 bp. For the other parameters we used the default settings. This step creates the bam files that will be used for the data analysis. The script used is: Alignment-by-novoalign.pl. It as a dependency the program NOVOALIGN (www.Novocraft.com).

8.- *Duplicate removal:* The elimination of duplicate sequences is always an important step in the data preparation. If this step is omitted the final coverage will be over estimated, which will negatively affect downstream analysis. This is especially relevant in the case of using genotype likelihood estimations. The program PICARD (http://broadinstitute.github.io/picard/) uses bam files as input to identify and eliminate duplicates.

9.- *Coverage estimations:* In order to have a realistic estimation, these calculations have to be done in the bam files after duplicate removal. To calculate coverage on each library we used the tool DEPTH in the program SAMTOOLS (Li et al., 2009). The following is the line of command to obtain the average per library:

```
>samtools depth -q 20 individual1.bam | awk '{sum+=$3} END { print "Average = ",sum/NR}'
```

***Data analysis***

Given the low coverage of our data and its demonstrated high performance in middle coverage data sets, for the data analysis we used the program Analysis of Next Generation Sequencing Data (ANGSD). It can be downloaded from: http://popgen.dk/wiki/index.php/ANGSD. Before using ANGSD it is necessary to select the sites that will be used in the analysis. This selection requires applying several site-specific filters implemented on SNP Cleaner (https://github.com/MVZSEQ/Data-quality-control). After the allele frequency estimation based on genotype likelihood estimations has been done, it is possible to use the output files from ANGSD in other analysis. They

are implemented in other programs as ngsTOOLS (Fumagalli et al., 2014), SplitsTree 4,13,1 (Huson and Bryant, 2006), ngsADMIX (Skotte et al., 2013), PopGenTools (https://github.com/MVZSEQ/Population-genomics) and several R packages (R Development Core Team, 2008). The analytical pipeline that we used is as follows:

1.- *VCF file, required for SNP Cleaner:* This file can be created with the program MPILEUP implemented in SAMTOOLS (Li et al., 2009). The input for this program corresponds to the reference file (FASTA file produced on step 7 of the "Data preparation")

In case the version SNPcleaner231.pl is used, it would be necessary to create a Pileup file. That was not our case, as we used the version SNPcleaner02.pl.

2.- *SNP cleaner:* We used the version SNPcleaner02.pl. It requires the VCF file on the same folder as the program (https://github.com/MVZSEQ/denovoTargetCapture). We used a minimum coverage (-d) of 3x in 70% of the libraries (-k). Note the parameter –k is the actual number of libraries.

3.- *keep file:* One of the output of SNPCleaner is a bed file <*.bed>. It corresponds to a file with all the sites present on each contig that passed the filters. The file has three columns, which correspond to: contig, start position and end position. Because the program is looking sites, the file will look like this:

```
m.0001      0     1
m.0001      1     2
m.0001      2     3
…
```

The actual position of the site corresponds to the third column. In order to create the keep file we selected only the columns 1 and 3.

4.- *rf file:* The rt file corresponds to all the unique contigs that are present in the keep file. We generate it using the following command line:

```
>cat Name.keep | cut -f1 | sed "s/$/\:/g" | uniq > Name_rf
```

5.- *bam.filelist:* The bam.filelist is a text file with the paths to all the bam files to be used. To create it, we open the folder with all the bam files and used the following command line:

```
>find `pwd` *bam | grep "^/Home" | grep ".bam$" > bam_filelist.txt
```

6.- *ANGDS:* The details about how to run ANGSD can be found on the website of the program. To estimate the SFS and use it as an input for ngsTOOLS we used the following options: <-only_proper_pairs 0 -minMapQ 0 -minQ 20 -GL 1 -fold 1 -doSaf 1 -doMajorMinor 1 -doMaf 2 -doPost 1 -doGeno 32> We also tried several values for "-

SNP_pval", which corresponds to a filter that penalizes mostly the sites present in the first SFS category.

7.- *Site Frequency Spectrum (*.sfs):* We used the program emOptim2 and a *.saf file as input. Note that the *.saf file correspond to one of the outputs of ANGSD. For plotting the SFS in R, the first element of the file has to be deleted as it corresponds to the invariant sites.

8.- *Principal Component Analysis (PCA):* The first step is to create a Covariate file. For this we used the program ngsCOVAR present in the package ngsTOOLS (Fumagalli et al., 2014). The number of sites corresponds the number of lines in the *.mafs file (one of the outputs of ANGSD). Note that this file has a header, so the total number of sites is the number of lines minus one. The output can be plotted in R.

9.- *Fst calculations for two populations*: For a couple of populations, after running ANGSD we merged the files that contained all the used sites (*.pos). Then, we selected only the unique sites. This can be accomplished with the following line of command:

```
>cat population1.saf.pos population2.saf.pos | sort | uniq -d > intersect.keep
```

After creating a file with all the common sites <intersect.keep>, we re-run ANGSD for each population using that file instead of the original keep file for the sites (-sites). Also, we had to remove the option -fold 1. This last action makes ANGSD run the analysis with the default with is unfolded genome. For each population a new SFS is created using the program emOptim2, but considering 2 times the number of individuals as it is an unfolded analysis. The number of sites is estimated the same way as to get the Covarfile when using ngsCOVAR.

The new pair of SFS files is used to estimate the 2dSFS following this command:

```
>ngs2dSFS -postfiles pop1.sfs pop2.sfs -outfile pop1pop2.2dsfs -relative 1 -nind 10 11  -maxlike 1 -block_size 10000 -islog 1  -nsites NNN
```

The 2dSFS files provides a prior that is later used in combination with the individual SFS files to estimate the pairwise Fst using the program ngsFST. For that we used the following command line:

```
>ngsFST -postfiles  pop1.sfs pop2.sfs -priorfile pop1pop2.2dsfs -outfile  pop1pop2.FST -nind 10 11 -nsites NNN  -block_size 1000 -islog 1
```

Finally, the Global FST corresponds to: Summation 1$^{st}$ col (within)/ Summation 2$^{nd}$ col (between+within).

10.- *Pairwise Fst for more than two populations*: The program PopGenTools_2.76.pl (https://github.com/MVZSEQ/Population-genomics) has an implementation to run all the previously described process in multiple comparisons across different populations. We applied that for each population.

11.- *Tajima's D*: First, we estimated SFS and theta for each population. It cannot be the same SFS file used for Fst calculations, because for Tajima's D has to be folded. As part of the options we had to include <–GL 1 –fold 1 –anc>. The –anc option corresponds to the "ancestral genome". As we did not have an outgroup, we just used the same file as the reference (-ref). The result will be a file with the Tajima's D for each contig. In order to get a global Tajima's D we created a new theta file where all the sites were present in the same chromosome. In order to do that, we replaced all the values of the #Chromosome (column 1) for the same number. All of this was done with the script GlobalTajimaInput.pl (available upon request). Then we recompressed the file, and used the thetaStat program to estimate Tajima's D (Korneliussen et al., 2013).

12.- *Nucleotide diversity (Pi)*: From Tajima's D output file *.thetas.gz.pestPG we took the value of theta P (tP) and divided by the number of sites used on that population.

13.- *ngsADMIX*: In order to verify the most likely distribution of the samples given a fixed number of groups and detect admixture we used the program ngsADMIX (Skotte et al., 2013). The first step was to map all the specimens against the same reference (in this case we used *T. brevignatha*), which required generating new bam files and repeating the Mapping step with NOVOALING (Step 7 of "Data preparation"). Then, using those bam files (duplicates removed) we generate a Beagle file using ANGSD. The commands for generating the Beagle file and running ngsADMIX are on the website of the program (http://www.popgen.dk/software/index.php/NgsAdmix).

For each value of k, we performed 10 repetitions and plotted the one with the highest likelihood value that was repeated (Skotte *pers. comm.*). We tested the k values from 1 to 9.

To have an idea of the optimal value of k, we implemented on an Excel spread sheet the Evanno method (Evanno et al., 2005) to select the optimal value of k. This implementation already exists for microsatellite data (STRUCTURE HERVASTER, Earl, 2011), but it does not exist for NGS applications.

The plotting of the results was done in R following a modified version of the code that appears on the website of the program. It was provided by one of the authors of ngsADMIX (Skotte *pers. comm.*).

14.- *SplitTree*: In order to represent the phylogenetic relationships among the different populations of these three closely related species, we constructed an unrooted phylogenetic network using the program SplitsTree 4,13,1 (Huson and Bryant, 2006). This approach allows representing a collection of incompatible trees that are equally consistent with the data given a model of evolution (Bandelt and Dress, 1992). This situation will be common on events of hybridization, horizontal gene transfer, recombination, and gene formation/duplication/loss. In particular, due to the known role

of admixture in early speciation events (Seehausen, 2004; Martin et al., 2013; Rius and Darling, 2014), this method is particularly suitable for our system.

We took the clean 1, 2 and u files for each specimen and mapped them against the same reference (*T. brevignatha*). The new bam files were used to re-run ANGSD, but in this case we added a few more extra filters < -doGeno 2 -SNP_pval 0.1 –postCutoff 0.75 -geno_minDepth 3 -minInd 81 -doCounts 1 -doGlf 4 >. These filters were applied in order to obtain high quality data and remove sites with lower coverage (-geno_minDepth) than a determined cutoff (-minInd). In the case of our analysis we chose to keep sites with a minimum of 3x in at least 70% percent of the individuals (81 specimens out of the 116 analyzed).

The resulting geno file (*.geno) it is used as an input for PopGenTools_2.80_New.pl, this new version it is still not on the GitHub site. It was used to produce an *.adegenet file, which is an intermediate file for Splitstree 4,13,1. The line of command used was:

```
>perl ../PoulationGenomics/PopGenTools_2.80_New.pl Adegenet -g <genofile> -n <number of specimens> -s <number of sites. wc –l *.pos> -o saa.adegenet -h 1 –m 0.3
```

The option -m corresponds to the percentage of individuals with missing data for a particular site that is allows. Above that percentage the site was removed. A good recommendation is 0.3 (Bi *pers. comm.*). We used that filter to take care of missing data.

In the *.adegenet file columns correspond to the sites and the lines to the individuals. It was used as an input file for the R script splittrees_generator.docx (available upon request), which produces a file *.D. This file was used as an input for the script splittree.pl (available upon request) that generates a nexus file that can be opened by SplitsTree 4,13,1.

# CHAPTER 4

**Color polymorphism on *Selkirkiella* spieders (Theridiidae) from the Juan Fernández archipelago and the Valdivian temperate rainforest.**


## ABSTRACT

Convergent evolution occurs on a wide variety of traits present on different groups of organisms. The convergence in color polymorphisms has been widely studied, as it corresponds to a unique situation where the convergence leads to diversity. In the spider family Theridiidae, the independent evolution of the abdominal color polymorphism has been well described in at least four species (*Enoplognatha ovata*, *E. latimana*, *Theridion grallator* and *T. californicum*). Among the shared characteristics on these cases of color polymorphism are: (1) the presence of "Yellow" as the double recessive and more common variant, (2) single loci Mendelian inheritance (except in Big Island population of *T. grallator*), (3) constant frequencies of variants among poorly connected populations and (4) evidence of the action of natural selection. The genus *Selkirlkiella* from the temperate rainforest of southern South America has several species with some degree of color polymorphism. Here, we documented the color polymorphisms of *S. alboguttata* (Robinson Crusoe Island) and *S. luisi* (Valdivia, Chile). *Selkirlkiella alboguttata* displays 6 morphs and *S. luisi* two morphs. The "Yellow" morph was the more common in both species. Based on a molecular phylogeny, we confirm that this genus is closely related to *Enoplognatha*. The presence of color polymorphism in this genus appears to be an event of convergent evolution at the family level, while between species it is likely due to common ancestry. There is also evidence of other cases of color polymorphism in the family. Finally, this phenomenon seems to be associated with the "under leaf community".

