UNIVERSITY OF CALIFORNIA SAN DIEGO

High Throughput Sequencing of Ultraconserved Elements Reveals Inter-island Relationships of a California Channel Island Endemic Ant Species (*Aphaenogaster patruelis*)

A Thesis submitted in partial satisfaction of
the requirements for the degree Master of Science

in

Biology

by

Bo Huey Chiang

Committee in charge:

Professor Davie A. Holway, Chair
Professor Joshua R. Kohn
Professor Greg Rouse

2018

The Thesis of Bo Huey Chiang is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

Chair

University of California San Diego

2018

TABLE OF CONTENTS

LIST OF SUPPLEMENTARY FILES

Chiang_Library_Prep.pdf. This is the unpublished procedure by Marek Borowiec of UC Davis for pre-sequencing DNA library preparation that includes instructions for bead clean-up, blunt end repair, adapter ligation, and index PCR.

Chiang_Target_Capture_Enrichment.pdf. This is the procedure by Arbor Biosciences for capturing and enriching the UCE sequences in the prepared libraries.

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

Funding for this study was provided by the National Geographic Society, the US Navy, and UC San Diego. We thank the following people for their help: Marek Borowiec, Melissa Booker, Christie Boser, Michael Branstetter, Amanda Chisholm, Kate Faulkner, Aaron Hebshi, Bill Hoyer, David Mazurkiewicz, Korie Merrill, and Phil Ward. We acknowledge the Channel Islands National Park, Catalina Island Conservancy, US Navy and the Isla Guadalupe Pacific Islands Biosphere Reserve for logistical support.

Special thanks to Dr. Bernice DeMarco, Dr. Philip Ward, the collecting team on Isla Guadalupe and the Santa Catalina Fox Project for the generous donation and loaning of precious specimens; to Jessica Hammond, Jennifer Chiang, and Stephen Chiang for spending hours going through hundreds of UCE loci data one at a time looking and recording best models for partitioned Bayesian Inference; to my thesis committee chair and members Dr. David Holway, Dr. Joshua Kohn, and Dr. Greg Rouse for approving and giving me the opportunity to conduct this research; to PhD candidate Ida Naughton for specimen collection, making all the materials available, and proofreading and editing the first draft; to Jess Davids for helping with the photographs of the specimens; to Bryan Huang for working with me through repetitive wet lab work; finally, to Dr. Jeffrey A. Cole of Pasadena City College whom introduced me to the field of Evolutionary Biology and Systematics.

ABSTRACT OF THE THESIS

High Throughput Sequencing of Ultraconserved Elements Reveals Inter-island Relationships of a California Channel Island Endemic Ant Species (*Aphaenogaster patruelis*)

by

Bo Huey Chiang

Master of Science in Biology

University of California San Diego, 2018

Professor David A. Holway, Chair

This study employs a phylogeographic analysis aimed at revealing inter-island relationships among populations of *Aphaenogaster patruelis*, an ant species endemic and restricted to the southern California Channel Islands and Isla Guadalupe. We collected samples from each of the five islands on which *A. patruelis* occurs, extracted the DNA, and constructed DNA libraries. We then enriched the libraries for ultraconserved elements (UCEs) and sent enriched product out for sequencing on an Illumina HiSeq 4000 at the Vincent J. Coates

Genomic Sequencing Laboratory at UC Berkeley. After quality control and assembly, 847 UCE loci were recovered. We analyzed aligned sequence data using PAUP (parsimony), RAxML under GTR+G model for maximum likelihood, and MrBayes for partitioned Bayesian Inference. The resulting trees from all analyses were concordant in major splits between island populations. Reconstructing the biogeographic history of this species is complicated by the lack of any extant mainland populations, but the topology of these phylogenies supports the hypothesis that the distribution of *A. patruelis* resulted from island hopping after a colonization event from the mainland. Santa Catalina Island could have been colonized first followed by San Clemente Island, which might have served as the origin of colonists that reached outlying islands to the south (Isla Guadalupe) and north (Santa Barbara and San Nicolas Islands). The distance between San Clemente Island and Isla Guadalupe (428 km) seems too great for an ant to reach via winged dispersal of alates; rafting thus seems a more likely mode of colonization for this ground-nesting ant species.

Introduction

Oceanic islands have long been important and popular study systems for testing ideas about evolution (Gillespie & Roderick 2002). Ever since Charles Darwin and his famous voyage to the Galapagos (Darwin, 1859), the unique geological makeup and isolation of islands has attracted evolutionary biologists. Discrete geographic barriers, which are characteristic of island systems, reduce or prevent genetic exchange with mainland populations and those on neighboring islands (Emerson, 2002). Patterns of migration among adjacent islands or between mainland and nearshore islands can be shaped by stochastic colonization events, resulting in puzzling distributions, relationships with living relatives, and origins that are difficult to infer (Garb & Gillespie, 2006; Hollocher & Williamson, 1996; Kane, Ochoa, Mathurin, & Erbe, 2005; Tinghitella, Zuk, Beveridge, & Simmons, 2011). The advancement of next generation sequencing techniques promises to advance an understanding of the evolutionary relationships among island organisms.

The California Channel Islands are an archipelago off the coast of southern California and consist of eight islands ranging in size from 2.6 to 249 km², 20 to 100 km from the continental mainland (Moody, 2000). None of these islands has ever been connected to the mainland. The northern islands (San Miguel, Santa Rosa, Santa Cruz, and Anacapa) were formerly connected into a single island during the last glacial maximum (Muhs et al., 2012), and thus share many biological affinities with one another. In contrast to the northern Channel Islands, the southern islands (San Nicolas, Santa Barbara, San Clemente, and Santa Catalina) were never connected to one another, and thus share fewer biological affinities with one another. Isla Guadalupe lies 280 km off of northern Baja California, and is not considered part of the

California Channel Islands archipelago, but shares numerous biological affinities including about 20 insular endemic plant species that reach their southern range limit on Isla Guadalupe. The isolated nature of Isla Guadalupe and the southern channel islands along with their complex topography and year-round Mediterranean climate (Power, 1980; Hochberg, 1993; Halvorson & Maender, 1994) gave rise to numerous endemic plant and animal taxa, and possibly provided an evolutionary refuge for species that colonized from the mainland, resulting in both relictual and *in situ* endemics occuring on the islands today. Currently, the southern channel islands are managed by several different agencies: San Nicolas and San Clemente Islands are controlled by the US Navy, Santa Catalina Island is privately owned with open space managed by the Catalina Island Conservancy, Santa Barbara Island is part of Channel Islands National Park, and Isla Guadalupe is part of the Pacific Islands Biosphere Reserve. Table 1 provides an island distance matrix.

The Myrmicine genus *Aphaenogaster* (Mayr, 1855) is diverse, abundant, and cosmopolitan with over 200 described species distributed across Australasia, Indomalaya, Malagasy, Nearctic, Neotropical, and Palearctic regions (Antweb.org). *Aphaenogaster* species are fairly common in Nearctic region and can be found throughout the United States. *Aphaenogaster patruelis* occurs on the four southern Channel Islands (Santa Catalina, San Clemente, Santa Barbara, and San Nicolas) and Isla Guadalupe, the species description appeared in "Espèces nouvelles de fourmis américaines" (Forel, 1886) with the type specimen from San Nicolas Island. Features of *A. patruelis* include a dark brown to black coloration with lighter appendage coloration, a pair of reduced propodeal spines, hair present all over the head,

mesosoma, and gaster, and a four-segmented antennae club (Fig. 1). *Aphaenogaster patruelis* from Santa Catalina Island are reddish in color (Fig. 1).

