UNIVERSITY OF CALIFORNIA SAN DIEGO

An analysis of mitochondrial DNA variation in *Aphaenogaster patruelis:* an island endemic ant species

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

in

Biology

by

Bryan Huang

Committee in charge:

Professor David A. Holway, Chair

Professor Joshua R. Kohn

Professor Katherine L. Petrie

2020

The thesis of Bryan Huang is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

Chair

University of California San Diego

2020

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to thank my committee chair and members Dr. David Holway, Dr. Joshua Kohn, and Dr. Katherine Petrie for giving me the opportunity to conduct this research. I would also like to thank PhD candidate Ida Naughton for her support throughout the project by providing materials and technical help. Finally, I would like to thank Charles Chiang for introducing me to the Holway lab and inspiring me to pursue this study.

ABSTRACT OF THE THESIS

An analysis of mitochondrial DNA variation in *Aphaenogaster patruelis:* an island endemic ant species

by

Bryan Huang

Master of Science in Biology

University of California San Diego, 2020

Professor David A. Holway, Chair

This thesis examines mitochondrial DNA variation exhibited by *Aphaenogaster patruelis*, an ant species that is confined to five oceanic islands off the coast of southern California and Baja California. The aims of this study are to assess inter-island relationships and to determine if there is cytonuclear discordance. *Aphaenogaster patruelis* is a putative relict known only from San Clemente Island, Santa Catalina Island, San Nicolas Island, Santa Barbara Island, and Isla Guadalupe. Samples were collected from each of the islands, and DNA extraction was performed to sequence the 12S and CO1 mitochondrial genes. Based on these sequence data, I constructed

median-joining haplotype networks and a CO1 phylogenetic tree. Samples on three islands were represented by a single haplotype, but samples on two islands were represented by two, distinct haplotypes. The presence of two, distinct haplotypes suggests multiple colonization events on these islands and indicates the presence of cytonuclear discordance as the split is not evident in previous analyses based on nuclear DNA. Given the lack of an extant mainland population of *Aphaenogaster patruelis*, reconstructing evolutionary relationships among island populations is not possible, but our results do suggest repeated colonization on individual islands, either through movements between islands or between islands and a now extinct mainland population.

Introduction

Situated off the coast of Southern California, the California Channel Islands are an archipelago of islands that range from 2.6 to 250 square kilometers in size (Rick et al. 2014). The California Channel Islands are located 20 to 98 kilometers off the Southern California coast and are often the focus of studies of island ecology and biodiversity due to the wide array of endemic life found among them (Rick et al. 2014). Differences in size and location of the islands make them a prime study system for observing the development of various endemic species, while being relatively isolated from the mainland. This isolation fosters the development of differentiation among populations of a wide variety of organisms, leading to questions about the evolutionary history of various species found on these islands. One such species that is prevalent among the Southern California Channel Islands is the ant species *Aphaenogaster patruelis*.

Containing over 200 different species, the genus *Aphaenogaster* is both diverse and widely distributed with species in the following biogeographic regions: Afrotropical, Australasia, Indomalaya, Malagasy, Neartic, Neotropical, Oceania, and Paleartic regions (Antweb.org). *Aphaenogaster patruelis* is an interesting species in that it is a putative relict, known only to occur on the Southern Channel Islands (Santa Barbara, Santa Clemente, Santa Catalina, and San Nicholas) and Isla Guadalupe. As *A. patruelis* does not have an extant mainland population, it would not be possible to completely reconstruct the evolutionary relationships of the populations on the islands; however, phylogenetic analysis can still reveal information about potential colonization patterns.

Phylogenetic reconstruction can involve the use of a variety of data sets, including nuclear and mitochondrial DNA. Studies using mitochondrial DNA are relatively common for species-level identification and provide a simple method of identification using DNA barcoding

(Meusnier, et al. 2008). In addition, this method of using mitochondrial genomes for comparative analyses in the order Hymenoptera has yielded tangible results (Song, et al., 2016). Building upon a previous analysis based on nuclear DNA (Chiang 2018), this study employs the use of mitochondrial gene analysis to examine evolutionary relationships among island populations of *A. patruelis*. By assessing the relationships among different populations using mitochondrial genes, it is possible to gain perspective on patterns of colonization by this island endemic. Mitochondrial genes can exhibit an accelerated rate of mutation, making them a useful tool to visualize differences among populations of the same species (Yang et. al, 2014). In addition, recombination in the mtDNA genome occurs rarely, making it a feasible candidate for analysis as mutations in the genome would be easier to attribute to evolution (Petráková et al. 2017).

As mitochondria have their own separate genomes and are inherited maternally, the evolutionary history suggested by these genes may not always agree with the history provided by nuclear DNA, leading to the potential of cytonuclear discordance, which is commonly observed in phylogenetic studies (Lee-Yaw, et al. 2018). Despite the potential for cytonuclear discordance, the data provided by mitochondrial gene analysis remain crucial to understanding patterns of colonization as they provide an additional view on the evolutionary history of a species. Complex biological processes that contribute to cytonuclear discordance may not be noticed if only analysis on nuclear genes was done (Ivanov, et al. 2018). Addressing the existence of cytonuclear discordance becomes a key aspect in assessing a phylogenetic analysis as it may suggest routes of colonization.