## INTRODUCTION

Convergent evolution occurs at different levels of biological complexity, from molecules to behaviors, including morphology and life history traits (Manceau et al., 2010). The origin of convergences can be explained by the adaptation of organisms to similar environmental conditions or life styles (niche occupancy) (Harmon et al., 2005). However, the molecular and\or mechanistic bases for the convergences are not always the same (Manceau et al., 2010). In the context of adaptive radiations convergence occurs between species that colonize similar microhabitats. Examples for this are the *Anolis* lizards in the Caribbean, *Mandarina* snails in the Bonin Islands, *Tetragnatha* spiders in Hawai'i and the Hawaiian Honey Creepers (Gillespie, 2013).

The convergence in color polymorphisms has been widely studied (Protas and Patel 2008), because correspondence of color to unique situations often leads to diversification and higher diversities. The evolutionary origin for this condition is related to one or a combination of three factors: sexual selection (guppy fish, *Poecilla reticulata*; Hughes et al., 2013), predator avoidance (peppered moth; Kettlewell, 1958) or drift (*Cepaea* land snails; Cain and Currey, 1963) (Gray and McKinnon, 2006). A unique opportunity to study the genetic and ecological mechanisms underneath those traits occurs when closely related species evolve convergent phenotypes. For example, in the butterfly genus *Papilio* (Nijhout 2003) and *Heliconius* (Kapan et al 2006), the color variation has been associated with Müllerian mimicry. On the other hand, in several species of African cichlid fishes (Seehausen et al, 1999), it is possible to see different color variant associated with sexual selection (Kocher, 2004).

In the spiders from the cobweb family Theridiidae, the independent evolution of the abdominal color polymorphism has been well described in at least four different species: *Enoplognatha ovata* (Reillo and Wise 1988), *E. latimana* (Oxford, 1991), *Theridion grallator* (Oxford and Gillespie 1996a,b) and *T. californicum* (Oxford 2005).

All of these species occur in very different areas. *E. ovata* has a wide distribution in Europe (Hippa and Oksala 1981; Oxford 1991), Asia and it was introduced to the northeast coast of the United States (Oxford and Reillo 1994). *E. latimana*, sister species of *E. ovata,* can be found living in sympatry with *E. ovata* on the Pembroke Peninsula (Southwest England) (Oxford, 1991). *Theridion grallator* (Hawaiian happy face spider) is endemic to the islands of O'ahu, Moloka'i, Maui and Big Island of Hawai'i (Croucher et al, 2012). Finally, *T. californicum* is distributed in the coastal forest of western North America (Levi, 1957; Oxford, 2009).

There are several characteristics that are shared across all of these examples of color polymorphism in the family Theridiidae, while other characteristics are exclusive to particular species. The first common element corresponds to the observation that the most common morph behaves as it is controlled by a recessive allele. In all species it corresponds to the "Yellow" morph (Reillo and Wise, 1988; Oxford, 1991; Oxford and Gillespie 1996a). There are also other two morphs ("redimita": yellow with two dorsolateral red lines; and "ovata": solid red), which are less common and also present in all these species. The degrees of dominance appear to be directly related with the amount of red pigmentation present in the phenotype (Oxford, 1983; Reillo and Wise, 1988 and Oxford and Gillespie 1996a). *T, grallator* (Oxford and Gillespie, 2001) and *T,*

*californicum* (Oxford 2009) present a total of 20 and 11 morphs, respectively, while the species in *Enoplognatha* have only the three previously described.

In all the species the inheritance of this trait is Mendelian and is associated with somatic chromosomes (Reillo, 1989; Oxford, 1989; Oxford and Gillespie 1996a), except on the Big Island population of *T. grallator* (Oxford and Gillespie 1996b). On the Big Island some of the morphs are restricted to males ("Red front" and "Red ring") and others to females ("Yellow" and "Red blob"). The inheritance of these sex-linked morphs has two loci instead of one (Oxford and Gillespie 1996b). Experimental crosses between Big Island and Maui specimens show that: (1) sex-linked morphs from Big Island behave identical as they do in pure Big Island crosses, (2) there is no effect of mixed genetic background in the non-sex-linked morphs, (3) the "Red front" morph (males) from the Big Island is expressed only in the presence of X chromosome from the Big Island, (4) the dominance hierarchies are maintained on mixed genetic backgrounds and (5) the over dominance of the white morph is less strong on mixed genetic backgrounds (Oxford and Gillespie 1996c). All of these suggest a more complicated genetic mechanism for color determination in this population.

The frequencies of the different morphs within each species tend to be very consistent (Reillo, 1989; Gillespie and Tabashnik, 1990) even if there is low connectivity between populations (Gillespie and Oxford, 1998; Croucher et al., 2012). Study of differentiation in nuclear markers has shown that the polymorphism is maintained by natural selection (Croucher at el., 2011a). In the case of *T. grallator* it has been suggested that balancing selection due to bird predation is responsible (Gillespie and Oxford, 1998). In particular, the Hawaiian Honey Creepers could develop a search image and any deviation from it will favor a reduction in predation (Gillespie and Tabashnik, 1990). For *E. ovata,* Oxford (2005) proposed a genetic model based on empirical data. He suggested that under certain range of allele frequencies genetic drift plays a major role on the maintenance of this polymorphism, but outside that range natural selection becomes important (Oxford, 2005). Also, the number of black spots has been postulated to a selective regime associated with thermal regulation at juvenile stages (Oxford, 1989; Oxford and Gummarsson, 2006). An aspect of the color polymorphism that has received less attention in terms of a comparative framework across species corresponds to the variation on the cephalothorax. It has been documented only in *T. grallator* (Oxford and Gillespie 1996b; Oxford and Gillespie 1996c).

In 2011, another example of color polymorphism was recognized in the Teridiidae in the endemic species, *Selkirkiella albogutata*, from Robinson Crusoe Island in the Juan Fernández archipelago of Chile (Fig. 1). Some of the recorded morphs were identical to the Hawaiian species (Fig. 2) while others were unique. Here we describe the color variation of *S. alboguttata* and place the phenomenon in a phylogenetic context to determine whether the color polymorphism of *S. albogutatta* represents a case for independent evolution and convergence, or whether there is common ancestry with another polymorphic species. To address this question we also examined color morphs in collections of related species from the continent (Valdivian temperate rainforest). In particular, we ask: (1) Which morphs are present in *S. alboguttata* and the continental species?, (2) What are the relative frequencies of the morphs and are they sex linked?, and (3) make a preliminary assessment of how wide spread the polymorphic condition is in the genus *Selkirkiella*.

A morphological phylogeny of the family placed the genus *Selkirkiella* as a sister to *Enoplognatha* (Agnarsson, 2004), while molecular analysis places it as a sister to Latrodectines (Arnedo et al. 2004); both a relatively basal placement on the family tree. *Selkirkiella alboguttata* belongs to a genus of 8 species distributed in the temperate rainforest of southern Chile and Argentina including the Falkland Islands (Agnarsson, 2004; World Spider Catalog, 2014). We expect to see on the close related species from the continent a similar degree of color variation (morphs, frequencies and sex linked morphologies) based on the original descriptions of these species (initially under the genus *Anelosimu*s). Levi (1963) stated that, "No two specimens of Chilean species have the same coloring".

Our study also addresses understanding the repeated evolution of complex traits at different taxonomic scales (Oxford, 2009). Having a second genus within the Theridiidae that has color polymorphism will allow the determination of which elements of this trait are convergent across the family and which ones conservered. At the same time, if the continental species also show similar polymorphisms it will be possible to address questions associated to the evolution of the polymorphism in recent divergences.

## MATERIALS AND METHODS

### Background on the genus *Selkirkiella*

Lucien Berland (1924) originally proposed the genus *Selkirkiella* based on a type species from the Juan Fernández archipelago. Regarding abdominal color polymorphism the original description of *S. alboguttata* states, "This pigmentation is somewhat variable, the bands can be fragmented and reduced to rows of spots, spots can happen to disappear, but the type of drawing is always found" (Berland. 1924). There is also a note about the similarity of this species to the specimens from Valdivia in Simon's collection. All other species (*S. carelmapuensis* (Levi, 1963), *S. luisi* (Levi, 1967), *S. magallanes* (Levi, 1963), *S. michaelseni* (Simon, 1902), *S. purpurea* (Nicolet, 1849), *S. ventrosa* (Nicolet, 1849), *S. wellingtoni* (Levi, 1967)), were initially described as *Anelosimus*, *Theridion* or subsequently transferred amongst those genera. The relocation of all these species into the *Selkirkiella* genus was done by Agnarsson (2004) following Arnedo et al. (2004).

There are three species with documented polymorphism. The description of *S. magallanes* states: "Abdomen variable in color with considerable amounts of white or black pigment. No two specimens alike" (Levi, 1963). Similarly, Nicolet (1849) mentioned the existence of several variants for *S. ventrosa*. Its coloration was described as: "Dark yellow abdomen mixed with white, black, red and dark brown". In Levi's (1976) publication there is a note saying that this particular specimen in Nicolet's collection is a misidentification of *E. ovata*, which has not been introduced to Chile. Perhaps this specimen that was accidently mixed with the Chilean collection or this species is extremely similar and was misidentified as *E. ovata*.

The best-documented species in terms of color polymorphism is *S. purpurea* (Nicolet, 1849). The original description explicitly states that there are 5 variants: "(1) Red abdomen with black spots and black in the lateral sides of the ventral area, (2) Dark on dorsal area of the abdomen with small red lateral markings with two white spots on the base of the markings, (3) Red abdomen with white spots and lateral markings almost

absent, (4) Dark abdomen with lateral markings that are white in the anterior area and red in posterior and (5) smaller in size, dark brown with white and red spots in the dorsal area" (Translated from Nicolet, 1849). Descriptions of the other four species in the genus do not include any mention of polymorphism, but it is not possible to rule out the possibility that the additional variants have not been recorded yet.

The descriptions for abdominal coloration in *S. carelmapuensis* mentions: "Dorsum of abdomen white with series of spots which contain some dark pigment" (Levi, 1963) and "yellow with grey" (Levi, 1967). In *S. luisi* the description says: "Abdomen yellowish with indistinct white dorsal spots continuing, in many specimens, as two longitudinal white lines; sometimes with four lines. Posterior portion of lines sometimes containing dark spots" (misidentified as *Anelosimus albolineatus* in Levi, 1967; World Spider Catalog, 2014). The description of *S. wellingtoni* also has no mention of color polymorphism. However, it does refer to a "purplish brown pattern of large spots" in the abdomen (Levi, 1967). This color is not mentioned in any other description and it could be a result of alcohol preservation. Finally, the description in Simon (1902) of *S. michaelseni* gives only a general idea of the abdominal coloration: "Protruding abdomen brief, shining, sparingly white-hair, albedo-gray, white-opaque line on both edges". A later description by Levi (1963) indicates the presence of two white longitudinal bands on the dorsal side of the abdomen and the occasional presence of black patches within the white area. Though inconclusive, this might also be an indication of polymorphism.