Apart from being described as an endemic species on California Channel Islands and Isla Guadalupe (Wetterer et al., 2000; Johnson & Ward, 2002), *A. patruelis* has not been researched extensively, and information such as the origin, behavior, and habitat preferences are lacking. Given the worldwide distribution of *Aphaenogaster*, one might hypothesize that the species in this genus are fairly competitive and aggressive, however *A. patruelis* appears competitively subordinate. *Aphaenogaster patruelis* appears to be a relictual species that occurs in most open habitat (scrub and grassland), can be found throughout the year, and forages diurnally and nocturnally. Like its congeners, *Aphaenogaster patruelis* is ground nesting. Due to its southern distribution, this species may have evolved in regions with warmer temperature and higher humidity relative to mainland California. The current distribution of *Aphaenogaster patruelis* does not overlap with any congeneric species, however *Aphaenogaster occidentalis* occurs (sparingly) on Santa Rosa Island.

Here, we examine in more detail the distribution and phylogeography of *A. patruelis* by utilizing high-throughput sequencing of ultraconserved elements (UCEs). UCEs are an efficient and cost effective method for obtaining large amounts of genome-wide sequence data, which can be used to reconstruct deep and shallow time scale evolutionary relationships (Smith, Harvey, Faircloth, Glenn & Brumfield, 2013). First described in Gil Bejerano et al. (2004) and later identified in organisms across numerous classes by Siepel et al. (2005), ultraconserved elements are short DNA regions that are highly conserved across broad taxonomic units with their function yet unclear (Dermitzakis, Reymond, & Antonarakis, 2005), however research has

shown correlations between UCEs and gene regulation and development (Ni et al., 2007; Visel et al., 2008). Ultraconserved elements are an extremely useful tool for phylogenetic reconstruction of evolutionary relationship because the "flanking regions" of UCEs contain enough variation to distinguish both deep and shallow evolutionary relationships (Faircloth, Branstetter, White, & Brady, 2014; Smith et al., 2013; Faircloth, Sorenson, Santini, Alfaro, 2013; Faircloth et al., 2012; McCormack et al., 2012). Numerous studies (Faircloth, Branstetter, White, & Brady, 2015; JEŠOVNIK et al., 2017; Blaimer et al., 2015; Blaimer, Lloyd, Guillory, & Brady, 2016) have demonstrated the effectiveness of UCEs in resolving phylogenetic relationships in ants.

Here, we used UCE sequence data from samples of *A. patruelis* from multiple locations on each island where the species occurs to address two contrasting hypotheses regarding its species-level phylogeography. Hypothesis 1: *A. patruelis* (or a related ancestor) was once a widespread on the adjacent mainland, colonized the closest island (Santa Catalina Island) and then sequentially island hopped to the other islands as a function of inter-island distance (Table 1) Hypothesis 2: *A. patruelis* became established on the islands through multiple, independent colonization events from the mainland. Fig 2 contains hypothetical topologies for each proposed hypothesis.

Methods

The complete list of specimens (including collection information) used in this study can be found in Supplementary Table. 1 and visual representation of geographic location of islands with specimen collection sites can found in Fig 3. All specimens were preserved in 90~100% ethanol following collection and stored in -20 celsius refrigerator. Prior to DNA extraction, specimens were removed from the refrigerator and dried before having the gaster removed by sterilized razor blade in order to avoid contamination from gut contents. We extracted total genomic DNA using a DNeasy Blood and Tissue Kit (Qiagen Inc., Valencia, CA) following the instructions included in the kit along with modifications optimized for obtaining higher final product concentration from arthropods. Tissues were ruptured using Benchmark Scientific D1030 BeadBug™ Homogenizer in centrifuge tubes instead of traditional mortar and pestle to minimize sample loss during transfer and increase surface area for proteinase K. We added of DTT to lysate for prevention of oxidation damage, and RNaseA for removal of RNA, and increased incubation time in proteinase K mixture to 48 hours for better breakdown of tissues. Lastly, we conducted the first elution in nuclease-free water while the second elution in product of first elution to increase final concentration. Before proceeding, the concentrations of every sample were tested using Invitrogen™ Qubit™ 1X dsDNA HS Assay Kit on a Qubit 2.0 Fluorometer to ensure the requirement of minimum sample concentration of 0.05 ng/ul was met. We sheared genomic DNA with two shearing cycles of 30 seconds each and then treated each sample with a 1-minute cooldown period after each shearing session on a Diagenode Bioruptor 300, targeting fragment sizes of 300 - 1000 bp in length.

The complete guide for library preparation (KAPA Biosystems DNA Hyperprep kit) and target capture and enrichment(Arbor Biosciences MyBaits) as well as pooling, are included in the Supplementary Files. For library preparation, we added Sera-Mag™ Magnetic SpeedBeads and PEG mixture to sheared DNA samples to retrieve desired fragments. The volume of speedbead added is extremely important in this process for fragments captured by beads will be determined by the amount of beads added. Too many beads result in the preservation of short and unuseful fragments, while too few beads result in the loss of desired fragments. The second and third steps involve blunt-end repair and stub ligation using KAPA DNA Hyperprep kit. Each end of the fragments are repaired to ensure the presence of functional 5' phosphate and 3' hydroxyl groups then add on stubs that contain the primer binding sites for illumina adapters. Before proceeding, a round of bead clean-up is required for size selection. This process is similar to the first step but here we release the DNA from beads at the end of the process. Next we amplify the samples using HiFi ready mix from KAPA kit in a thermocycler with unique i5 and i7 primers (Illumina Tru-Seq style adapters) for 12 cycles to amplify the adapters onto the fragments of each sample. The adapters are unique indices used to decipher fragments from different samples. The final step of library preparation is another bead clean-up for another round of size selection. The post library preparation is comprised of two processes and one cleaning step in between; these preparations use arbor bioscience myBaits® Kit and MyBaits Hymenoptera v. 2. In the hybridization process the libraries are first denatured, then have blockers hybridized with adapters and baits hybridized with the target gene region. The cleaning process binds targeted libraries to streptavidin-coated magnetic beads then washes away most non-targeted libraries. During the target enrichment process, the targeted libraries are first

released from magnetic beads then amplified in thermocycler. All our libraries are then pooled and sequenced on an Illumina HiSeq 4000 at the Vincent J Coates Genomic Sequencing Lab at UC Berkeley.

All sequencing data were downloaded from UC Berkeley server and processed using phyluce 1.6.7 software package (Faircloth, 2016). Phyluce is based on python version 2.7 and is only supported on Mac and Linux platform, however this can be easily overcome using Oracle VM VirtualBox with a Linux Mint 19 system. Complete codes used in this project are included can be found on first author's GitHub repository (https://github.com/discipleofdarwin/UCE-Analysis-Codes). Sequences were first trimmed and ambiguous data were removed using illumiprocessor 2.0.8 and then assembled into contig and written into individual fastq files by Abyss 1.5.2 (Simpson et al., 2009) with kmer setting of 60 as suggested by Faircloth (2016). After assembly, the data from each library along with probe set Hymenoptera 2.5Kv2 (Branstetter, Longino, Ward, & Faircloth, 2017; Faircloth et al., 2012) were used to search through each library data for the generation of lastz files that contained all the contigs that matched the probes and a combined sqlite matrix that maps the contigs matched in each library. We then generated a config file(.conf) that included a sample names block and UCE region names block under. If done without the "--incomplete-matrix" command, then phyluce will only list the name of UCE regions shared by all libraries, and depending on the dataset can sometimes lead to no UCE region names listed if no UCE regions are shared by all libraries, however if "--incomplete-matrix" was included then all UCE region that were not shared by all will be listed as well. The combined fasta file that included contigs of each library tagged with library name and UCE region name was produced using the config file together with