To establish relationships among populations using mitochondrial analysis and to investigate the potential of cytonuclear discordance, the 12S and CO1 genes were selected. The CO1 gene is widely regarded as the standard DNA barcode, making it a priority for investigating

mitochondrial genetics; however, the 12S gene is also noted to be a viable alternative if CO1 gene sequencing is unreliable (Lv, et al. 2014). Including both in the analysis will provide additional information on the results. Haplotype network analysis was selected as the primary method of discerning relationships among the five island populations of *A. patruelis*. Haplotype networks are a web of haplotypes that provide a view of population genetics in addition to addressing the differences in haplotypes found, which would ultimately aid in determining relationships among samples (Leigh and Bryant, 2015). Median-joining networks were used as they build upon minimum-spanning tree methods and incorporate potential extant sequences that are not present in the sample pool or extinct ancestral sequences (Bandelt, et al. 1999). These potential sequences appear as black circles within the median-joining networks and are known as "median vectors," which are based upon a maximum-parsimony heuristic algorithm and can provide additional information on potential extinct populations that may have been involved in colonization patterns (Bandelt, et al. 1999).

To put together these haplotype networks, *Aphaenogaster patruelis* samples were collected from multiple locations on all five islands on which this species is known. DNA was then extracted from all samples and sequenced to obtain the data for the CO1 and 12S genes used in the study. The gene sequences were then processed into a 12S haplotype network (Figure 1), CO1 haplotype network (Figure 2), concatenated CO1-12S network (Figure 3), and a CO1 phylogenetic tree (Figure 4), for discussion.

Methods

Aphaenogaster patruelis samples were collected into ethanol on all five islands on which the species is known to occur: Isla Guadalupe, San Clemente Island, Santa Catalina Island, San Nicolas Island, and Santa Barbara Island. Samples from Isla Guadalupe were labeled with the letter G, San Clemente Island with C, Santa Catalina Island with X and L, San Nicolas Island with N, and Santa Barbara Island with B. Prior to DNA extraction, specimens were air dried for 30 min and I used a sterilized razor blade to remove gasters to prevent contamination of samples with gut contents. I extracted total genomic DNA using a DNeasy Blood and Tissue Kit (Qiagen Inc. Valencia, CA) following the kit instructions and an additional Rnase step. I used a Benchmark Scientific D1030 Bead Bug Homogenizer to rupture ant tissue and to incubate tissue lysate for 24 hours. I quantified genomic DNA for each sample using a Qubit 4.0 fluorometer, then used a Diagenode Bioruptor to shear genomic DNA into fragments 400 – 1000 bp in length. I then prepared DNA libraries for each sample as follows. I added a custom Sera-Mag Magnetic SpeedBeads and PEG mixture at 2X volume to the sheared samples and eluted DNA from the beads to remove small fragments from each sample). I then used Roche Diagnostic DNA Hyper Prep kits for blunt end repair and A-tailing followed by another SeraMag bead clean at 1.6X volume. I amplified each library with unique Illumina TruSeq style i5 and i7 adapters, using KAPA HiFi Hot Start polymerase and 12 PCR cycles. Indexed libraries were purified with a final bead clean at 1.6X volume.

Indexed libraries were used as template for sanger sequencing of mtDNA for most samples, but total genomic DNA was used as template for a small subset of samples (template type for each sample is listed in Table 2). For each sanger sequencing reaction, I added 2 uL of DNA template and the primers and annealing temperatures for each reaction are listed in Table 1.

Our amplification protocol included 2 min at 95 C, followed by 35 cycles of 95C for 15 sec, annealing temperature for 30 sec, and 72 C for 1 min, then a hold at 72C for 7 min after cycling. I visualized samples on a gel following PCR amplification, and sent samples to Eton Bioscience for PCR purification and sequencing.

Raw data for the CO1 and 12S genes were collected in '.abi' files and were imported into MEGAX. Poor quality base-calls were trimmed from each of the sample files manually by examining their respective chromatograms and the files were exported as FASTA files. Forward and reverse sequences of each sample were then incorporated into a consensus sequence post-trimming through BioEdit. For samples that were unable to be properly matched, the sequence with the highest quality base-calls were used after re-checking the chromatograms. Discrepancies in the base-calls were manually fixed by checking with the forward and reverse sequence chromatograms. Consensus sequences were then imported into CO1 and 12S multiple-alignment files respectively through BioEdit. The multiple-alignment files were then aligned using the Clustal-W plug-in and exported as FASTA files.

The multiple-alignment FASTA files were then imported into Mesquite 3.61 and trimmed further so that the lengths of the sequences within each multiple-alignment file were identical. Blank data was then filled in with a "?" in order to complete processing of the files for use in creating haplotype networks. The processed files were then exported as NEXUS files.