**Collecting sites and specimens**

The material analyzed in this study comes from 2 field collections on Robinson Crusoe Island, one museum collection (California Academy of Sciences) from the Valdivia area (Chile) and a field collection from Berkeley, California (USA). The author made the first field collection on Robinson Crusoe Island on August 2011 (14 individuals), while a second collection (48 individuals) was made by Gustavo Hormiga and Miquel Arnedo in April 2012. The spiders were collected during the day and night using hand collecting and a beating sheet. The specimens collected in 2011 were photographed before preservation. All the samples were preserved in 95% ethanol at -20°C for genomic work.

A total of 8 jars of unsorted arachnid material from the California Academy of Sciences collection came from tree fogging in the Valdivia area of Chile (PI Elizabeth Arias, February 2008) was examined. From this collection a total of 80 *Selkirkiella* specimens were identified. Lastly, specimens of *T. californicum* were collected in the hills of Berkeley, California (USA). These specimens were also preserved for genomic work. The exact collecting sites for all these samples are given in Table 1.

The morphs of the different specimens of *Selkirkiella* were scored using a classification from previous studies of color polymorphism in theridiid spiders (Reillo and Wise 1988; Oxford and Gillespie 1996a; Oxford and Gillespie 1996b; Oxford, 2005).

**DNA sequences**

DNA was extracted form 4 legs from the right side of each the spider using the QIAGEN DNA easy kit. The exact localities for each specimen are given in Table 1.

Mitochondrial [Cytochrome oxidase subunit I (COI) and 16S)] and nuclear [(28S, 18S and Histone 3 (H3)] genes were amplified in order to include these samples with a previously published data set (Arnedo et al. 2007). We sequenced 2 individuals of each of the following species: *S. alboguttata, S. luisi* and *T. californicum* (Table 2).

The master mix for all the PCR reactions consisted in: 1 µl of each primer at 10 µM, 2 µl of AmpliTaq buffer 10x, 0,5 µl of $MgCl_2$ 25 mM, 11.7 µl $H_2O$, 1 µl BSA 1x, 0.2 µl DreamTag and 1 µl of DNA extraction. The amplification profile for the gene COI (HCO:LCO; Folmer et al, 1994) started with 2 minutes at 95°C, followed by 35 repetitions of a cycle which started with 30 seconds at 95°C, then 45 seconds at 42°C and finally one minute at 72°C; there was an extra step of extension at 72°C for ten minutes. This is the same profile used for all the other markers, the only elements that differed were the annealing temperature and the number of cycles.

For the amplification of 16S we used the primers LR-N-13398 and LR-J-12864 (Simon et al., 1994). Annealing temperature was 47°C and we used 25 cycles. The primers used for the amplification of 18S corresponded to 18S-5F and 18S-9R (Giribet et al., 1999) with an annealing temperature of 59°C and 40 repetitions. For 28S we used the primers 28SA and 28SB (Whiting et al., 1997) with an annealing temperature of 48°C and 40 cycles. Finally, for the amplification of Histone 3 the primers H3F and H3R were used (Colgan et al., 1998). The annealing temperature corresponded to 48°C and we did 40 repetitions.

The PCR products were verified on an agarose/TBE gel. PCR products were cleaned using Axygen AxyPrep MagPCR Clean-up kit. DNA was sequenced directly in both directions through the cycle sequencing method using dye terminators (Sanger et al., 1977) using Life Technologies' BigDye Terminator v3.1 Cycle Sequencing Kit. Sequenced products were cleaned using Axygen AxuPrep Mag DyeClean kit and run out on ABI 3730XL DNA Analyzer automated sequencer. The sequences were edited, concatenated and aligned with Geneious Pro 5.6.7 (Biomatters). The sequence alignment used the algorithm MAFFT (Katoh et al., 2002) with the default parameters.

**Phylogenetic analysis**

Model parameters for estimating molecular evolution were determined using PARTITIONFINDER (Lanfear et al., 2012). The concatenated sequence (2,360 bp) had 8 partitions, including codon positions. The estimated models were: H3 position 1: SYM+G; 18S position 2, 28S position 2, COI position 3 and H3 position 2: K80+I+G; H3 position 3: JC; 16S position 1, 16S position 2 and 16S position 3: GTR+I+G; 18S position 1 and 18S position 3: K80+G; 28S position 1 and 28S position 3: K80+I+G; COI position 1: HKY+I+G; and COI position 2: GTR+I+G. *Synotaxus sp.* (Synotaxidae) was used as an outgroup following Arnedo et al., (2004). For comparison purposes we used the sequences present in Arnedo et al. (2007), but we also included sequences of *Enoplognatha caricis* and two species of *Anelosimos* (Arnedo et al., 2004). This would help us to determine where the Juan Fernández Islands and the Valdivian forest species reside relative to other polymorphic taxa in the tree.

The phylogenetic reconstruction was performed with Mr. Bayes 3.2.2 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003) available from the remote server CIPRES Science Gateway version 3.3 (Miller et al., 2012). The

phylogenetic reconstruction utilized 2 runs of 4 independent chains each. It was run for 35,000,000 generations with a sampling frequency of 1,000. The burn-in was set at 25%. Parameter convergence was determinate by checking the potential scale reduction factor (PSRF=1.0), standard deviation of split frequencies (<0.01) and with estimation of Effective Sample Size (ESS>200) combining both runs with TRACER version 1.7.5. (Rambaut et al., 2014). The phylogenetic tree was visualized with FigTree version 1.4.0 (http://tree.bio.ed.ac.uk/software/figtree/).

## RESULTS

### Species identification from California Academy of Sciences collection

The museum specimens from the tree fogging collections were all identified as *S. luisi* based on the morphological key of Levi (1967).

### Color polymorphism

Both *Selkirkiella* species showed color polymorphism. In the case of *S. albogutatta* there were 7 variants identified from a total of 63 specimens from different 5 localities on the island of Robinson Crusoe (Table 1). The description of the morphs is presented in Fig 3. The guanine layer can be seen underneath all the markings except for the black spots. There is also a black bar on the cephalothorax, which is more pronounced in some morphs than others.

The frequencies of the different morphs are summarized in Table 3. All of the different color polymorphisms are found amongst the males, while two are absent from females and juveniles ("Red front" and "Belt Red front"). In males the two most common morphs are "Yellow" and "Arch" representing 28.6% of the population each. "Yellow" occurs on 50% of the females, while "Belt" is the second most abundant (25%). In our analysis we did not distinguish between juvenile males and females. Most juveniles were members of the "Yellow" morph (62.1%), followed by "Arch" (10.3%). In the aggregation for the whole species, "Yellow" is the most common morph (48.4%) followed by "Two spots" (24.2%).

In *S. luisi*, the samples were preserved in 95% ethanol at -20°C, so it is very likely that some of the original pigments were degraded. There are 2 morphs from a total of 80 specimens collected in three different locations (Table 1). One of the locations, Fundo Paipahueño, corresponded to 3 different tree species (*Aextoxicon punctatum*. *Podocarpus nubigenus* and *Myrceugenia planipes*). The variants recorded were: "Yellow" and "Red bands" (for descriptions see Fig. 4) In males, "Yellow" represented an overwhelming majority (90.5%), while for females it was 67.7%. In juveniles 75% correspond to "Yellow" and 25% to "Red bands" (Table 4). In both species sample size was too small to make robust statistical conclusions regarding spatial variation of the polymorphism, and in the case of *S. alboguttata* there was also more extensive sampling effort in Plazoleta El Yunque.

## Phylogenetic reconstruction

After completion of both runs, all parameters had a potential scale reduction factor equal to 1.0 and the Effective Sample Size for the combined runs was greater than 200. The standard deviation of split frequencies between runs was 0.006295. A value of <0.01 is consider as "very good" (Ronquist et al., 2011).

In the tree (Fig. 5) there is a clade consisting on *E. carcis* and the two species of *Selkirkiella*. The *Selkirkiella* spp. are monophyletic and the island species have a relatively long branch perhaps due to their isolation. Most of the species from the ingroup form a polytomy, and there is insufficient resolution to determine relative relationships between some of the species and genera. All of these species correspond to the clade Lost Colular Setae (LCS) described in Arnedo et al., (2004). It includes species from the genera *Wamba, Keijia, Theridion, Thymoites, Neottiura, Coleosoma, Achaearanea, Takayus, Meotipa, Rugathodes, Ameridion* and *Simitidon*. The species from the genus *Theridion* are clustered on three main clades with some degree of geographic grouping (World Spider Catalog, 2014). The first one includes species mostly from North America: *T. pictum* (Holartic), *T. glaucescens* (USA and Canada), *T. differens* (USA and Canada), *T. flavonotatum* (USA and Cuba), *T. murarium* (North America) and *T. melanosticum* (Mediterranean, Aldabra, Seychelles, China, Japan, Polynesia, USA, Canada, Hispaniola). The second group includes different species across the whole American continent: *T. pictipes* (USA), *T. albidum* (USA and Canada), *T. cf. frondeum* (USA and Bahama Islands), *T. nigroannulatum* (Ecuador and Perú), *T. calcynatum* (Venezuela to Argentina) and *T. longipedatum* (Colombia). The third cluster has three Hawaiian species (*T. grallator, T. kauaiense* and *T. posticatum*), all of them sister to *Exalbidion pallisterorum*. This last species was described originally as *Theridion* and lives in México (World Spider Catalog, 2014). The two specimens sequenced of *T. californicum* are linked directly to the polytomy on the base. Finally, *T. melanum* appears sister to two species of *Keijia* and *T. acutitarse* to other two species of *Rugathodes*.

Nested within this polytomy there is a clade that includes two big groups. The first includes the species of *Chrysso*, *Theridula* and *Helvibis*. Then, there is another with *Anelosimus, Ariamnes, Selkirkiella* and *Enoplognatha*. The support values of these nodes are in general high.

# DISCUSSION

## Another case of convergence in color polymorphism

This paper is the first attempt to characterize the color polymorphism present in an endemic genus, *Selkirkiella* (Theridiidae), from the temperate rainforest of South America. In particular we described the color variation and phylogenetically assign two species within this group.

The first species, *S. alboguttata*, is endemic to Robinson Crusoe Island (Juan Fernández archipelago, Chile). It presents a total of 6 variations with "Yellow" the most common. As in the other well-studied polymorphic species (Reillo and Wise, 1988; Oxford, 1991; Oxford and Gillespie, 1996a,b and Oxford, 2005) most of the morphs are present in both genders. There are only two morphs that are absent from females and

juveniles. This can be interpreted in two ways: (1) that pigmentation is deposited late in development or (2) no females with that morphology were collected. The possibility that pigmentation is deposited later in development has been shown for different European populations of *E. ovata* (Hippa and Oksala, 1979; Oxford, 1983; Oxford and Gillespie 1996b). While, the second option is also likely as these three morphs are only represented by one specimen, which suggests that they could correspond to a very rare phenotype.