contigs and sqlite matrix. The combined fasta then gives rise to fasta files that were named and categorized by UCE regions that contained alignment of contig of each library finished by mafft 7.130 (Katoh & Stanley, 2013). The contig labels inside the file were removed so only the library name was left as indicator for each contig, the removal of contig labels were done because programs for phylogenetic analysis do not accept names that are too long. The resulting files from the last step were further processed by removing files with alignment that includes less than 75% of the taxa and shorter than 400 bp in length, this process was done to remove less informative or badly sequenced data, and in this project the process retrieved 847 UCE regions. All alignments are then concatenated into one complete dataset with total of 661,721 characters (gap and missing included) on SequenceMatrix 1.8 (Vaidya, Lohman, & Meier, 2011), the data were exported once in nexus format for PAUP parsimony analysis and a second time in phylip format for RAxML maximum likelihood (ML). All phylogenetic analyses done for this project employed Cipres Science Gateway (Miller, Pfeiffer, & Schwartz, 2010). For Parsimony analysis, PAUP (Swofford, 1991) was chosen with set parameter of heuristic search with sequences added randomly and swapped on all the best tree found for $2.0*10^5$ repetition with 1000 heuristic bootstrap replicate and 1000 sequences added randomly for each replicate. Models for ML and partitioned bayesian inference (BI) were selected by jmodeltest2 (Darriba, Taboada , Doallo , & Posada, 2012; Guindon & Gascuel, 2003) with parameter of 3 substitute scheme, ML optimized, and NNI for base tree selection. ML analysis ran on RAxML-HPC v.8 under GTR+G model with 1000 bootstrap. BI completed with MrBayes 3.22 (Ronquist, Huelsenbeck,& Teslenko, 2011) following set parameter of four chains mcmc, for $2*10^7$ generation, assuming all starting trees to be equally likely, branch unlinked, analysis repeated once, and a relative burn-in of 25% selected

using Tracer v.1.6 (Rambaut, Drummond, & Suchard, 2014). All analyses were built using *Novomessor albisetosus* and *Novomessor cockerelli* as outgroup taxa. Results were visualized and annotated with TreeGraph2 (Stöver & Müller 2010) and can be found in Fig 4, 5, and 6. Out group branch is removed so the branch length of the rest of the clades can be seen.

Results

As shown in figures, the use of genome-wide sequence data obtained from UCEs produced phylogenies with with high bootstrap for each species considered. *Aphaenogaster* species from eastern North America (*A. picea, A. miamiana, A. tennesseensis, A. fulva, A. ashmeadi, A. lamellidens,* and *A. rudis*) form a distinct clade separated from *Aphaenogaster* species from western North America (*A. patruelis, A. occidentalis,* and *A. mutica*). *A. patruelis* shares a more recent common ancestor with southerly distributed *A. mutica* from Baja California rather than with the more northerly distributed *A. occidentalis*.

Within the *A. patruelis* clade, specimens from each island form a strongly supported monophyletic clade except the specimens from Santa Barbara Island that is rendered paraphyletic with respect to samples from the monophyletic from San Nicolas island. Inter-island relationships observed within *A. patruelis* resembled the hypothetical tree in Fig. 2, which is based on a scenario of nearest-distance, island hopping from Santa Catalina Island. While we are unable to determine for certain the origins of colonizers for each island, the paraphyletic relationship of Santa Barbara Island specimens with respect to those from San Nicolas Island depicted in our results indicate island hopping from Santa Barbara to San Nicolas Island.

Discussion

According to the *Aphaenogaster* phylogeny by DeMarco & Cognato (2016) *A. boulderensis* appears to be the sister taxon to *A. mutica*, and that the most recent common ancestor of *A. patruelis* was the ancestor to the clade that includes *A. boulderensis* and *A. mutica*. This shows that *Aphaenogaster patruelis* is not a derivative of *Aphaenogaster occidentalis*, which does occur on the northern Channel Islands. This relationship also rules out the possibility that *Aphaenogaster patruelis* shared an *Aphaenogaster occidentalis* ancestor that colonized one of the southern Channel Islands, evolved into *Aphaenogaster patruelis* and then subsequently colonized the other southern Channel Islands and Isla Guadalupe.

The close agreement between our observed phylogenies and that proposed from Hypothesis (Fig. 2), suggests that the colonization of the Channel Islands by *A. patruelis* (or its ancestor) originated with a colonization event from the mainland to Santa Catalina Island, followed by subsequent colonizations to adjacent islands to the south (San Clemente and Isla Guadalupe) and to the north (San Nicolas and Santa Barbara). The monophyly of island clades (excluding Santa Barbara Island) suggest single colonization events followed by geographic isolation and no gene flow from nearby islands. It should be emphasized that we can't rule out a scenario of multiple, independent colonization events from a hypothetical mainland population of *A. patruelis*.

Under any scenario of colonization the monophyly of four out of five island clades at least suggests that inter-island colonization events by *A. patruelis* are relatively rare. A hypothesized colonization of Santa Barbara Island by San Clemente Island *A. patruelis* seems

surprising given that the distance between San Clemente Island and Santa Barbara Island is 1.4 times greater than distance between Santa Barbara Island and Santa Catalina Island. Assuming that *A. patruelis* colonized Santa Barbara Island from San Clemente Island, this phenomenon illustrates that island hopping events are stochastic in nature and the result of chance dispersal.

Currently, few studies document the phylogeography and diversification of arthropods from the California Channel Islands. Chatzimanolis, Norris, & Caterino (2010) examined the phylogeography of a multi-island endemic beetle *Coelus pacificus* (Coleoptera:Tenebrionidae) on the Channel Islands and found evidence for multiple colonization events from the mainland. Other examples of multi-island endemics on the Channel Islands include *Ypsolopha lyonothamnus* that feed on ironwood (*Lyonothamnus floribundus* Gray) described by Miller (1985) & Powell (1994) for being distantly related to adjacent mainland congeners, instead related to either desert taxa or are relicts with remote mainland relationships.

Although the results of this study seem consistent with the island hopping hypothesis, we can't rule out independent colonizations from a hypothetical (i.e., now extinct) mainland population of *A. patruelis* (or its ancestor). A scenario of multiple, independent colonization events from the mainland, however, would have had to have produced a tree that by coincidence matches that produced by nearest-distance island hopping. Evidence to support the likelihood of nearest distance island hopping comes from the relationship between *A. patruelis* from San Nicolas and Santa Barbara islands. Under any scenario of dispersal and colonization, the notion that geographic separation of islands prevent gene flow is still supported. Given the lack of gene flow among island populations, conservation management strategies should recognize that each island population (perhaps excluding San Nicolas and Santa Barbara Islands) should be treated

and managed as independent units. More information about these species seems increasingly important in the face of current and future threats to *Aphaenogaster patruelis* on the Channel Islands, such as non-native species like the Argentine ant (*Linepithema humile*) and anthropogenic disturbances such as fire.