In order to process the files in PopArt to generate haplotype networks for analysis, a traits file is needed in conjunction with the NEXUS file. Two traits files were created through Text Edit to correspond with the two multiple-alignment NEXUS files and the files were uploaded into PopArt where a median-joining network analysis was done. This procedure then generated the 12S and CO1 haplotype networks presented in the results.

Samples that included both a 12S and CO1 sequence were then selected and used to create a concatenated haplotype network for further analysis. Multiple-alignment files of the 12S and CO1 sequences of the samples were created respectively using the processed sequences from the 12S and CO1 multiple-alignment NEXUS files created previously. These files were then imported into Mesquite 3.61 and concatenated. A new traits file was then created for the concatenated multiple alignment file and a haplotype network was generated using PopArt with the median-joining method as well. 18 samples were used in Fig 1, 25 samples were used in Fig 2, and 15 samples were used in Fig 3.

To create the CO1 phylogenetic tree, I used the CO1 multiple-alignment FASTA file. Outgroup sequences were selected from *A. mutica* samples and added into the file with the AM1 and AM2 labels after being processed in the same manner as the *A. patruelis* samples. Alignment using the Clustal-W plug in was performed again, and the file was exported as a separate FASTA file. This file was then imported into Mesquite 3.61 and exported as a Phylip file for analysis. The file was then uploaded into the CIPRES gateway for RAxML analysis. RAxML-HPC BlackBox was selected as the tool, a bootstrap value was set to 1000, and the outgroups were selected. The results file was then downloaded and opened in FigTree v1.4.3 where the CO1 tree was visualized and modified to be easier to read.

Results

Figure 1 provides a view of the relationships of *Aphaenogaster patruelis* based on the 12S gene. Two distinct haplotypes were detected on samples both from Santa Catalina Island and from San Clemente Island, whereas samples from the other three islands just had one haplotype per island. Santa Catalina and San Clemente Islands also jointly shared a haplotype, indicating a close relationship. Another multi-island haplotype was also evident in the set of samples from Santa Catalina, San Nicolas, and Santa Barbara Islands.

Figure 2 provides a view of the relationships between *Aphaenogaster patruelis* samples based on the CO1 gene. Samples from Santa Barbara and San Nicolas Islands each appear to have multiple closely related haplotypes, while Isla Guadalupe appears to have two distant haplotypes. Two distinct haplotype groups for Santa Catalina Island as well as two distinct haplotypes groups for San Clemente Island are also apparent in the network. San Clemente Island sample C5 is separate from the rest of the San Clemente Island cluster, and there is one Santa Barbara Island sample that is separate from the Santa Barbara Island cluster as well. The existence of the two distinct haplotype groups for samples from Santa Catalina Island and San Clemente Island are consistent with Figure 1 as well.

Figure 3 illustrates the relation among *Aphaenogaster patruelis* samples by using a concatenated network created from combining the data from 12S and CO1 genes. Only 15 samples in the study had both a 12S and CO1 gene sequenced, so the network has a smaller sample size than the Figure 1 and Figure 2. Samples from Santa Barbara and San Nicolas Islands have multiple closely related haplotypes in the network. Isla Guadalupe only has one haplotype as sample G3 was omitted from this network due to the lack of a 12S sequence. Two distinctly separated haplotype groups for San Clemente Island and two distinctly separated groups

haplotypes for Santa Catalina Island are visible in this network as well, confirming the findings in Figs. 1-2.

Figure 4 highlights the existence of a paraphyletic San Clemente Island cluster, along with a paraphyletic Santa Catalina Island cluster, with the exception of a C5 San Clemente Island sample. The X1 Santa Catalina Island sample that is not grouped with the rest of the Santa Catalina Island samples also appears to be closely related to the samples from Santa Barbara and San Nicolas Islands. These differences show a distinct split occurring in both the Santa Catalina Island and San Clemente Island samples; this split was also evident from the haplotype networks discussed previously.

Discussion

Based on the findings from Figs. 1-4, it becomes evident that San Clemente Island and Santa Catalina Island differ from the other islands in that they have two distinct haplotypes instead of one. Specifically, the C5 San Clemente Island sample appears separate from the cluster of San Clemente Island samples, and the L1 and X10 Santa Catalina Island samples appear different from the other Santa Catalina Island cluster. The presence of two haplotypes on one island can be attributed to the colonization over time by two differentiated populations, which could occur by colonization from an extinct mainland population or from another island. In addition, the ability of ants to disperse throughout large geographic areas via flight supports this idea of island-hopping and mainland-to-island colonization as they could ride wind currents to reach the farther islands (Helms IV, 2017). As *Aphaenogaster patruelis* no longer exists on the mainland, it is possible that two populations that previously existed may have been responsible for distinct haplotype split.