The other species that we studied was *S. luisi* from the Valdivian temperate rainforest. It has two moprhs ("Yellow" and "Red bands"), both present in males and females. In this species "Yellow" morph (75%) is quite common, which could explain why previous descriptions make no mention of the polymorphic condition (Levi, 1967).

While five polymorphic species are now known, there remain three species without reported color polymorphism (*S. carelmapuensis, S. wellingtoni* and *S. michaelseni*). It is possible that the apparent lack of polymorphism in these species corresponds to individual researcher bias. Three of the five species [*S. alboguttata* (Berland, 1924), *S. carelmapuensis* (Levi, 1963), *S. luisi* (Levi, 1967), *S. michaelseni* (Simon, 1902), *S. wellingtoni* (Levi, 1967)] with non-reported polymorphism were described by Levi, while two of the three (*S. magallanes* (Levi, 1963), *S. purpurea* (Nicolet, 1849), *S. ventrosa* (Nicolet, 1849)) with reported polymorphism, were described by Nicolet. Note Levi (1963) did describe one species with polymorphism (*S. magallanes*).

The description of *S. carelmapuensis* and *S. michaelseni* appears to represent a yellow morph, and it is possible as in *S. luisi*, the type specimen corresponds to the most common morph. Also it is probable that there is an underestimation of the morphs with red coloration due to degradation of pigments after their preservation (Oxford, 2009).

There are several factors that support calling this pattern of polymorphism an evolutionary convergence. The first and more essential is the polymorphic condition. In all cases there is more than one variant of abdominal coloration. As seen in the other well-studied species the "Yellow" morph is always present. In particular, the black spots present on this morph of *S. alboguttata* are more similar to *T. grallator* or *E. ovata*, than to *T. califronicum*. The variants "Arch" and "Red front" of *S. alboguttata* are also present in *T. grallator*. Indeed, "Red front" corresponds to the one that gives the common name to the species (Hawaiian Happy face spider). The morph "Red blob" present in *T. grallator*, *T. californicum*, *E. ovata* and *E. latimana* has not been reported for any *Selkirkiella*. It is important to mention, that this morph is a rare variant and the overall sampling of *Selkirkiella* is far from complete. On the other hand, the morph "Two red lines" has also not been reported in *S. albogutatta*. However, in *S. luisi,* the morph "Red bands" is similar to "Two red lines", which has been described on both *Theridion* and *E. ovata* (Gillespie and Tabashnik, 1989; Oxford, 2009). Note that on *S. luisi* the bands/stripes are shorter.

In comparison with other polymorphic *Selkirkiella*, it is possible that one of the variants of *S. magallanes* corresponds to the "Belt" morph of *S. alboguttata*. For *S. magallanes* Levi (1967) describes the presence of a "broad transverse band", however he claimed it was "made up of white pigment spots" and further mentions that the colors are altered after a year in alcohol.

In the species previously studied all of the variants have been reported in both genders and appear to be coded by a single locus. The only exception is the Big Island

population of *T. grallator* where the inheritance of the trait seems to be controlled by two loci and some of the morphs are restricted to a particular sex. In all of these the "Yellow" condition is double recessive. Rearing experiments are necessary to identify the inheritance mode of this trait in *Selkirkiella.* It is also not possible to estimate the number of loci that control this character. However, the similarity of the abundance of the "Yellow" morph in both species of *Selkirkiella* and its presence in males and females suggests possible similarities of the genetic mechanism with the other species. In other families, such as *Phiale* (*P.tristis, P.mimica, P.croceand* and *P.ortrudae*) in Salticidae, color polymorphism are also inherited by simple Mendelian alleles. However, the polymorphic condition is present only on females (Galiano, 1981).

This general color pattern has been explained by a model, which treats the pattern as 12 cells that can or cannot express pigmentation (Oxford, 2009). These cells are related with the embryonic segments and can be recognized in the adults by the boundaries between patches of guanine. Unfortunately, this model is difficult to test on *Selkirkiella,* because of the lack of background deposition of guanine.

The consistency of the relative proportions of the different morphs between species and across isolated populations (Gillespie and Oxford, 1998) provides evidence for a polymorphism maintained by natural selection. For *T. grallator* and *T. californicum* it has been hypothesized that predatory pressures (birds) are responsible for this condition (Gillespie and Oxford, 1998). The *Selkirkiella* system provides a unique opportunity to test this hypothesis. This would be explained more extensively in the section "The "under leaf fauna" and color polymorphism at a community level" below.

In the case of *E. ovata* the selective pressures that maintain the polymorphism appear more complex. There are similar proportions of the different morphs across populations, but there is also a clinal tendency where more northern populations are over represented by the "Yellow" morph (Oxford, 1985; Reillo, 1989). However, this pattern has an exception in Finland (Reillo and Wise, 1988), where the "redimita" morph is more common (Hippa and Oksala 1979; Hippa and Oksala, 1981). The genus *Selkirkiella* has a latitudinal distribution from Los Ríos district on Chile to southern Patagonia and the Falkland Islands, corresponding to a range of more than 1,300 km projecting towards the Southern Pole. This could be considered as a mirror image of the geographic distribution of *E. ovata*, and provides an opportunity to test for latitudinal gradients in the relative morph frequencies in *Selkirkiella* species.

Another situation in which a comparative studied could be done corresponds to the impact of habitat fragmentation on the genetic structure of the populations of *T. californicum* in the San Francisco region of California (Croucher et al., 2011b). Such a study could look for the effects of population expansion from refugial patches into managed areas, and provided comparisons for understanding the demographic processes occurring in *Selkirkiella* species present in the fragmented temperate rainforest of southern Chile (Ketl, 2001).

**Phylogenetic reconstruction**

*Selkirkiella* species appear as a sister group to *E. carcis*, which is identical with previous morphological phylogenetic reconstructions (Agnarsson, 2004). However, in the molecular phylogeny presented by Arnedo et al. (2004), there is a specimen of

*Selkirkiella*, which is sister to the Latrodectines clade. These species were not included in this study. Future analyses should add species from more continental taxa as well as additional *Enoplognatha.* These taxa from the same region as *Selkirkiella* will may provide better information for the sister clade of this group. In Chile there is only one species described of *Enoplognatha, E. zapfei*, and its abdominal coloration is unlike the polymorphic species: "Abdomen gray with darker gray mottled pattern on dorsum" (Levi, 1962).

This reconstruction shows that *Selkirkiella* represents another independent origin of color polymorphism in Theridiidae. This new case appears closer to *E. ovata*/*E. latimana* (Europe), than with *T. grallator* (Hawai'i) or *T. californicum* (California to British Columbia).

Related to the phylogenetic position of the other polymorphic species, *T. grallator* and *T. californicum* are not sister species and actually belong to separate sub-clades that are related by a polytomy (Fig. 4). The close relative to *T. grallator* is the *T. kauaiense*/*T.posticatum* pair, and does not represent a polymorphic condition. For *T. californicum*, it is impossible to refer to its sister species, as it is part of a polytomy in the sub-family Theridiinae. However, at least based on the current data, it is unlikely that it will be a polymorphic species given fact that in the analysis there is no other taxon with this condition. The fact that the molecular analysis was congruent with the previous molecular (Arnedo et al 2004) and morphological (Agnarsson, 2004) studies supports the idea of an independent evolution of this color polymorphism.

**Similarities between *S. alboguttata* and *T. grallator***

Even though they are positioned in very different parts of the tree, *S. alboguttata* and *T. grallator* share many similarities in the color polymorphism and the environment where they live. First, there is a striking similarity between some of the morphs, in particular, "Yellow" and "Red front". For the shape of the dorsal design of these variants, *T. grallator* is more similar to *S. alboguttata* (Fig. 2), than to *T. californicum*. Second, both spiders live in forest with an important presence of trees from the Myrtaceae family. In the case of *T. grallator* it is the O'hia (*Metrosideros polymorpha*) and in *S. alboguttata* it is the Luma (*Myrceugenia fernandeziana*), and both species also occur in other plants. Finally, there are many floristic and faunistic elements shared between the Hawaiian Islands and the Juan Fernández archipelago. Skottsberg (1925) studied the comparative flora of both archipelagoes and concluded that all the shared elements correspond to old-Pacific taxa that had originated in the Antarctic continent before its glaciation, which explained their wide distribution across the South Pacific Ocean.

The fauna shared between these two islands groups also corresponds to clades with Pacific wide distribution. For example, in Hawai'i the planthoppers from the genus *Nesosydne* (Delphacidae) radiated into 82 species (Hasty, 2005) with a very high fidelity to their host plant (silversword alliance; Gillespie & Roderick, 2002). In the Juan Fernández there are two endemic species, *N. minos* and *N. calypso*, with host plants belong to the genera *Gunnera*, *Pernettya* and *Drimys* (Fennah, 1957). Another example is found in land snails from the families Succineidae and Achatinellidae. In Hawai'i , there are 42 endemic Succineidae species (Cowie, 1995), while in the Juan Fernández there are 6 (*Succinea fernandi, S. texta, S. pinguis, S. cumingi, S. semiglobosa* and *Omalonyx*

*gayana*) (Odner, 1922). The Achatinellinae snails have a total of 209 species in the Hawaiian Islands (Cowie, 1995), while in the Juan Fernández, there are two endemic genera, *Tornatellina* and *Fernandezia*, with 7 and 12 species, respectively. The dispersal of invertebrates between these two groups of islands is difficult to explain based on current wind patterns, however the Pink-footed Shearwater (*Puffinus creatopus),* which nests in the Juan Fernández Islands and Isla Mocha (southern Chile), has been recorded as a vagrant visitor in Hawai'i (Pyle and Pyle, 2009). So bird transport remains a possibility.

**Color polymorphism in other species in Theridiidae**

There are more species within the Theridiidae family that present some degree of polymorphism or some of their documented morphologies match variants present in the species with well-studied polymorphism (in particular the "Yellow" morph). Among the species with documented polymorphism are: *Theridion frondeum* (Bradley, 2012) and *Thwaitesia glabicauda* (Zhang, 2011). While among the species that have morphology which resemblances at least one of the variants from the polymorphic species are: *Chrysso foliata* (Chikuni, 1989), *Chrysso octomaculata* (Chen, 2001; Namkung, 2002), *Chrysso scintillans* (Yoshida, 2003), *Chryso spiniventris* (Chen, 2001), *Chrysso vesiculosa* (Yoshida, 2003), *Enoplognatha margarita* (Chikuni, 1989; Namkung, 2002), *Phycosoma martinae* (Sebastian and Peter, 2009) and *Theridion bimaculata (*Chikuni, 1989). There are also other records of spiders with similar polymorphism from Japan, Pohnpei and Fiji (Gillespie *pers. comm.*). Detailed descriptions of morphs, the relative frequencies and the microhabitats in which they live are unknown for most of these species. This information in a phylogenetic framework is essential to understand the evolutionary patterns of this trait. In particular, examining the similarities in convergence and the ecological conditions under which it evolves (Oxford, 2009).