Figures

Figure 1. Photo representation of the species *Aphaenogaster patruelis* with scale in cm and mm. A & B) Facial and body shot of specimen from Santa Barbara island displaying general look among different island populations. C & D) Shots of Santa Catalina Island, displaying the unique coloration in Santa Catalina *Aphaenogaster patruelis.*

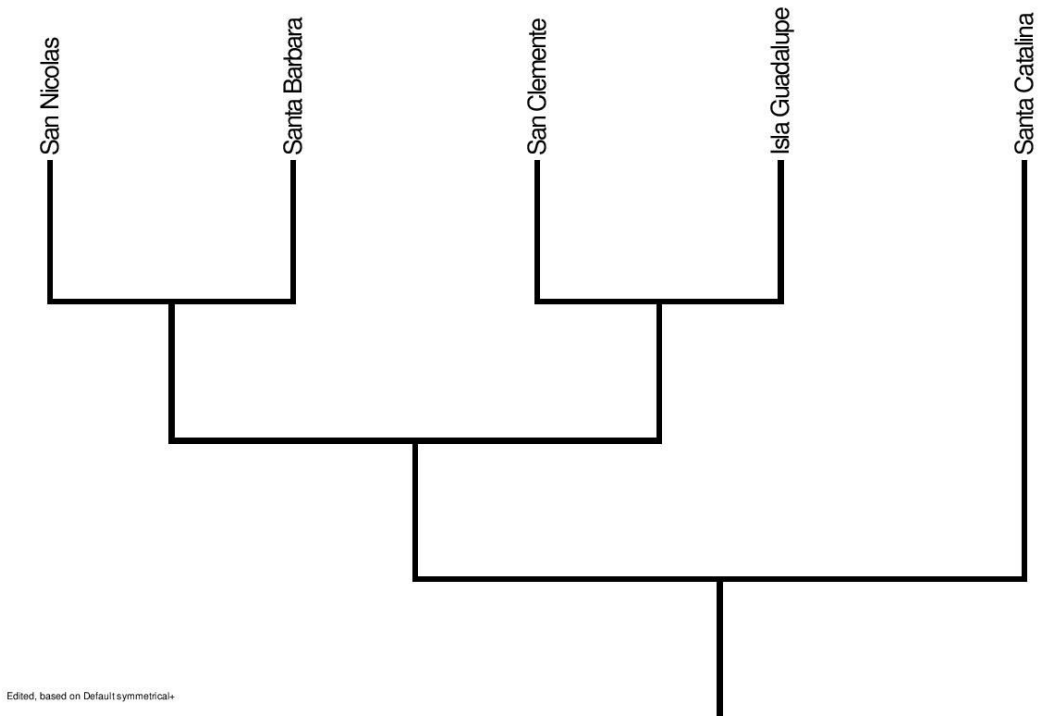

Figure 2. Hypothetical phylogenetic topology constructed based on closest distance relationship shown in Table 1.

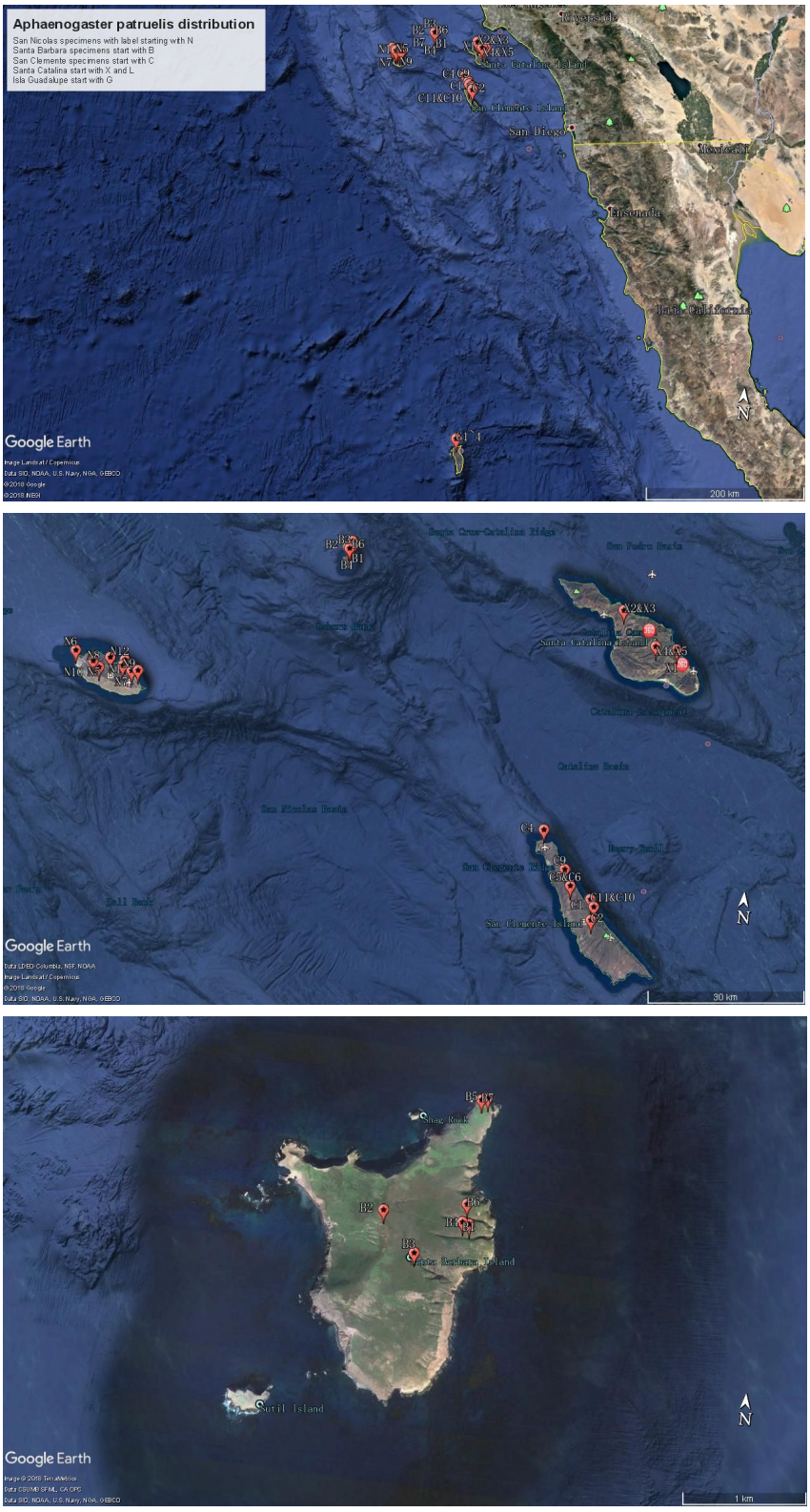

Figure 3.1, 3.2, 3.3 Display of geographical location of California Channel Island and Isla Guadalupe and specimen locality on each island.