By assigning two theoretical differentiated extinct mainland populations as populations A and B, there is potential for multiple mainland-to-island colonization events to occur. Population A would be responsible for colonizing San Clemente Island and another island (potentially Santa Barbara Island), where island-hopping events would lead it to colonize Santa Catalina Island. Population B would then be responsible for colonizing San Clemente and Santa Catalina Islands from the mainland. The potential for multiple colonization events from the mainland remains a plausible hypothesis that is supported by evidence from other arthropods as well. A study on four beetle species (*Thinopinus pictus, Hadrotes crassus, Hypocaccus lucidulus, and Nyctoporis carinata*) that occur on the mainland and multiple California Channel Islands provide evidence of potential for multiple colonization events from mainland-to-island as well (Caterino,

Chatzimanolis, and Richmond, 2015). This proposed route of colonization; however, is based upon the CO1 phylogenetic tree (Figure 4) and does not rule out the possibility of alternate methods of colonization.

Previous research on *Aphaenogaster patruelis* phylogeny using Ultraconserved Elements (nuclear DNA) indicated that each of the islands had strongly supported clades that were monophyletic, with the exception of the Santa Barbara Island clade, which appeared paraphyletic to the San Nicholas Island clade (Chiang, 2018). Based on these findings, it was speculated that the colonization of the California Channel Islands by *A. patruelis* occurred with a colonization event from a mainland population to Santa Catalina Island (Chiang, 2018). It was then proposed that subsequent colonization of the other islands occurred from this stepping point due to the monophyletic nature of the clades (Chiang, 2018). This monophyletic grouping proposed by nuclear DNA is inconsistent with the paraphyletic San Clemente Island and Santa Catalina Island groups, seen in Figure 4. If the mitochondrial data had presented monophyletic grouping of the samples by island, it would be considered concordant. In short, the pattern of colonization suggested by nuclear DNA does not directly address the split in haplotype grouping of Santa Catalina Island and San Clemente Island, indicating the presence of cytonuclear discordance.

Instances of cytonuclear discordance are common. In populations of the black-tailed brush lizard (*Urosaurus nigricaudus*) of Baja California, for example, the mitochondrial genealogy suggests less differentiation as opposed to the nuclear genealogy (Lindell, Méndez-de La Cruz, and Murphy, 2008). Despite this pattern, there are still many unknowns about what drives discordance within the *A. patruelis* populations. As explained in a study on the North American Tree Squirrels (*Tamiasciurus*), two potential causes of cytonuclear discordance could arise from incomplete lineage sorting, or from hybridization (Chavez, et al, 2014). Further

studies on cytonuclear discordance conducted by Toews and Brelsford (2012) also indicated that adaptive introgression of mitochondrial DNA, disparities in demographics, and sex-based asymmetries are possibilities that may account for the differences. In short, more data are required before identifying the most plausible factor behind this cytonuclear discordance in the populations of *A. patruelis*. Although the cause of cytonuclear discordance remains to be answered, the data provided by the mitochondrial analysis still provide information on the relationships between the populations of *A. patruelis*.

Furthermore, as haplotype networks are unrooted, the networks do not account for time (Kong, et al. 2015). Data from Figs 1-3 essentially provide information about the relationships between the populations of *A. patruelis* found on the islands based on "relatedness," or the differences between the haplotypes by the number of mutations. Further studies incorporating additional mitochondrial genes or a greater sample size may provide more evidence to revise or support this proposed route of colonization. Although the results of the study appear consistent with the proposal that two extinct mainland populations were responsible for the colonization of the Southern California Channel Islands and Isla Guadalupe, it is impossible to rule out different patterns of colonization as the lack of an extant mainland population makes rooting the samples through a phylogenetic tree a challenge. More research on the influence of geographic events and a larger sample size would be needed to contribute to developing a more concrete pattern of colonization; however, additional research on the factors behind cytonuclear discordance would help reveal any intricacies of the way in which *A. patruelis* colonized the islands.

Tables

Table 1. List of primers used for each target gene. Sequences and Annealing Temperatures included

Target Gene	Primer Name	Sequence	Annealing Temperature
CO1-Forward	LCO-1490	5’ - GGT CAA CAA ATC ATA AAG ATA TTG G	51C
CO1-Reverse	HCO-2198	5’ - TAA ACT TCA GGG TGA CCA AAA AAT CA	51C
12S-Forward	12SAI	5’ - CTA GGA TTA GAT ACC CTA TT	49C
12S-Reverse	12SBI	5’- AAG AGC GAC GGG GCG	49C

Table 2. List of samples used in Figures with sequencing source preparation method. Asterisk '*' denotes samples that were re-sequenced to check accuracy.

Sample	CO1 Network	12S Network	Concatenated Network	Unenriched Library CO1	Total Genomic DNA CO1	Unenriched Library 12S	Total Genomic Data 12S
B1		x		x			x
B2	x	x	x	x		x	
B3	x			x		x	
B4	x	x	x		x		x
B5	x			x		x	
B6	x			x		x	
B7	x	x	x	x		x	
C1	x	x	x	x		x	
C2	x	x	x		x	x	
C4	x			x		x	
C5	x	x	x	*	x	*	x
C10	x	x	x		x		x
C11	x				x	x	
C14		x					x
N2	x	x	x	x		x	
N5		x		x		x	
N7	x	x	x	x		x	
N8	x	x	x	x		x	
N9	x	x	x	x		x	
N10	x			x		x	
N12	x			x		x	
X1	x	x	x	x		x	
X4	x			x		x	
X5	x			x		x	
X10	x	x	x		x		x
L1	x	x	x		x		x
G2	x	x	x	x		x	
G3	x			x		x	

Table 3. List of samples used in project with abbreviations, locality, collection dates, and name of collector.