**The "under leaf fauna" and color polymorphism at a community level**

All the spiders that have been studied for color polymorphism tend to live under leaves (Oxford & Gillespie 1998; Oxford, 2009, Croucher et al., 2011a). In many cases it has been proposed that the color polymorphism evolved as a way to avoid predation (Gillespie and Tabashnik, 1990). However, the spiders are not the only ones with this life style. There seems to exist a "under leaf fauna" in which yellow/green organisms have different kinds of red and black markings. For example in the Hawaiian rainforest it is possible to find the green morph of the *Tetragnatha* spider (Araneae: Tetragnathidae), *Nesiomiris* (Hemiptera: Miridae), Nabidae (Hemiptera), *Nesophrosyne* larvae (Hemiptera: Cicadellidae) and other insects with similar markings (Fig. 6). If the predators represent a selective pressure and there is apostatic selection associated with this color polymorphism, a consideration of the color morphs present in the whole community of the "under leaf fauna" might be more informative than studies of single taxon.

On Robinson Crusoe Island there are other species of Theridiid that live in sympatry with *S. alboguttata* (*Chrysso backstromi, Theridion anson, Kochiura attrita, K. rosea* and *Styposis camoteensis*), and it would be interesting to see if they are also

polymorphic and the frequency of the different morphs if present. Also, from the evolutionary ecology perspective, *S. alboguttata* provides an ideal system to study the ecological origins of color polymorphism. On Robinson Crusoe there are only two endemic birds, which are insectivorous: the Cachudito de Juan Fernández, (*Anairetes fernandezianus*) and the Picaflor de Juan Fernández (*Sephanoides fernandensis*). Moreover, there are no records of human induced land bird extinctions, thus studying the birds predatory behavior would provide insights about the selective pressure that could be helping to maintain this condition.

In the context of the whole family the color polymorphism present in *Selkirkiella* should be considered an example of convergence, but at the same time within the genus it is possible to study the changes of this trait, which have been inherited most likely via common ancestry. The ecological mechanisms behind this convergence could be also studied at the community level including other species of this "under leaf fauna."

## CONCLUSION

The genus *Selkirlkiella* present on the temperate rainforest of southern South America has several species with documented color polymorphism. Here we documented that *Selkirlkiella alboguttata* displays 6 morphs and *S. luisi* only two. The "Yellow" morph was the more common in both species. By using a molecular phylogenic analysis, we confirm that this genus is close related to *Enoplognatha*, where other polymorphic species have been also described.

The presence of color polymorphism in this genus appears to be an event of convergent evolution at the family level, while between species it is more likely to be due to common ancestry. Indeed, at least five of the eight described species support this pattern. In order to rigorously test this statement it is necessary to add to the analysis other species of *Selkirkiella* and *Enoplognatha,* especially *E. ovata*, *E. latimana* and the closest relatives from continental Chile (Latrodectines clade). There are also other species within the family that also exhibit evidence of color polymorphism. Their detailed study and inclusion on a phylogenetic framework will inform our understanding of the phenomenon at a larger taxonomic scale.

The evolution of color polymorphisms in spiders may be related to a more widespread phenomenon associated with the whole "under leaf community" (spiders and insects), and may be associated with avoidance of bird predation.

This study shows a new case of well-documented color polymorphism on this family and provides a more complete phylogenetic background for the study of repeated evolution in color polymorphism in spiders.

## ACKNOWLEDGEMENTS

We thank Aaron Ramirez, which invitation to join his expedition to the Robinson Crusoe Island (founded by the Tinker Gran -Center of Latinamerican Studies, UC Berkeley-, UC Berkeley Integrative Biology Summer Research Grant and the Graduate Research Fellowship Travel Grant -NSF-) lead to this research. We also thank Iván Leiva, director of the Parque Nacional Archipiélago Juan Fernández, the park rangers Ramón Schiller and Alfonso Andaur, and the guide Rosa María Schiller for their field

assistance. Moreover, we appreciate all the help provided by Javiera Meza (CONAF V Región) and Lynne Hollyer (UC Berkeley Industry Alliances Office) in obtaining the permits. We also acknowledge the specimens collected by Gustavo Hormiga (The George Washington University) and Miquel Arnedo (Universidad de Barcelona) in Robinson Crusoe Island, as well as the ones collected by Elizabeth Arias in Valdivia. The access to this latter collection was facilitated by curator Charles Griswold (California Academy of Sciences). Finally, we would like to thank Peter Croucher for his collectingds of *T. californicum* and help in the lab work, and to Andrés Muñoz for a discussion about the plant community associated with this spiders. Darko Cotoras was funded by a Fulbright/CONICYT fellowship.

## LITERATURE CITED

Agnarsson I. 2004. Morphological phylogeny of cobweb spiders and their relatives (Araneae, Araneoidea, Theridiidae). *Zoological Journal of the Linnean Society* 141: 447–626.

Arnedo M.A., Agnarsson I. and Gillespie R.G. 2007. Molecular insights into the phylogenetic structure of the spider genus *Theridion* (Araneae, Theridiidae) and the origin of the Hawaiian *Theridion*-like fauna. *Zoologica Scripta* 36: 337–352.

Arnedo M.A., Coddington J., Agnarsson I. and Gillespie R.G. 2004. From a comb to a tree: phylogenetic relationships of the comb-footed spiders (Araneae, Theridiidae) inferred from nuclear and mitochondrial genes. Molecular Phylogenetics and Evolution 31: 225–245

Berland L. 1924. Araignées de l'ile de Pàques et des iles Juan Fernandez. In: The Natural History of Juan Fernandez and Easter Island. Vol. 3(3), 419-437.

Bradley R.A. 2012. Common Spiders of North America. University of California Press. 288pp.

Cain A.J. and Currey J.D. 1963. Area effects in *Cepaea*. Philosophical Transactions of the Royal Society London B 246: 1–81

Chen S. 2001. A guide to common spiders of Taiwan (Ebook). China Scientific Book Services. 476 pp.

Chikuni Y. 1989. Pictorial Encyclopedia of Spiders in Japan. Kaisei-sha Publishing Co., Tokyo, 310 pp.

Colgan,D.J., McLauchlan A., Wilson G.D.F., Livingston S.P., Edgecombe G.D., Macaranas J., Cassis G. and Gray M.R., 1998. Histone H3 and U2 snRNA DNA sequences and arthropod molecular evolution. Aust. J. Zool. 46, 419–437

Cowie, R.H., 1995. Variation in species diversity and shell shape in Hawaiian land snails: In situ speciation and ecological relationships. Evolution 49, 1191–1202.

Croucher P.J.P, Oxford G.S., Lam A. and Gillespie R.G. 2011a. Stabilizing selection maintains exuberant colour polymorphism in the spider *Theridion californicum* (Araneae, Theridiidae). Molecular Ecology 20: 206–218

Croucher P.J.P., Oxford G.S. and Gillespie R.G. 2011b. Population structure and dispersal in a patchy landscape: nuclear and mitochondrial markers reveal area effects in the spider *Theridion californicum* (Araneae: Theridiidae). *Biological Journal of the Linnean Society* 104: 600–620

Croucher P.J.P., Oxford G.S., Lam A., Mody N., and Gillespie R.G. 2012. Colonization

history and population genetics of the color-polymorphic Hawaiian happy-face spider *Theridion grallator* (Araneae, Theridiidae). Evolution 66: 2815-2833
Drummond AJ, Ashton B, Buxton S, Cheung M, Cooper A, Duran C, Field M, Heled J, Kearse M, Markowitz S, Moir R, Stones-Havas S, Sturrock S, Thierer T, Wilson A (2012) Geneious v5.6, Available from http://www.geneious.com
Eddy T.D., Gardner J.P.A. and Pérez-Matus A. 2010. Applying Fishers' Ecological Knowledge to Construct Past and Future Lobster Stocks in the Juan Fernández Archipelago, Chile. *PLoS ONE* 5: e13670.
Fennah R.G. 1957. Los insectos de las Islas Juan Fernández. 29. Fulgoroidea (Homoptera: Fulgoroidea). Rev. Chil. Entomol. 5, 375-384.
Folmer O., Black M., Hoeh W., Lutz R. and Vrijenhoek R.1994. DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. Molecular Marine Biology and Biotechnology. 3: 294-299
Galiano M.F. 1981. Revisión del género *Phiale* C.L.Koch, 1946 (Araneae, Salticidae) I. Las especes polimórficas del grupo mimica. Arachnol. 9: 61—85.
Gillespie R.G. 2013. Adaptive Radiation: Convergence and Non-equilibrium. Current Biology. 23: R71–R74
Gillespie R.G. and Oxford G.S. 1998. Selection on the color polymorphism in Hawaiian Happy-face spiders: evidence from genetic structure and temporal fluctuations. Evolution 52: 775-783
Gillespie R.G. and Tabashnik B.E. 1990. Maintaining a happy face: stable colour polymorphism in the spider *Theridion grallator* (Araneae, Theridiidae). Heridity 65: 67-74
Gillespie R.G. and Tabashnik. 1989. What makes a happy face? Determinants of color pattern in the Hawaiian happy face spider Theridion grallator (Araneae, Theridiidae). Heredity 62: 355-363
Giribet G., Carranza S., Riutort M., Baguna J. and Ribera, C. 1999. Internal phylogeny of the Chilopoda (Myriapoda, Arthropoda) using complete 18S rDNA and partial 28S rDNA sequences. Philos. Trans. R. Soc. Lond. B 354: 215–222
Gray S.M. and McKinnon J.S. 2006. Linking color polymorphism maintenance and speciation. TRENDS in Ecology and Evolution 22: 71-79
Harmon L.J., Kolbe J.J., Cheverud J.M. and Losos J.B. 2005. Convergence and the multidimensional niche. Evolution 59: 409–421
Hasty G.L. 2005. Host-plant specialization affects species diversity and vulnerability to endangerment in a group of island-endemic phytophagous insects (Delphacidae, Hawaiian Islands). Tesis doctoral, University of California, Berkeley,
Hippa H. and Oksala I. 1981. Polymorphism and reproductive strategies of *Enoplognatha ovata* (Clerck) (Araneae, Theridiidae) in northern Europe. Annales Zoologici Fennici. 18: 179-190.
Huelsenbeck J.P. and Ronquist F. 2001. MRBAYES: Bayesian inference of phylogeny. Bioinformatics 17: 754-755
Hughes K.A., Houde A.E., Price A.C. and Rodd F.H. 2013. Mating advantage for rare males in wild guppy populations. Nature 503: 108–110
Kapan D.D, Flanagan N.S., Tobler A., Papa R., Reed R.D., Gonzalez J.A., Restrepo M.R., Martinez L., Maldonado K., Ritschoff C., Heckel D.G., and McMillan W.O.