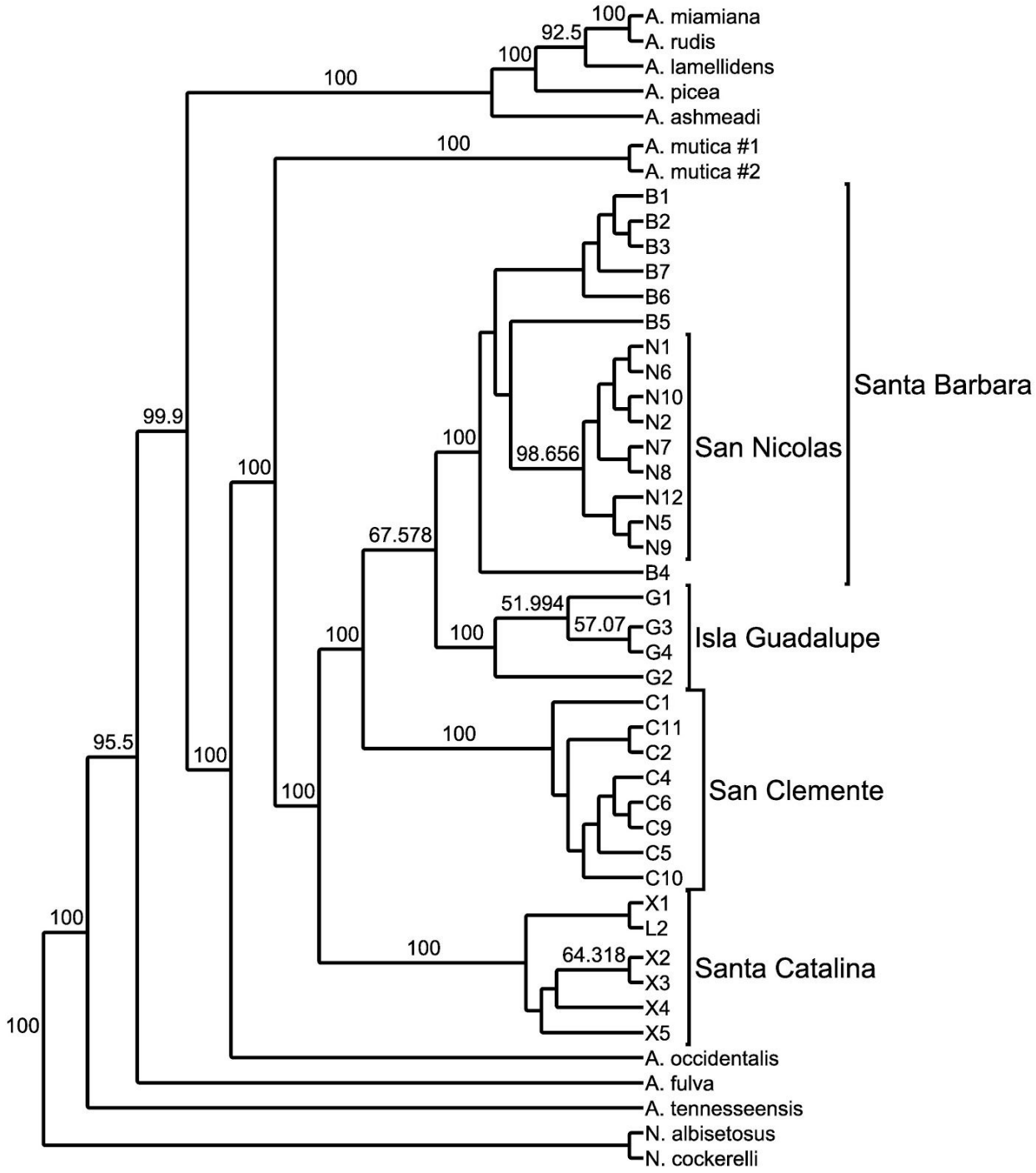

Figure 4. Parsimony tree generated using concatenated data of 847 UCE loci with $2x10^5$ heuristic search and 1000 bootstrap replicates. Node labels display bootstrap support values.

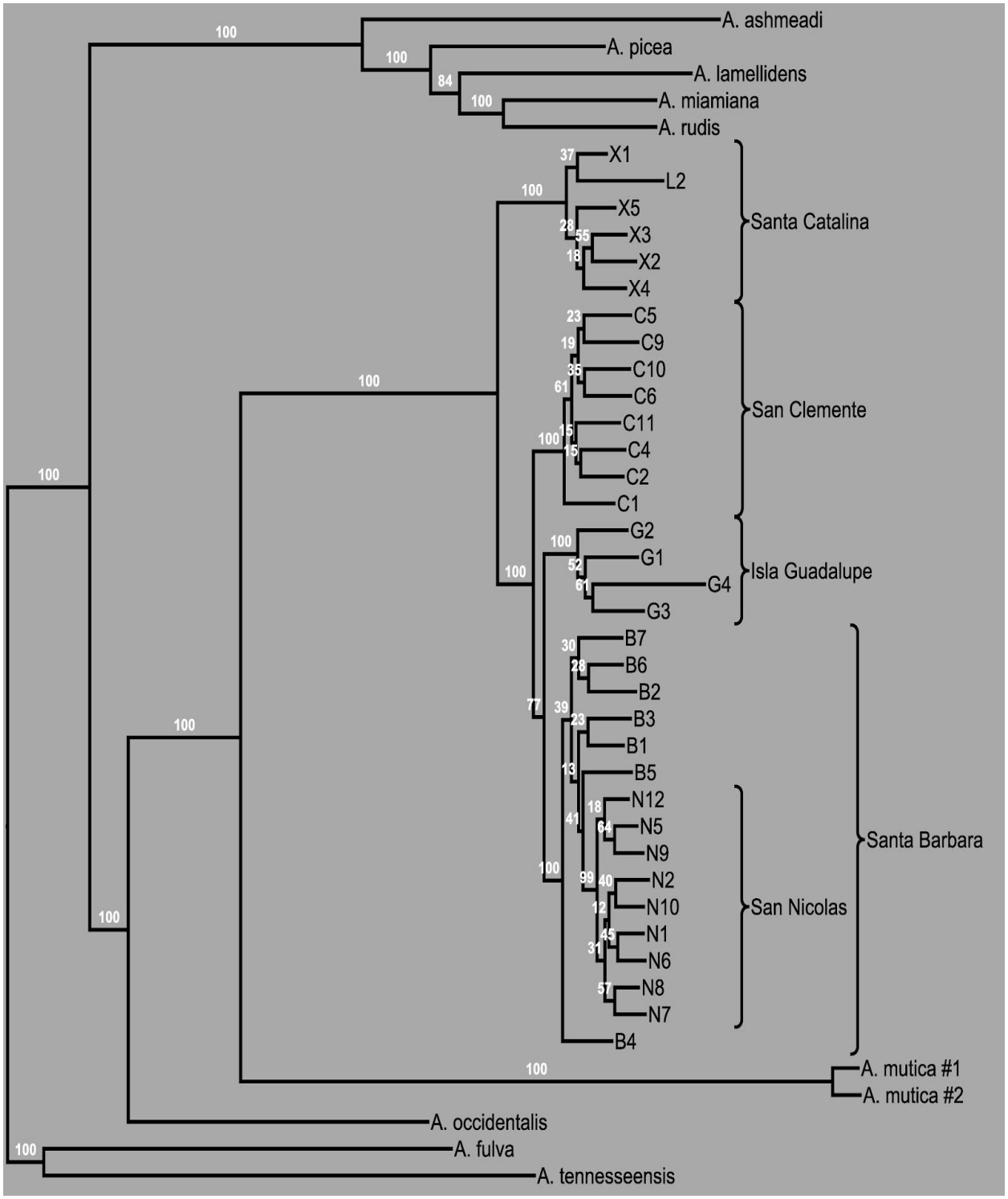

Figure 5. Maximum Likelihood tree constructed under jmodeltest2 selected GTR+G model. Legends display island clades and with outgroup removed for display of branch length.