Label	Species	State	County	Island	Location	Coordinates	Date (MM/DD/YYYY)	Collector
Am1	*A. mutica*	Baja California	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	X	X	1km W La Laguna	(23°33'0.00"N,110° 0'0.00"W, 1800m)	12/30/2003	Philip S. Ward
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
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
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
C14	*A. patruelis*	California	Los Angeles Co.	San Clemente	Horton Canyon	N 32.90560 W 118.47855	5/22/2016	Ida Naughton
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
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
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
N2	*A. patruelis*	California	Ventura Co.	San Nicolas	X	X	X	X
N5	*A. patruelis*	California	Ventura Co.	San Nicolas	Grand Canyon	(33°14'5.50"N,119°31'18.40"W)	9/14/2017	Ida Naughton
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
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
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
X10	*A. patruelis*	California	Los Angeles Co.	Santa Catalina	Middle Canyon, Johnso	x	6/8/2015	David A. Holway
L1	*A. patruelis*	California	Los Angeles Co.	Santa Catalina	Fox Project Trap 441	x	8/10/2004	x

Figures

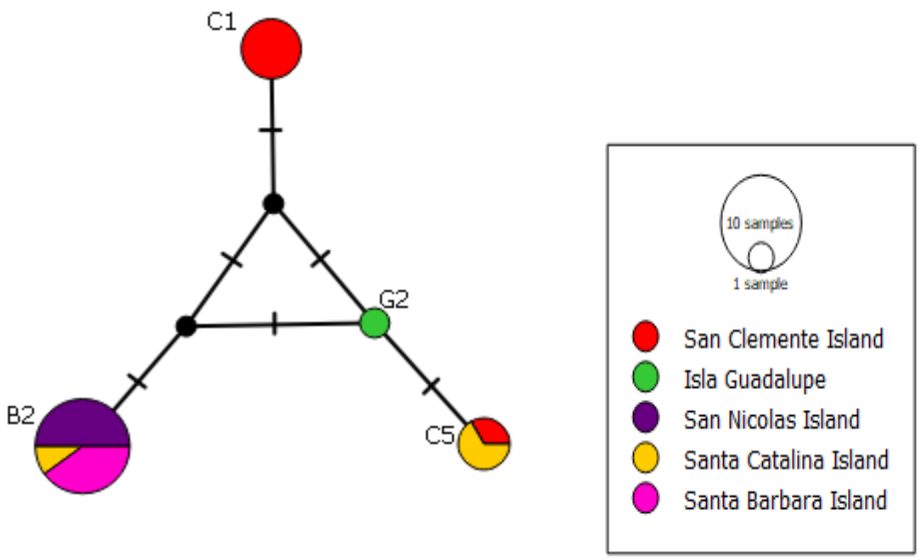

Figure 1. Median-joining haplotype network of the 12S Mitochondrial Gene of *Aphaenogaster patruelis*. 18 samples were used in the development of this network. Black dots indicate hypothetical haplotypes and hashes indicate mutations. The size of circle is proportional to the number of samples.