2006. Localization of Müllerian mimicry genes on a dense linkage map of *Heliconius erato*. Genetics. 173: 735-757
Katoh K., Misawa K., Kuma K. and Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 30: 3059-3066
Kelt D.A. 2001. Differential effects of habitat fragmentation on birds and mammals in Valdivian temperate rainforests. Rev. chil. hist. nat. 74: 769-777
Kettlewell B. 1958. The importance of the micro-environment to evolutionary trends in the Lepidoptera. Entomologist 91: 214-224
Kocher T.D. 2004. Adaptive evolution and explosive speciation: the cichlid fish model. Nat. Rev. Genet. 5: 288-98.
Lanfear, R., Calcott, B., Ho, S.Y. & Guindon, S. 2012. Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Molecular Biology and Evolution. 29: 1695–1701
Levi H.W. 1957. The spider genera *Enoplognatha*, *Theridion* and *Paidisca* in America north of Mexico. Bull. Am. Mus. Nat. Hist. 112: 1-123.
Levi H.W. 1962. The Spider Genera *Steatoda* and *Enoplognatha* in America (Araneae, Theridiidae). Psyche 69: 11-36
Levi H.W. 1963. The American spiders of the Genus Anelosimus (Araneae: Theridiidae). Transactions of the American Microscopical Society. 82: 30-48
Levi H.W. 1967. The theridiid spider fauna of Chile. *Bulletin of the Museum of Comparative Zoology* 136: 1–20.
Manceau M., Domingues V.S., Linnen C.R., Rosenblum E.B. and Hoekstra H.E. 2010. Convergence in pigmentation at multiple levels: mutations, genes and function. *Phil. Trans. R. Soc. B* 365: 2439-2450.
Miller M.A., Pfeiffer W. and Schwartz T. 2012. "The CIPRES Science Gateway: Enabling High-impact Science For Phylogenetics Researchers With Limited Resources", XSEDE '12, Proceedings of the 1st Conference of the Extreme Science and Engineering Discovery Environment: Bridging from the eXtreme to the campus and beyond: pg: -, ACM, 978-1-4503-1602-
Namkung J. 2002. The spiders of Korea. Kyo-Hak Publishing Co., Seoul, 648 pp.
Nicolet A.C. 1849. Aracnidos. In: Gay, C. (ed.) Historia física y política de Chile. Zoología 3, 319-543
Nijhout H. F. 2003. *Polymorphic mimicry in Papilio dardanus: mosaic dominance, big effects and origins*. Evolution & Development. 5: 579-592.
Odhner N.H. 1922. Mollusca from Juan Fernandez and Easter Island. The Natural History of Juan Fernandez and Easter Island 3: 219-265
Oxford G. 1976. The colour polymorphism in *Enoplognatha ovatum* (Clerck) (Araneae:Theridiidae) .Heredity 36: 369-381
Oxford G.S. 1983. Genetics of colour and its regulation during development of the spider *Enoplognatha ovata* (Clerk) (Araneae: Theridiidae) Ibid 51: 621-634
Oxford G.S. 1985. A countrywide survey of colour morph frequencies in the spider *Enoplognatha ovata* (Glerck) (Araneae: Theridiidae): evidence for natural selection Biol. J. Linn. Soc. 24: 103-142

Oxford G.S. 1989. Genetics and distribution of black spotting in *Enoplognatha ovata* (Araneae: Theridiidae), and the role of intermittent drift in population differentiation. Biological Journal of the Linnean Society 36: 111-128.

Oxford G.S. 1991. Visible morph-frequency variation in allopatric and sympatric populations of two species of *Enoplognatha* (Araneae: Theridiidae). Heredity 67: 317-324.

Oxford G.S. 2005. Genetic drift within a protected polymorphism: enigmatic variation in color-morph frequencies in the candy-stripe spider, *Enoplognatha ovata*. Evolution. 59: 2170–2184.

Oxford G.S. 2009. An exuberant, undescribed colour polymorphism in *Theridion californicum* (Araneae, Theridiidae): implications for a theridiid pattern ground plan and the convergent evolution of visible morphs. Biological Journal of the Linnean Society. 96: 23-34.

Oxford G.S. and Gillespie R.G. 1996a. Genetics of a colour polymorphism in the Hawaiian happy face spider, *Theridion grallator* (Araneae: Theridiidae) from Greater Maui. Heredity 76: 238-248.

Oxford G.S. and Gillespie R.G. 1996b. Quantum shift in the genetic control of a colour polymorphism in *Theridion grallator* (Araneae: Theridiidae). The Hawaiian happy-face spider. Heredity 76: 249-256

Oxford G.S. and Gillespie R.G. 1996c. The effects of genetic background on the island-specific control of a colour polymorphism in Theridion grallator (Araneae: Theridiidae), the Hawaiian happy-face spider. Heredity 76: 257-266

Oxford G.S. and Gillespie R.G. 2001. Portraits of evolution: studies of coloration in Hawaiian spiders. BioScience 51: 521-528

Oxford G.S. and Gunnarsson B. 2006. Spatial variation in colour morph, spotting and allozyme frequencies in the candy-stripe spider, *Enoplognatha ovata* (Theridiidae) on two Swedish archipelagos. Genetica 128: 51–62

Oxford, G. S. and Gillespie R.G. 1996b. The effects of genetic background on the island-specific control of a colour polymorphism in the Hawaiian happy face spider, *Theridion grallator* (Araneae: Theridiidae). Heredity 76: 257-266.

Oxford, G.S. and Reillo P.R. 1994. The world distributions of species within the *Enoplognatha ovata* group (Araneae: Theridiidae): implications for their evolution and for previous research. Bull. Br. Arachnol. Soc. 27: 225-249.

Protas M.E. and Patel N.H. 2008. Evolution of Coloration Patterns. Annual Review of Cell and Developmental Biology. 24: 425-446.

Pyle, R.L., and P. Pyle. 2009. The birds of the Hawaiian Islands: occurrence, history, distribution, and status. B.P. Bishop Museum, Honolulu, HI, U.S.A. Version 1 (31 December 2009) http://hbs.bishopmuseum.org/birds/rlp-monograph/

Rambaut A, Suchard MA, Xie D & Drummond AJ (2014) Tracer v1.6, Available from http://beast.bio.ed.ac.uk/Tracer

Reillo P.R. 1989. Color Polymorphism in the Spider *Enoplognatha ovata* (Araneae: Theridiidae): Broad-scale Morph-frequency Variation in Northeastern North America American Midland Naturalist. 122: 199-203

Reillo P.R. 1989. Color polymorphism in the spider *Enoplognatha ovata* (Araneae: Theridiidae): Broad-scale morph-frequency variation in Northeastern North America. American Midland Naturalist. 122: 199-203

Reillo P.R. and Wise D.H. 1988. Genetics of Color Expression in the Spider *Enoplognatha ovata* (Araneae: Theridiidae) from Coastal Maine. American Midland Naturalist. 119: 318-326

Reillo P.R. and Wise D.H. 1988. Temporal and spatial patterns of morph-frequency variation among coastal Maine populations of the polymorphic spider *Enoplognatha ovata* (Araneae: Theridiidae). *American Midland Naturalist.* 120: 337-354.

Ronquist F. and Huelsenbeck J.P. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19: 1572-1574

Ronquist F., Huelsenbeck J. and Teslenko M. Draft MrBayes version 3.2 Manual: Tutorials and Model Summaries. November 15, 2011

Sanger F., Nicklen S. and Coulson A.R. 1977. DNA sequencing with chain-terminating inhibitors. Proceedings of the National Academy of Sciences USA 74: 5463-5467

Sebastian P.A. and Peter K.V. 2009. Spiders of India. Orient Blackswan. 734pp.

Seehausen O., Mayhew P.J. and Van Alphen J.J.M. 1999. Evolution of colour patterns in East African cichlid fish. Journal of Evolutionary Biology. 12: 514-534.

Simon C., Frati F., Beckenbach A., Crespi B., Liu H. and Flook P. 1994. Evolution weighting and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Ann Entomol Soc Am 87: 651–701

Simon E. 1902. Arachnoideen, exclu. Acariden und Gonyleptiden. In: Ergebnisse der Hamburger Magalhaensische Sammelreise. Hamburg 6(4), 1-47.

Skottsberg. 1925. Juan Fernandez and Hawaii: A phytographical discussion. Bernice P. Bishop Museum Bulletin 16

Stuardo J.R. and Vargas.Almonacid P. 2000. Moluscos terrestres de Chile. Sinonimia y problemas relacionados: 1. Familias Veronicellidae, Pipillidae y Achatinellidae (Gastrópoda: Pulmonata) *Gayana (Concepc.):* 64: 171-188

Vargas R., Reif A. and Faúndez M.J. 2011. The forests of Robinson Crusoe Island, Chile: an endemism hotspot in danger. Bosque 32: 155-164

Whiting M., Carpenter J., Wheeler Q. and Wheeler W. 1997. The Strepsiptera problem: phylogeny of the holometabolous insect orders inferred from 18S and 28S ribosomal DNA sequences and morphology. Syst. Biol. 46: 1–68

World Spider Catalog. 2014. World Spider Catalog. Natural History Museum Bern, online at http://wsc.nmbe.ch, version 15.5, accessed on September 13th, 2014.

Yoshida, H. 2003. The spider family Theridiidae (Arachnida: Araneae) from Japan. Arachnological Society of Japan, 224 pp.

Zhang Z. 2011. A Photographic Guide to Spiders of China. Chongqing University Pub. 175 pp.

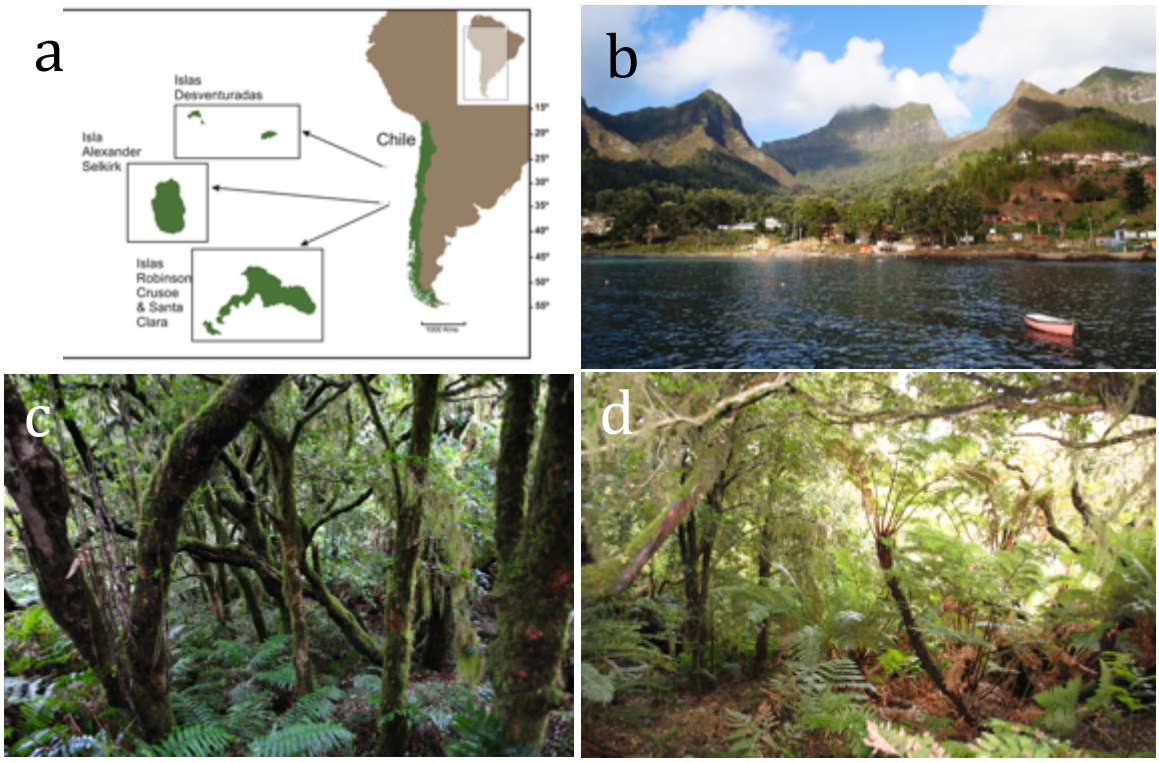



**Figure 1: Field site.** (a) Robinson Crusoe island (Juan Fernández Archipelago, Chile) is a volcanic island (approx. 4 MY) located 680 km from the coast of central Chile (image from Eddy et al., 2010). (b) Its only town, San Juan Bautista, has around 800 people. It has an area of 47.9 km² and its highest peak is Cerro El Yunque (915 meters). (c) The vegetation is mainly related to the Valdivian coastal formations of southern Chile. It is dominated by Luma (*Myrceugenia fernandeziana,* Myrtaceae) and Caleno (*Drimys confertifolia,* Winteraceae). Both, sister to continental species from the Valdivian forest. (d) It also shares characteristics with subtropical islands of the New Zealand region and the Hawaiian rainforest. The annual precipitation is 957 mm and the average temperature 15.2°C (Vargas et al., 2011).