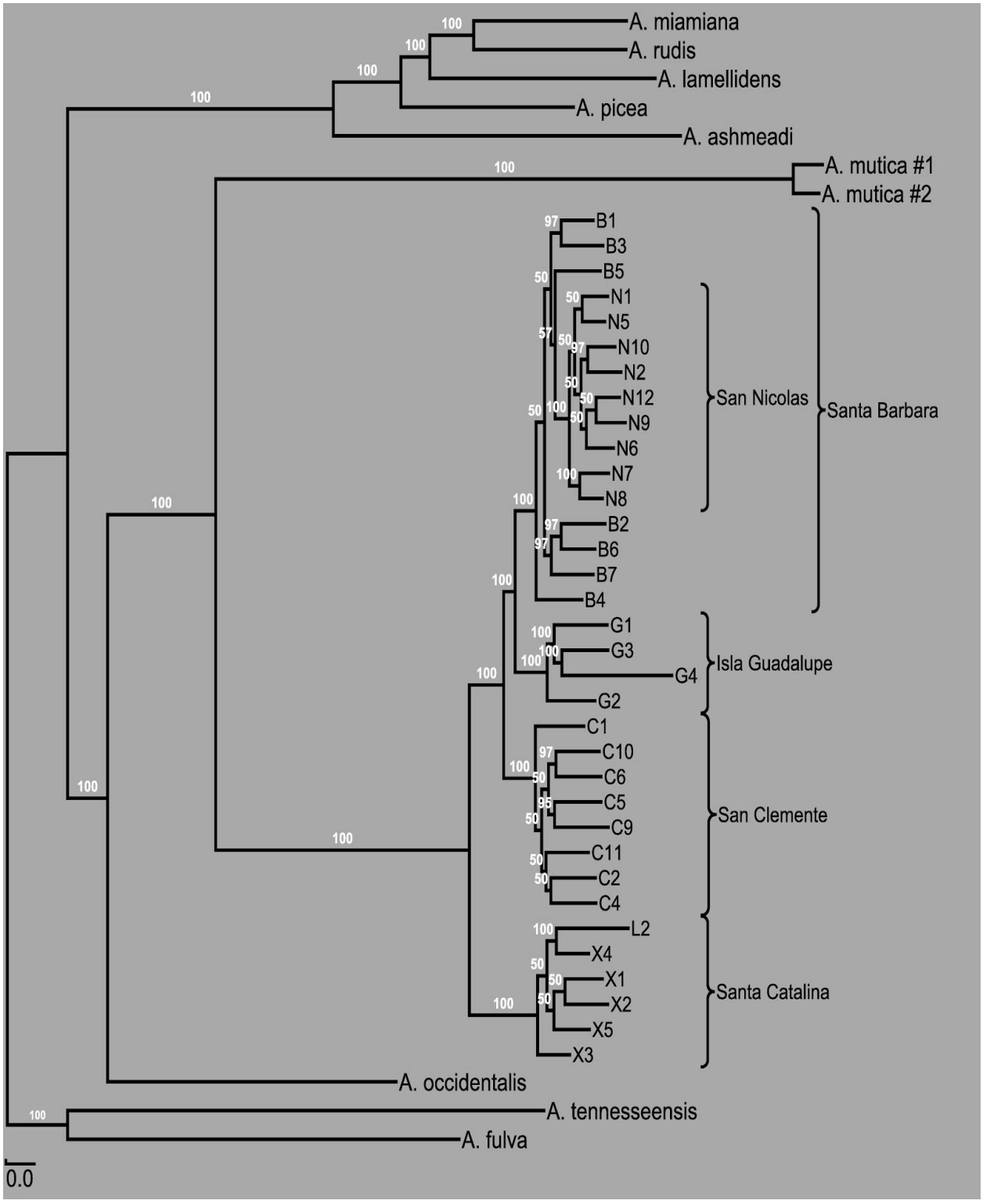

Figure 6. Partitioned Bayesian Inference tree generated using partitioned dataset with specific substitution scheme and rate selected for all 847 UCE loci.

Tables

Table 1. Island matrix distance shown with abbreviations (SCA: Santa Catalina Island, SCL: San Clemente Island, SBA: Santa Barbara Island, SNI: San Nicolas Island, and ISG: Isla Guadalupe) that display the distance in kilometers between islands and between each island and the mainland.

	SCA	SCL	SBA	SNI	ISG	Closest distance to mainland
SCA	----------					35.4 km
SCL	54.4 km	----------				66 km
SBA	58.4 km	81 km	----------			61.2 km
SNI	102 km	101 km	49.9 km	----------		98 km
ISG	482 km	428 km	497 km	480 km	----------	241 km

Appendix

Supplementary Table 1. Table of collection information of specimens used in this project, including shorthand abbreviation, locality, collection dates, and name of collectors.

Label	Species	State	County	Island	Location	Coordinates	Date (MM/DD/YYYY)	Collector
A30	*A. picea*	Michigan	Ingham Co.	X	Rose Lake	(42°48'1.44"N, 84°22'53.70"W)	5/31/2009	Bernice B. DeMarco
A113	*A. miamiana*	Mississippi	Oktibeha Co.	X	Osborne Prairie	(33°30'52.98"N, 88°43'54.00"W)	03/08/2010	Bernice B. DeMarco
A142	*A. tennesseensis*	Minnesota	Olmsted Co.	X	Quarry Hill Nature Center	(44° 3'35.40"N, 92°29'13.80"W)	7/11/2010	Bernice B. DeMarco
A209	*A. fulva*	Virginia	Surry Co.	X	York River State Park	(37°24'34.02"N, 76°42'49.02"W)	8/12/2010	Bernice B. DeMarco
A282	*A. ashmeadi*	North Carolina	Bladen Co.	X	Culbreth Smith & Lula Loop Rd	(34°45'31.74"N, 78°35'23.04"W)	7/27/2011	Bernice B. DeMarco
A295	*A. lamellidens*	North Carolina	Orange Co.	X	Duke Forest Warming Project	(36° 2'11.16"N, 79° 4'38.88"W)	7/28/2011	Bernice B. DeMarco
Am1	*A. mutica*	Baja California Sur	X	X	1km W La Laguna	(23°33'0.00"N, 110°0'0.00"W,1800m)	12/30/2003	Philip S. Ward
Am2	*A. mutica*	Baja California Sur	X	X	1km W La Laguna	(23°33'0.00"N, 110° 0'0.00"W, 1800m)	12/30/2003	Philip S. Ward
Ar	*A. rudis*	New Jersey	Middlesex Co.	X	Cheesequake State Park	(40°26'16.98"N, 74°15'36.00"W)	5/30/2010	Bernice B. DeMarco
B1	*A. patruelis*	California	Santa Barbara Co.	Santa Barbara	Middle Canyon	(33°28'39.60"N, 119° 1'50.90"W)	7/17/2014~7/18/2014	David A. Holway
B2	*A. patruelis*	California	Santa Barbara Co.	Santa Barbara	SW of N Peak	(33°28'40.60"N, 119° 2'16.60"W)	7/17/2014	David A. Holway
B3	*A. patruelis*	California	Santa Barbara Co.	Santa Barbara	X	(33°28'29.84"N, 119° 2'5.61"W)	7/19/2014	Ida Naughton
B4	*A. patruelis*	California	Santa Barbara Co.	Santa Barbara	Grave Yard Canyon	(33°28'39.20"N, 119° 1'48.70"W)	7/18/2014~7/19/2014	Ida Naughton
B5	*A. patruelis*	California	Santa Barbara Co.	Santa Barbara	North Point	(33°29'15.54"N, 119° 1'45.66"W)	7/XX/2014	Ida Naughton
B6	*A. patruelis*	California	Santa Barbara Co.	Santa Barbara	Cave Canyon	(33°28'44.70"N, 119° 1'49.90"W)	7/17/2014~7/18/2014	David A. Holway
B7	*A. patruelis*	California	Santa Barbara Co.	Santa Barbara	Under Atraplex	(33°29'14.82"N, 119° 1'47.76"W)	7/XX/2014	Ida Naughton
C1	*A. patruelis*	California	Los Angeles Co.	San Clemente	Horton Canyon	(32°54'20.16"N, 118°28'42.78"W)	5/22/2016	Ida Naughton
C2	*A. patruelis*	California	Los Angeles Co.	San Clemente	Norton Canyon	(32°52'59.05"N, 118°28'55.74"W)	5/21/2016	Ida Naughton
C4	*A. patruelis*	California	Los Angeles Co.	San Clemente	NW Wilson Cove	(33° 1'41.24"N, 118°35'28.19"W)	3/13/2015	David A. Holway