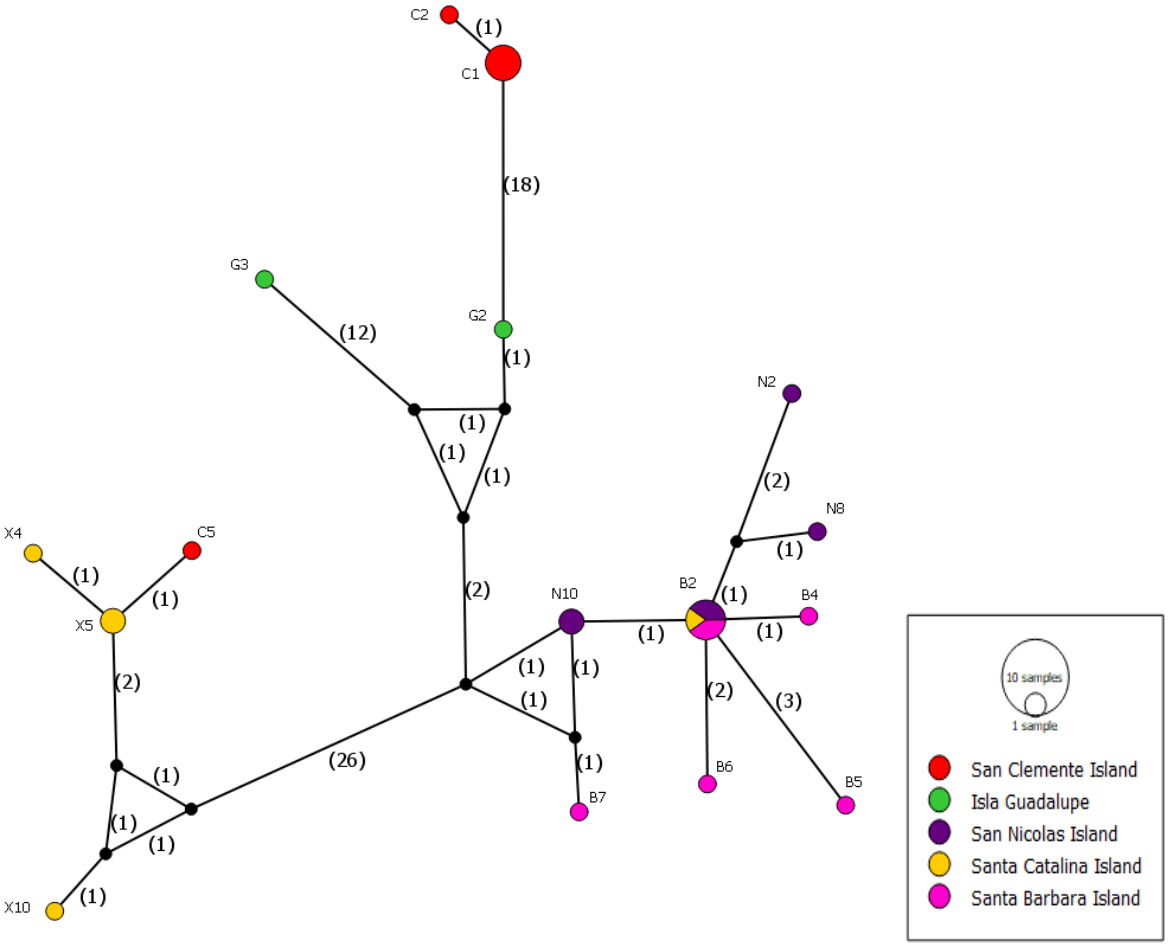

Figure 2. Median-joining haplotype network of CO1 mitochondrial gene of *Aphaenogaster patruelis*. 25 samples were used in the development of this network. Black dots indicate hypothetical haplotypes and the numbers in parentheses indicate the number of mutations. The size of circles is proportional to the number of samples.