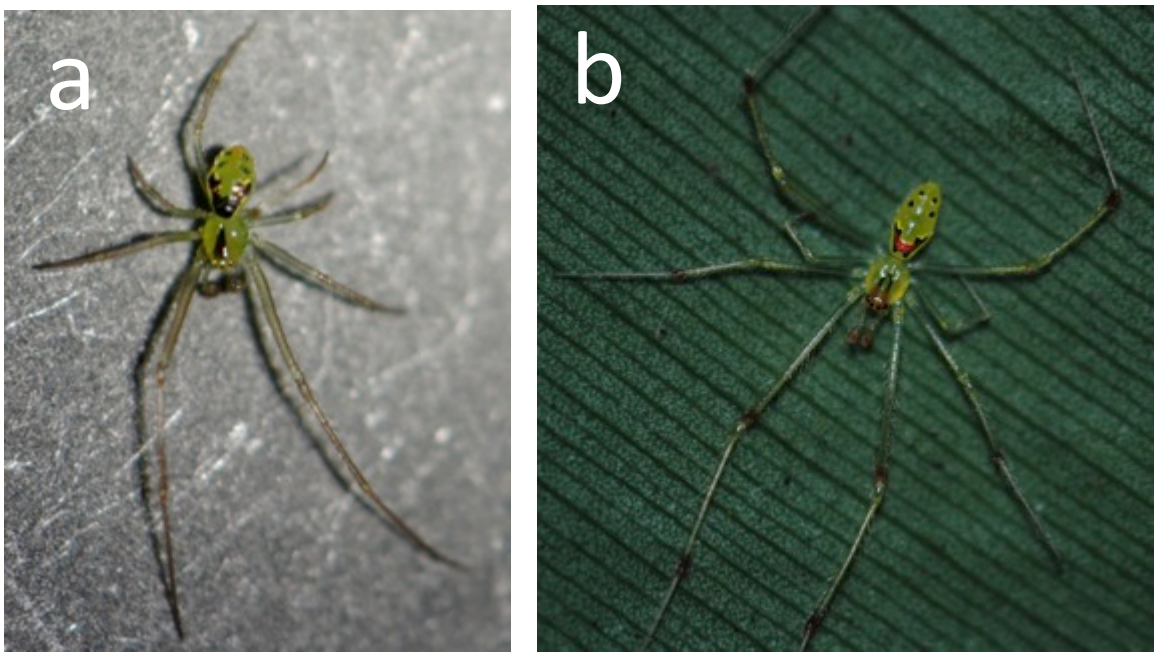



**Figure 2: Similarities between *S. alboguttata* and *T. grallator*.** Among the different variants of (a) *S. alboguttata*, one of them is strikingly similar to the "Red front" variant of (b) *T. grallator*. Indeed, this is the morph that gives the common name of "Hawaiian happy face spider" to *T. grallator*. Then, *S. alboguttata* could be also called "Araña de cara feliz de Robinson Crusoe" (from the spanish: "Happy Face Spider from Robinson Crusoe").

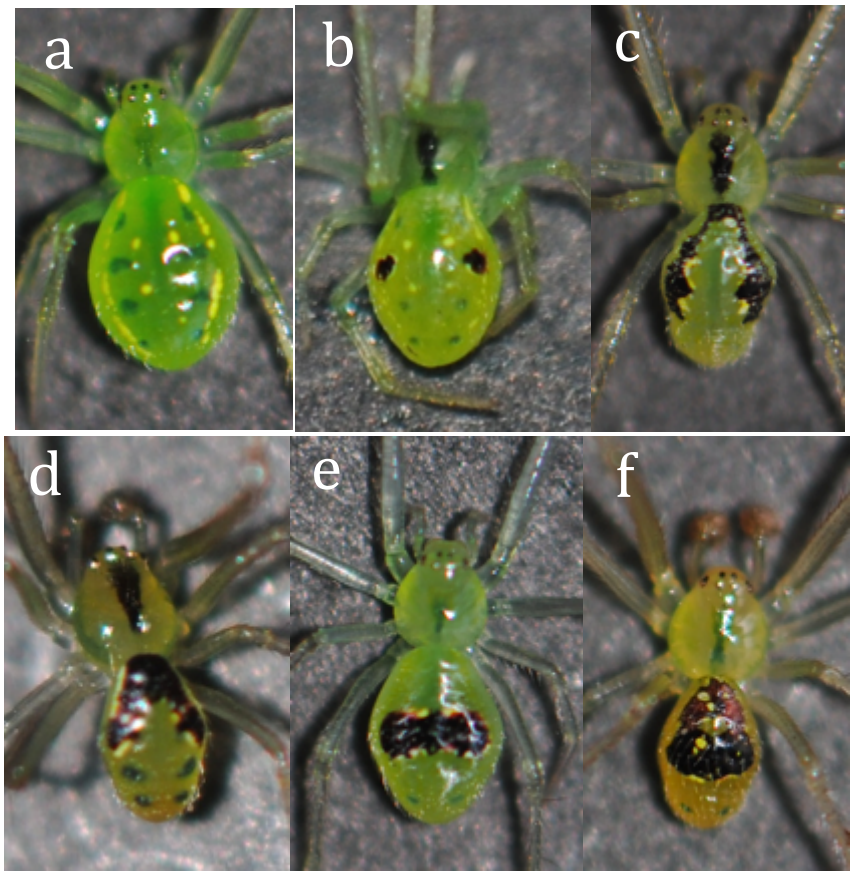



**Figure 3: Color polymorphism on *S. alboguttata*.** The species presents 7 different variants. They correspond to: (a) "Yellow", has four pairs of black spots on the sides; (b) "Two spots", the second pair of spots from anterior to posterior have enlarged concentrations of red pigment; (c) "Arch", it corresponds to a black U shape mark in the anterior part of the abdomen, there are no black spots; (d) "Red front", it has a U shape mark in the anterior part of the abdomen and two pairs of spots in the posterior part; (e) "Belt", an horizontal dark mark in the middle of the abdomen, there are also two very faint black spots in posterior; and (f) "Belt Red front" it looks like "Belt", but in the anterior part of the abdomen there is a red crescent. All the pictures where taking from live spider, except (g). The colors on this specimen could be altered by the preservation (95% ethanol in -20°C).

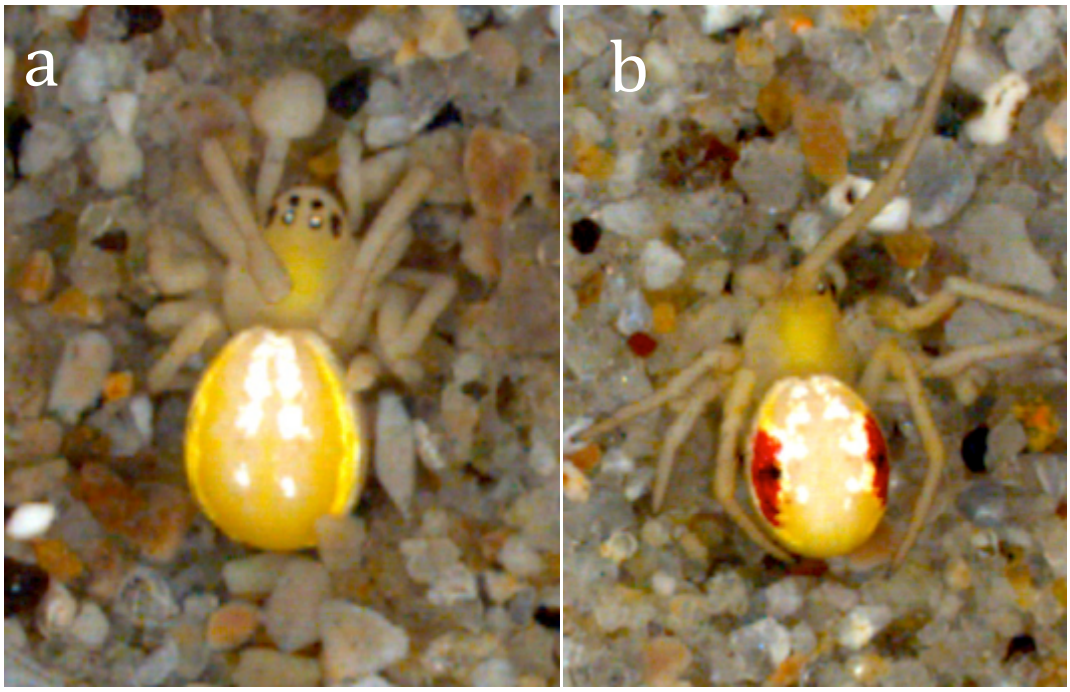



**Figure 4: Color polymorphism on *S. luisi*.** The species presents 2 different variants. They correspond to (a) "Yellow", which is very similar to the one in *S. alboguttata* (note there are two lateral bands with a different yellow coloration) and (b) "Red bands", which has two red markings on the sides that do not extend for the full extension of the abdomen. The pictures were taken from spider preserved on 95% ethanol in -20°C, so the original colors on this specimen could be altered.