Supplementary Table 1. Continued

C5	*A. patruelis*	California	Los Angeles Co.	San Clemente	Plot 7U	(32°56'15.30"N, 118°31'44.30"W)	3/14/2015	David A. Holway
C6	*A. patruelis*	California	Los Angeles Co.	San Clemente	Plot 7U	(32°56'15.30"N, 118°31'44.30"W)	3/14/2015	David A. Holway
C9	*A. patruelis*	California	Los Angeles Co.	San Clemente	Plot 3I	(32°57'52.60"N, 118°32'33.70"W)	7/28/2017	David A. Holway
C10	*A. patruelis*	California	Los Angeles Co.	San Clemente	Burns Canyon	(32°55'2.80"N, 118°29'9.00"W)	4/25/2018	David A. Holway
C11	*A. patruelis*	California	Los Angeles Co.	San Clemente	Burns Canyon	(32°55'2.80"N, 118°29'9.00"W)	4/25/2018	David A. Holway
G1	*A. patruelis*	Baja California	X	Isla Gradalupe	X	(29° 6'19.77"N, 118°19'11.37"W)	X	X
G2	*A. patruelis*	Baja California	X	Isla Gradalupe	X	(29° 6'19.77"N, 118°19'11.37"W)	X	X
G3	*A. patruelis*	Baja California	X	Isla Gradalupe	X	(29° 6'19.77"N, 118°19'11.37"W)	X	X
G4	*A. patruelis*	Baja California	X	Isla Gradalupe	X	(29° 6'19.77"N, 118°19'11.37"W)	X	X
L2	*A. patruelis*	California	Los Angeles Co.	Santa Catalina	X	X	8/3/2007	X
N1	*A. patruelis*	California	Ventura Co.	San Nicolas	Airfield	(33°14'2.04"N, 119°27'15.59"W)	6/6/2016	Ida Naughton
N2	*A. patruelis*	California	Ventura Co.	San Nicolas	X	X	X	X
N4	*A. occidentalis*	California	Marin Co.	X	Marin Headlands	(37°49'41.58"N, 122°29'51.06"W)	7/XX/2014	Ida Naughton
N5	*A. patruelis*	California	Ventura Co.	San Nicolas	Grand Canyon	(33°14'5.50"N, 119°31'18.40"W)	9/14/17	Ida Naughton
N6	*A. patruelis*	California	Ventura Co.	San Nicolas	Rock Crusher	(33°15'35.90"N, 119°34'19.30"W)	9/12/2017	David A. Holway
N7	*A. patruelis*	California	Ventura Co.	San Nicolas	X	(33°14'7.98"N, 119°26'29.64"W)	6/5/2016	David A. Holway
N8	*A. patruelis*	California	Ventura Co.	San Nicolas	X	(33°15'12.00"N, 119°30'1.68"W)	6/3/2016	David A. Holway
N9	*A. patruelis*	California	Ventura Co.	San Nicolas	X	(33°14'29.70"N, 119°28'25.08"W)	6/3/2016	David A. Holway
N10	*A. patruelis*	California	Ventura Co.	San Nicolas	X	(33°14'40.50"N, 119°32'4.50"W)	6/4/2016	David A. Holway
N12	*A. patruelis*	California	Ventura Co.	San Nicolas	X	(33°15'4.62"N, 119°28'12.84"W)	6/4/2016	David A. Holway
Na	*N. albisetosus*	Arizona	Cochise Co.	X	Chiricahua Mts SWRS 7kmW	(31°52'60.00"N, 109°12'0.00"W)	8/8/2009	Bernice B. DeMarco
Nc	*N. cockerelli*	Arizona	Cochise Co.	X	Chiricahua Mts SWRS 7kmW	(31°52'60.00"N, 109°12'0.00"W)	8/8/2009	Bernice B. DeMarco
X1	*A. patruelis*	California	Los Angeles Co.	Santa Catalina	X	(33°21'3.89"N, 118°21'10.26"W, 458m)	5/22/2018	David A. Holway
X2	*A. patruelis*	California	Los Angeles Co.	Santa Catalina	X	(33°24'36.43"N, 118°28'0.80"W, 165m)	5/22/2018	David A. Holway

Supplementary Table 1. Continued

X3	*A. patruelis*	California	Los Angeles Co.	Santa Catalina	X	(33°24'36.43"N, 118°28'0.80"W, 165m)	5/22/2018	David A. Holway
X4	*A. patruelis*	California	Los Angeles Co.	Santa Catalina	X	(33°21'12.17"N, 118°23'47.76"W, 262m)	5/22/2018	David A. Holway
X5	*A. patruelis*	California	Los Angeles Co.	Santa Catalina	X	(33°21'12.17"N, 118°23'47.76"W, 262m)	5/22/2018	David A. Holway

References:

AntWeb. Available from: https://www.antweb.org/description.do?subfamily=myrmicinae&genus=aphaenogaster&rank=genus. Accessed 21 November 2018

AntWeb. Available from: https://www.antweb.org/description.do?genus=aphaenogaster&species=patruelis&rank=species. Accessed 21 November 2018

Blaimer, B. B., Brady, S. G., Schultz, T. R., Lloyd, M. W., Fisher, B. L., & Ward, P. S. (2015). Phylogenomic methods outperform traditional multi-locus approaches in resolving deep evolutionary history: a case study of formicine ants. *BMC Evolutionary Biology*, 15(1), 271.

Blaimer, B. B., Lloyd, M. W., Guillory, W. X., & Brady, S. G. (2016). Sequence capture and phylogenetic utility of genomic ultraconserved elements obtained from pinned insect specimens. *PLoS One*, 11(8), e0161531.

Branstetter, M. G., Longino, J. T., Ward, P. S., & Faircloth, B. C. (2017). Enriching the ant tree of life: enhanced UCE bait set for genome‐scale phylogenetics of ants and other Hymenoptera. *Methods in Ecology and Evolution*, 8(6), 768-776.

Chatzimanolis, S., Norris, L. A., & Caterino, M. S. (2010). Multi-island endemicity: phylogeography and conservation of Coelus pacificus (Coleoptera: Tenebrionidae) darkling beetles on the California Channel Islands. *Annals of the Entomological Society of America*, 103(5), 785-795.

Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* 9(8), 772.

Darwin, C. (1999). *The origin of species: by means of natural selection or the preservation of favoured races in the struggle for life*. Bantam Classics.

DeMarco, B. B., & Cognato, A. I. (2016). A multiple‐gene phylogeny reveals polyphyly among eastern North American Aphaenogaster species (Hymenoptera: Formicidae). *Zoologica Scripta*, 45(5), 512-520.

Dermitzakis, E. T., Reymond, A., & Antonarakis, S. E. (2005). Conserved non-genic sequences—an unexpected feature of mammalian genomes. *Nature Reviews Genetics*, *6*(2), 151.