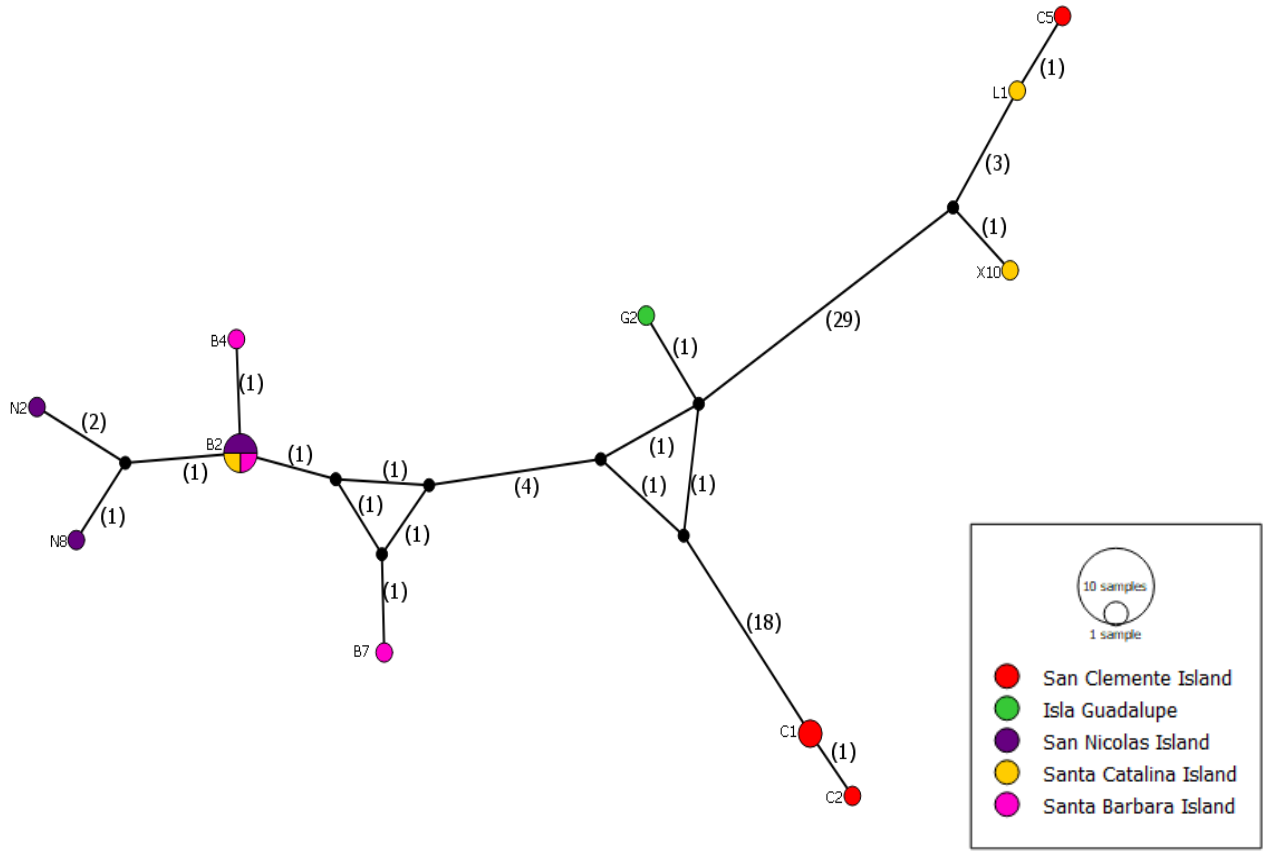

Figure 3. Median-joining haplotype network of concatenated 12S and CO1 genes of *Aphaenogaster patruelis*. 15 samples were used in the development of this network. Black dots indicate hypothetical haplotypes and the numbers in parenthesis indicate the number of mutations. The size of circles is proportional to the number of samples.

Figure 4. Phylogenetic CO1 Tree of *Aphaenogaster patruelis*. Created with RAxML-HPC BlackBox using the CIPRES gateway and visualized with FigTree v1.4.3. Outgroups were selected among *Aphaenogaster mutica* samples and consist of AM1 and AM2. A bootstrap value of 1000 was used. Values assigned reflect the number of times the pattern was determined out of a 100 with higher numbers reflecting higher confidence.

References

AntWeb. Available from: https://www.antweb.org/description.do?subfamily=myrmicinae&genus=aphaenogaster&rank=genus. Accessed 18 November 2020

AntWeb. Available from: https://www.antweb.org/description.do?genus=aphaenogaster&species=patruelis&rank=species. Accessed 18 November 2020

Bandelt, H. J., Forster, P., & Rohl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution, 16*(1), 37-48. doi:10.1093/oxfordjournals.molbev.a026036

Bell, R. C., Drewes, R. C., & Zamudio, K. R. (2015, February 26). Reed frog diversification in the Gulf of Guinea: Overseas dispersal, the progression rule, and in situ speciation. *Evolution, 69*(4), 904-915. doi:10.1111/evo.12623

Caterino, M. S., Chatzimanolis, S., & Richmond, M. P. (2015). On the Origins of the Insect Fauna of California's Channel Islands: A Comparative Phylogeographic Study of Island Beetles. *Monographs of the Western North American Naturalist, 7*, 276-296. doi:10.3398/042.007.0121

Chatzimanolis, S., Norris, L. A., & Caterino, M. S. (2010, September 1). 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). doi:10.1603/AN10037

Chavez, A. S., Maher, S. P., Arbogast, B. S., & Kenagy, G. J. (2014). Diversification and Gene Flow in Nascent Lineages of Island and Mainland North American Tree Squirrels (Tamiasciurus). *Evolution, 68*(4), 1094-1109. doi:10.1111/evo.12336

Chiang, B. (2018). High Throughput Sequencing of Ultraconserved Elements Reveals Inter-island Relationships of a California Channel Island Endemic Ant Species (Aphaenogaster patruelis). UC San Diego.

Edgington, H. A., Ingram, C. M., & Taylor, D. R. (2016). Cyto-nuclear discordance suggests complex evolutionary history in the cave-dwelling salamander, E urycea lucifuga. *Ecology and Evolution, 6*(17), 6121-6138. doi:10.1002/ece3.2212

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

Hakala, S. M., Seppä, P. S., & Helanterä, H. (2019, March 27). Evolution of dispersal in ants (Hymenoptera: Formicidae): A review on the dispersal strategies of sessile superorganisms. *Myrmecological News, 29*.

Hall, T. A. (1999) BioEdit: A User-Friendly Biological Sequence Alignment Editor and Analysis Program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, 41, 95-98.

Helms, J. A., IV. (2017, December 19). The flight ecology of ants (Hymenoptera: Formicidae). *Myrmecological News, 26*, 19-30.

Hunter, M. E., Johnson, N. A., Smith, B. J., Davis, M. C., Butterfield, J. S., Snow, R. W., & Hart, K. M. (2018). Cytonuclear discordance in the Florida Everglades invasive Burmese python (Python bivittatus) population reveals possible hybridization with the Indian python (P. molurus). *Ecology and Evolution, 8*(17), 9034-9047. doi:10.1002/ece3.4423

Ivanov, V., Lee, K. M., & Mutanen, M. (2018, April). Mitonuclear discordance in wolf spiders: Genomic evidence for species integrity and introgression. *Molecular Ecology, 27*(7), 1681-1695. doi:10.1111/mec.14564

Kong, S., Sánchez-Pacheco, S. J., & Murphy, R. W. (2015). On the use of median-joining networks in evolutionary biology. *Cladistics, 32*(6), 691-699. doi:10.1111/cla.12147