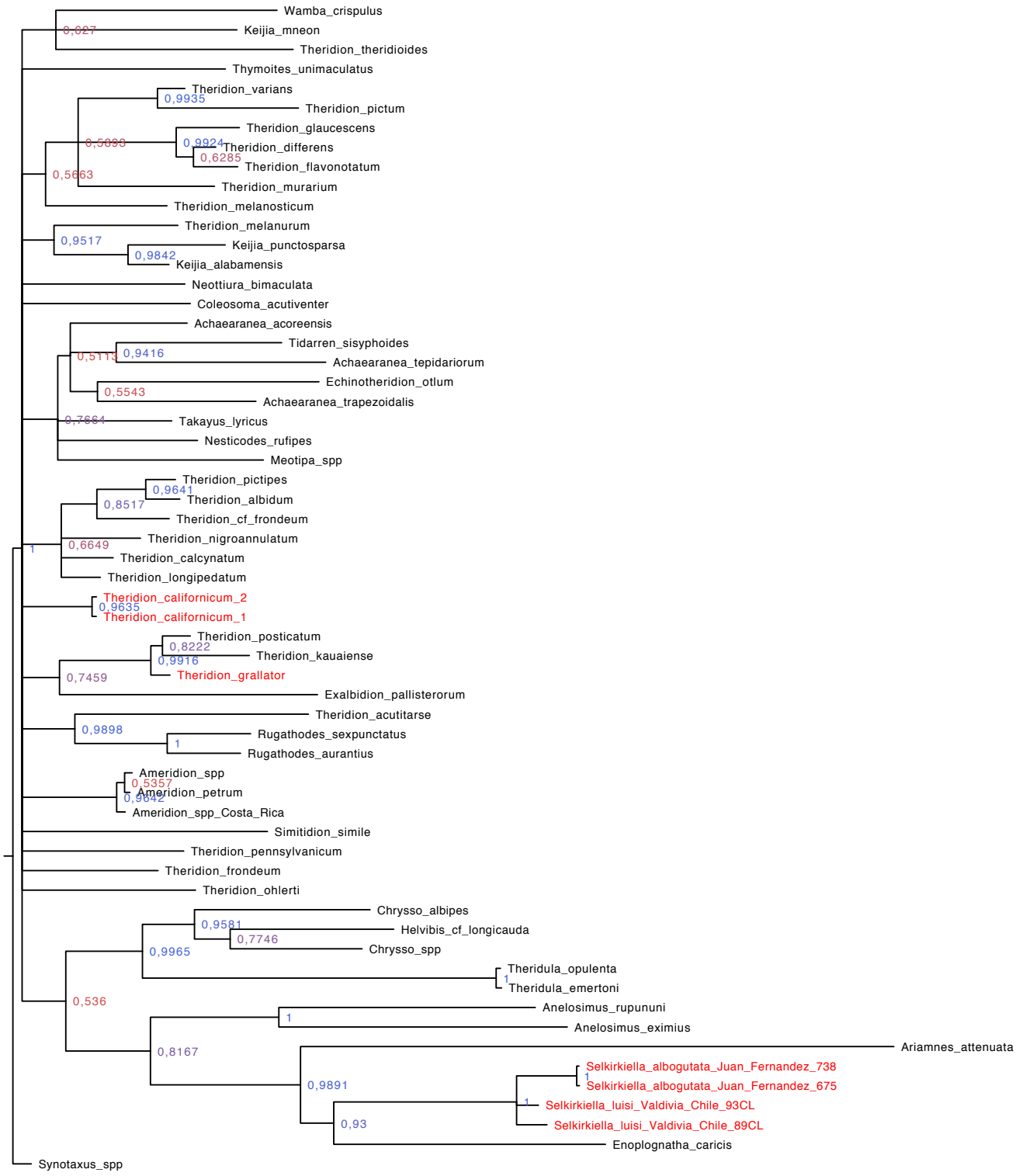



**Figure 5. Bayesian phylogenetic reconstruction.** The molecular phylogeny is based on five concatenated genes (COI, 16S, 18S, 28S and H3). The sequenced specimens (Table 2) were included to the specimens present in Arnedo et al., 2007 and some from Arnedo et al., 2004. Nodes with posterior probabilities <0.5 were collapsed into polytomies.

Posterior probabilities closer to 0.5 are in red and to 1 in blue. In red all the species names with described color polymorphism.


**Figure 6. The "under leaf fauna".** Different kinds of arthropods that live under the leaves share similar colors and body shapes. Some examples from Hawai'i: (a) *Tetragnatha brevignatha* (Araneae: Tetragnatidae), (b) *Nesiomiris* (Hemiptera: Miridae), Nabidae (Hemiptera) and *Nesophrosyne* larvae (Hemiptera: Cicadellidae). Note the similarity between *T. brevignatha* and *Nesiomiris*, as well as the colors of the *Nesophrosyne* larvae and the Hawaiian Happy face spider.

**Table 1: Samples utilized on this study**

| Species | Collecting site | Date | Number of specimens | Collector |
|---|---|---|---|---|
| *S. alboguttata* | Cerro Damajuana (S 33°38.984' W 78°50.645') | 08.26.2011 | 8 | Darko Cotoras |
| | Cordón Salsipuedes (S 33°37.955' W 78°50.236') | 08.24.2011 | 1 | |
| | Salto de la Pulga (S 33°38.201' W 78°51.201') | 08.21.2011 | 4 | |
| | Plazoleta El Yunque (S 33°38.984' W 78°50.645') | 08.24,27.2011 | 1 | |
| *S. alboguttata* | Plazoleta El Yunque (S 33°38.984' W 78°50.645') | 03.24.2012, 04.1,5,6,14.2012 | 41 | Gustavo Hormiga and Miquel Arnedo |
| | Valle Villagra (no GPS point) | 04.02.2012 | 7 | |
| *S. luisi* | Reserva Costera Valdiviana, Valdivia. Fogging of *Nothofagus nitida.* (S 39°49.687' W 73°35.227') | 02.26,27.2008 | 14 | Elizabeth Arias |
| | Fundo Paipahueño, Valdivia. Fogging of *Aextoxicon punctatum*. (S 39°42.806' W 73°21.390') | 02.24.2008 | 26 | |
| | Fundo Paipahueño, Valdivia. Fogging of *Podocarpus nubigenus*. (S 39°42.619' W 73°21.058') | 02.22,24.2008 | 27 | |
| | Fundo Paipahueño, Valdivia. Fogging of *Myrceugenia planipes*. (S 39°42.806' W 73°21.390') | 02.24.2008 | 9 | |
| | Fundo Manchao, Valdivia. Fogging of *Myrceugenia planipes*. (S 39°42.194' | 02.23.2008 | 4 | |

| | | | | |
|---|---|---|---|---|
| | W 73°21.418') | | | |
| *T. californicum* | Berkeley, California. USA (no GPS point) | 2011 | 2 | Peter Croucher |

All the samples from *S. alboguttata* were collected on the Robinson Crusoe Island (Juan Fernández Archipelago, Chile). The general location for the samples of *S. luisi* corresponds to Los Rios District, Chile.

**Table 2: Sequenced specimens**

| **Species** | **Locality** | **Collector** |
|---|---|---|
| *S. alboguttata (738)* | Cerro Damajuana (S 33°38.984' W 78°50.645') | Darko Cotoras |
| *S. alboguttata (675)* | Salto de la Pulga (S 33°38.201' W 78°51.201') | Darko Cotoras |
| *S. luisi (89CL)* | Fundo Paipahueño, Valdivia. Fogging of *Podocarpus nubigenus*. (S 39°42.619' W 73°21.058') | Elizabeth Arias |
| *S. luisi (93CL)* | Fundo Paipahueño, Valdivia. Fogging of *Myrceugenia planipes*. (S 39°42.806' W 73°21.390') | Elizabeth Arias |
| *T. californicum* | Berkeley, California. USA (no GPS point) | Peter Croucher |
| *T. californicum* | Berkeley, California. USA (no GPS point) | Peter Croucher |

**Table 3: Frequencies of color variants on *S. alboguttata***

| | Males | 21 | | | | |
|---|---|---|---|---|---|---|
| Location | Yellow | Arch | Belt | Two spots | Red front | Belt red front |
| Plazoleta El Yunque | 3 | 5 | 1 | 5 | 1 | 0 |
| Cerro Damajuana | 0 | 1 | 0 | 1 | 0 | 0 |
| Cordón Salsipuedes | 0 | 0 | 0 | 0 | 0 | 0 |
| Salto de la Pulga | 2 | 0 | 0 | 0 | 0 | 1 |
| Valle Villagra | 1 | 0 | 0 | 0 | 0 | 0 |
| TOTAL per category | 6 | 6 | 1 | 6 | 1 | 1 |
| PERCENTAGE per category | 0,2857 | 0,2857 | 0,0476 | 0,2857 | 0,0476 | 0,0476 |

| | Females | 12 | | | | |
|---|---|---|---|---|---|---|
| Location | Yellow | Arch | Belt | Two spots | Red front | Belt red front |
| Plazoleta El Yunque | 3 | 0 | 2 | 2 | 0 | 0 |
| Cerro Damajuana | 2 | 0 | 1 | 0 | 0 | 0 |
| Cordón Salsipuedes | 1 | 0 | 0 | 0 | 0 | 0 |
| Salto de la Pulga | 0 | 1 | 0 | 0 | 0 | 0 |
| Valle Villagra | 0 | 0 | 0 | 0 | 0 | 0 |
| TOTAL per category | 6 | 1 | 3 | 2 | 0 | 0 |
| PERCENTAGE per category | 0,5 | 0,0833 | 0,25 | 0,1666 | 0 | 0 |

| | Juveniles | 29 | | | | |
|---|---|---|---|---|---|---|
| Location | Yellow | Arch | Belt | Two spots | Red front | Belt red front |
| Plazoleta El Yunque | 11 | 2 | 1 | 6 | 0 | 0 |
| Cerro Damajuana | 2 | 0 | 0 | 1 | 0 | 0 |
| Cordón Salsipuedes | 0 | 0 | 0 | 0 | 0 | 0 |
| Salto de la Pulga | 0 | 0 | 0 | 0 | 0 | 0 |
| Valle Villagra | 5 | 1 | 0 | 0 | 0 | 0 |
| TOTAL per category | 18 | 3 | 1 | 7 | 0 | 0 |
| PERCENTAGE per category | 0,6206 | 0,1034 | 0,0344 | 0,2413 | 0 | 0 |

| | | | | | | |
|---|---|---|---|---|---|---|
| TOTAL (M+F+J) | 30 | 10 | 5 | 15 | 1 | 1 |
| PORCENTAGE (M+F+J) | 0,4838 | 0,1612 | 0,0806 | 0,2419 | 0,0161 | 0,0161 |

**Table 4: Frequencies of color variants on *S. luisi***

| | Male | 21 | Female | 31 | Juvenile | 28 |
|---|---|---|---|---|---|---|
| Location | Yellow | Red bands | Yellow | Red bands | Yellow | Red bands |
| Reserva Costera Valdiviana (*N. nitida*) | 4 | 1 | 3 | 1 | 2 | 3 |
| Fund. Paipahueño (*A. punctatum*) | 5 | 0 | 8 | 3 | 7 | 3 |
| Fund. Paipahueño (*P. nubigenus*) | 8 | 0 | 6 | 3 | 8 | 2 |
| Fund. Paipahueño (*M. planipes*) | 0 | 1 | 3 | 2 | 3 | 0 |
| Fund. Manchao (*M. planipes*) | 2 | 0 | 1 | 1 | 0 | 0 |
| | | | | | | |
| TOTAL per category | 19 | 2 | 21 | 10 | 20 | 8 |
| PERCENTAGE per category | 0,9047 | 0,0952 | 0,6774 | 0,3225 | 0,7142 | 0,2857 |

| | | |
|---|---|---|
| TOTAL (M+F+J) | 60 | 20 |
| PORCENTAGE (M+F+J) | 0,75 | 0,25 |