Emerson, B. C. (2002). Evolution on oceanic islands: molecular phylogenetic approaches to understanding pattern and process. *Molecular ecology*, 11(6), 951-966.

Faircloth, B. C., McCormack, J.E., Crawford, N.G., Harvey, M.G., Brumfield, R.T., Glenn, T.C. 2012. Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. *Systematic Biology* 61: 717–726. doi:10.1093/sysbio/SYS004.

Faircloth, B. C., Sorenson, L., Santini, F., & Alfaro, M. E. (2013). A phylogenomic perspective on the radiation of ray-finned fishes based upon targeted sequencing of ultraconserved elements (UCEs). *PLoS One*, 8(6), e65923.

Faircloth, B. C., Branstetter, M. G., White, N. D., & Brady, S. G. (2015). Target enrichment of ultraconserved elements from arthropods provides a genomic perspective on relationships among Hymenoptera. *Molecular Ecology Resources*, 15(3), 489-501.

Forel, A. (1886). Espèces nouvelles de fourmis américaines. *Annales*.

Garb, J. E., & Gillespie, R. G. (2006). Island hopping across the central Pacific: mitochondrial DNA detects sequential colonization of the Austral Islands by crab spiders (Araneae: Thomisidae). *Journal of Biogeography*, 33(2), 201-220.

Gillespie, R. G., & Roderick, G. K. (2002). Arthropods on islands: colonization, speciation, and conservation. *Annual review of entomology*, 47(1), 595-632.

Guindon S and Gascuel O (2003). A simple, fast and accurate method to estimate large phylogenies by maximum-likelihood". *Systematic Biology* 52: 696-704.

Halvorson, W. L. & Maender, G. J. (eds) (1994) The Fourth California Islands Symposium: update on the status of resources. Santa Barbara Museum of Natural History, Santa Barbara.

Hochberg, F. G. (ed.) (1993) The third California Islands Symposium: Recent Advances in California Islands Research. Santa Barbara Museum of Natural History, Santa Barbara.

Hollocher, H. & Williamson, M. (1996). Island Hopping in Drosophila: Patterns and Processes: Discussion. *Philosophical Transactions of the Royal Society of London Series B*, 351, 743.

JEŠOVNIK, A., SOSA‐CALVO, J. E. F. F. R. E. Y., Lloyd, M. W., Branstetter, M. G., Fernandez, F., & Schultz, T. R. (2017). Phylogenomic species delimitation and host‐symbiont coevolution in the fungus‐farming ant genus Sericomyrmex M ayr (Hymenoptera: Formicidae): ultraconserved elements (UCEs) resolve a recent radiation. *Systematic Entomology*, 42(3), 523-542.

Johnson, R. A., & Ward, P. S. (2002). Biogeography and endemism of ants (Hymenoptera: Formicidae) in Baja California, Mexico: a first overview. *Journal of Biogeography*, 29(8), 1009-1026.

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular biology and evolution*, 30(4), 772-780.

Kane, E. C., Ochoa, R., Mathurin, G., & Erbe, E. F. (2005, December). Raoiella indica Hirst (Acari: Tenuipalpidae): An island-hopping mite pest in the Caribbean. In *Abstract. ESA Meeting, Fort Lauderdale*.

Mayr, G. L. (1855). *Formicina austriaca: Beschreibung der bisher im österreichischen Kaiserstaate aufgefundenen Ameisen, nebst Hinzufügung jener in Deutschland, in der Schweiz und in Italien vorkommenden Arten*. Druck von Karl Ueberreuter.

McCormack, J. E., Faircloth, B. C., Crawford, N. G., Gowaty, P. A., Brumfield, R. T., & Glenn, T. C. (2012). Ultraconserved elements are novel phylogenomic markers that resolve placental mammal phylogeny when combined with species-tree analysis. *Genome research*, 22(4), 746-754.

Menard, Y. (2014, November 6). Scientist Shares News on Isolated Baja California Island: Isla Guadalupe. In National Park Service-Channel Islands National Park California. Retrieved from https://www.nps.gov/chis/learn/news/pr11062014.htm

Miller, S. E. (1985). The California Channel Islands--Past, present, and future: an entomological perspective.

Miller, M. A., Pfeiffer, W., & Schwartz, T. (2010, November). Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In *Gateway Computing Environments Workshop (GCE)*, 2010 (pp. 1-8). Ieee.

Moody, A. (2000). Analysis of plant species diversity with respect to island characteristics on the Channel Islands, California. *Journal of Biogeography*, 27(3), 711-723.

Muhs, D.R., Simmons,K.R., Schumann,R.R., Groves, L.T.,Mitrovica, J.X.,Laurel, D., 2012. Sea-level history during the last interglacial complex on San Nicolas Island, California: implications for glacial isostatic adjustment processes, paleozoogeography and tectonics. *Quaternary Science Reviews* 37, 1–25.

Ni, J. Z., Grate, L., Donohue, J. P., Preston, C., Nobida, N., O'Brien, G., ... & Ares, M. (2007). Ultraconserved elements are associated with homeostatic control of splicing regulators by alternative splicing and nonsense-mediated decay. *Genes & development*, 21(6), 708-718.

Powell, J. A. (1994). Biogeography of Lepidoptera on the California Channel Islands. In *The Fourth California Islands Symposium: Update on the Status of Resources* (pp. 449-64). Santa Barbara, CA: Santa Barbara Museum of Natural History.

Power, D. M. (ed.) (1980) The California Islands. Proceedings of a multidisciplinary symposium. Santa Barbara Museum of Natural History, Santa Barbara.

Rambaut, A., Drummond, A. J., & Suchard, M. (2014). Tracer v1. 6 http://beast. bio. ed. ac. uk. *Tracer* (visited on 2017-06-12).

Ronquist, F., Huelsenbeck, J., & Teslenko, M. (2011). Draft MrBayes version 3.2 manual: tutorials and model summaries. *Distributed with the software from http://brahms. biology. rochester. edu/software. html.*

Simpson, J. T., Wong, K., Jackman, S. D., Schein, J. E., Jones, S. J., & Birol, I. (2009). ABySS: a parallel assembler for short read sequence data. *Genome research*, gr-089532.

Smith, B. T., Harvey, M. G., Faircloth, B. C., Glenn, T. C., & Brumfield, R. T. (2013). Target capture and massively parallel sequencing of ultraconserved elements for comparative studies at shallow evolutionary time scales. *Systematic biology*, 63(1), 83-95.

Swofford, D. L. (1991) PAUP: Phylogenetic Analysis Using Parsimony, Version 3.1 Computer program distributed by the Illinois Natural History Survey, Champaign, Illinois.

Stöver, B. C., & Müller, K. F. (2010). TreeGraph 2: combining and visualizing evidence from different phylogenetic analyses. *BMC bioinformatics*, 11(1), 7.

Tinghitella, R. M., Zuk, M., Beveridge, M., & Simmons, L. W. (2011). Island hopping introduces Polynesian field crickets to novel environments, genetic bottlenecks and rapid evolution. *Journal of evolutionary biology*, 24(6), 1199-1211.

Vaidya, G., Lohman, D. J., & Meier, R. (2011). SequenceMatrix: concatenation software for the fast assembly of multi‐gene datasets with character set and codon information. *Cladistics*, 27(2), 171-180.

Wetterer, J. K., Ward, P. S., Weiterer, A. L., Longino, J. T., Träger, J. C., & Miller, S. E. (2000). Ants (Hymenoptera: Formiddae) of Santa Cruz Island, California. *Bull. Southern California Acad.* Sei, 99(1), 25-31.