Lee-Yaw, J. A., Grassa, C. J., Joly, S., Andrew, R. L., & Rieseberg, L. H. (2018, August 23). An evaluation of alternative explanations for widespread cytonuclear discordance in annual sunflowers (Helianthus). *New Phytologist, 221*(1), 515-526. doi:10.1111/nph.15386

Leigh, J.W. and Bryant, D. (2015), PopART: full-feature software for haplotype network construction. *Methods in Ecology and Evolution*, *6*, 1110-1116. doi:10.1111/2041-210X.12410

Levin, L., Blumberg, A., Barshad, G., & Mishmar, D. (2014, December 23). Mito-nuclear co-evolution: The positive and negative sides of functional ancient mutations. *Frontiers in Genetics*. doi:10.3389/fgene.2014.00448

Lindell, J., Cruz, F. R., & Murphy, R. W. (2008). Deep biogeographical history and cytonuclear discordance in the black-tailed brush lizard (Urosaurus nigricaudus) of Baja California. *Biological Journal of the Linnean Society*, *94*(1), 89-104. doi:10.1111/j.1095-8312.2008.00976.x

Lv, J., Wu, S., Zhang, Y., Chen, Y., Feng, C., Yuan, X., Jia G., Deng J., Wang C., Wang Q., Mei L., Lin, X. (2014). Assessment of four DNA fragments (COI, 16S rDNA, ITS2, 12S rDNA) for species identification of the Ixodida popart(Acari: Ixodida). *Parasites & Vectors*. doi:10.1186/1756-3305-7-93

Maddison, W. P. and D.R. Maddison. (2019). Mesquite: a modular system for evolutionary analysis. Version 3.61 http://www.mesquiteproject.org

Meusnier, I., Singer, G. A., Landry, J., Hickey, D. A., Hebert, P. D., & Hajibabaei, M. (2008, May 12). A universal DNA mini-barcode for biodiversity analysis. *BMC Genomics*, *9*. doi:10.1186/1471-2164-9-214

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

Petráková, L., Tóthová, A. and Schlaghamerský, J. (2017), Phylogeography of the rare velvety tree ant Liometopum microcephalum (Formicidae: Dolichoderinae). *Journal of Biogeography, 44*, 1652-1664. doi:10.1111/jbi.12960

Phuong, M. A., Bi, K., & Moritz, C. (2017). Range instability leads to cytonuclear discordance in a morphologically cryptic ground squirrel species complex. *Molecular Ecology, 26*(18), 4743-4755. doi:10.1111/mec.14238

Rick, T. C., Sillett, T. S., Ghalambor, C. K., Hofman, C. A., Ralls, K., Anderson, R. S., Boser C. L.,Braje T. J., Cayan D. R., Chesser R. T., Collins P. W., Erlandson J. M., Faulkner K. R., Fleischer R., Funk W. C., Galipeau R., Huston A., King J., Laughrin L., Maldonado J., McEachern K., Muhs D. R., Newsome S. D., Reeder-Myers L., Still C., Morrison, S. A. (2014, August). Ecological Change on California's Channel Islands from the Pleistocene to the Anthropocene. *BioScience, 64*, 680-692. doi:10.1093/biosci/biu094

Rubinoff, D., & Holland, B. S. (2005, December). Between Two Extremes: Mitochondrial DNA is neither the Panacea nor the Nemesis of Phylogenetic and Taxonomic Inference. *Systematic Biology, 54*(6), 952-961. doi:10.1080/10635150500234674

Song, N., Li, X., Yin, X., Li, X., Yin, J., & Pan, P. (2019). The mitochondrial genomes of palaeopteran insects and insights into the early insect relationships. *Scientific Reports, 9*(1). doi:10.1038/s41598-019-54391-9

Song, S., Tang, P., Wei, S., & Chen, X. (2016, February 16). Comparative and phylogenetic analysis of the mitochondrial genomes in basal hymenopterans. *Scientific Reports, 6*. doi:10.1038/srep20972

Stamatakis, A. (2014) RAxML Version 8: A tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies. *Bioinformatics 10.* doi:10.1093/bioinformatics/btu033

Sudhir Kumar, Glen Stecher, Michael Li, Christina Knyaz, and Koichiro Tamura (2018) MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Molecular Biology and Evolution 35*, 1547-1549

Toews, D. P., & Brelsford, A. (2012). The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology, 21*(16), 3907-3930. doi:10.1111/j.1365-294x.2012.05664.x

Tseng, S., Wetterer, J. K., Suarez, A. V., Lee, C., Yoshimura, T., Shoemaker, D., and Yang, C. S. (2019, September 17). Genetic Diversity and Wolbachia Infection Patterns in a Globally Distributed Invasive Ant. *Frontiers in Genetics, 10*. doi:10.3389/fgene.2019.00838

Yang, L., Tan, Z., Wang, D., Xue, L., Guan, M., Huang, T., and Li, R. (2014). Species identification through mitochondrial rRNA genetic analysis. *Scientific Reports, 4*. doi:10.1038/srep04089