DePaul University
UNIVERSITY LIBRARIES

DePaul University
Digital Commons@DePaul

College of Science and Health Theses and Dissertations

College of Science and Health

Winter 3-22-2019

Phylogeography of the Neotropical Fish Genus Rhoadsia (Teleostei: Characidae) in Ecuador

Roberto Valentino Cucalón Tamayo
DePaul University, rcucalon@mail.depaul.edu

Follow this and additional works at: https://via.library.depaul.edu/csh_etd

 Part of the Biology Commons

Recommended Citation

Cucalón Tamayo, Roberto Valentino, "Phylogeography of the Neotropical Fish Genus Rhoadsia (Teleostei: Characidae) in Ecuador" (2019). *College of Science and Health Theses and Dissertations*. 310.
https://via.library.depaul.edu/csh_etd/310

This Thesis is brought to you for free and open access by the College of Science and Health at Digital Commons@DePaul. It has been accepted for inclusion in College of Science and Health Theses and Dissertations by an authorized administrator of Digital Commons@DePaul. For more information, please contact digitalservices@depaul.edu.

Phylogeography of the Neotropical Fish Genus *Rhoadsia*

(Teleostei: Characidae) in Ecuador

A thesis presented in

partial fulfillment of the

requirement for the degree of

Master of Science

By

Roberto Valentino Cucalón Tamayo

March 2019

Advisor: Dr. Windsor E. Aguirre

Department of Biological Sciences

College of Science and Health

DePaul University

Chicago, Illinois

Table of Contents

Acknowledgments

I want to thank my Advisor Dr. Windsor Aguirre for the numerous opportunities that he opened to me since the first moment I met him. He has been an important support in my academic and professional life. I appreciate the time he always found to help me when I was struggling during my research. I could not be more grateful for having the luck of crossing his path having the opportunity to learn from his experiences.

Thanks to Dr. Jalene LaMontagne and Dr. Caleb McMahan for accepting being part of my thesis committee. Also, to Pedro Jimenez-Prado, Ronald Navarrete-Amaya, Jonathan Valdiviezo-Rivera, Antonio Torres, and his students, and my good friend Thomas Borders for helping me collect the specimens from different rivers in Ecuador. Thanks to Lenin Maldonado who provided technical assistance in the field. Special thanks to Jose Lara (president of agritourism in Cerro de Hayas) and Carlos Jara (president of the community 23 de Noviembre) who allowed the collection of samples from their community.

Thanks to Nathan Lujan from the Royal Ontario Museum (ROM) and the Smithsonian Tropical Research Institute in Panama for providing some of the samples used in this study. Gian Carlo Sánchez assisted with collecting the ROM samples. Thanks to Natasha Weirzal and Kirby Karpan who helped with the amplification of the mitochondrial genes from many of the samples collected.

Thanks to Dr. Silliker, Dr. Sparkes, and Dr. Brooke who contributed to the progress of my reading, writing, and oral skills during the development of my proposal and thesis. I want to thank all my professors for their support, Dr. Rajah, Dr. Dean, and Dr. Bell.

I would like to thank the members of the Aguirre lab, Joe Forberg, Sara Alharbi, Jessica Vaclav, and Franklin Joaquin, as well as my peers from the cohort 2018 for their unconditional support during my stay at the program.

I want to thank DePaul University and the Department of Biological Science for the opportunity to study here and pursue my master's degree. Also, the Graduate Research Funding program for the financial support that helped me to complete my thesis research.

I want to dedicate this work to my beloved family, my parents Valentin Cucalón and Mariana Tamayo and my wife Pierina García who always have supported me in the fulfillment of my goals.

Abstract

Western Ecuador is considered a biodiversity hotspot. Nevertheless, studies of population genetic structure and variation are rare, especially in aquatic species. The genus *Rhoadsia* is an endemic freshwater fish in this region with two recognized species, *Rhoadsia minor* and *R. altipinna*. Little is known about the evolutionary relationships of their populations, and due to morphological similarities, their validity as distinct species has been questioned. The present study uses a phylogeographic approach to examine the evolutionary history of the genus and the validity of the two described species. Furthermore, I investigated the possible geographical origin of the genus based on patterns of genetic diversity and genetic distance from two sister genera. I also estimated potential zoogeographic breaches based on patterns of genetic divergence that could reflect points of genetic divergence between populations of other freshwater fishes. I used several molecular markers with different rates of evolution: cytochrome oxidase I (*COI*) and cytochrome b (*Cyt-b*) from the mitochondrial genome (mtDNA) and 12 microsatellite loci from the nuclear genome (nDNA). Sampling was conducted in drainages across Western Ecuador and an effort was made to sample at different elevations (from 30 to 1260 m above sea level), given the known patterns of morphological divergence associated with elevation. The phylogenetic tree resolved with the mtDNA data confirmed the presence of two species exhibiting genetic introgression at the border between the species ranges (Northern Guayas). A Bayesian-based analysis of the microsatellite data revealed the existence of ten populations of *Rhoadsia* divided into three main groups (I, II, and III). Group I coincided with the distribution of *R. minor* while group II and III seemed to represent geographic subgroups of *R. altipinna*. Patterns of genetic divergence and diversity were used to suggest potential evolutionarily significant units within both species for conservation efforts. The most likely geographical origin of the genus appeared to be in the Guayas drainage. However, this result is a hypothesis and should be examined more carefully in the future using other approaches. Two major zoogeographic breaches were identified: the first one was between the two largest drainages (Esmeraldas and Guayas) and the

second one was located farther south cutting through the Jubones drainage and isolating the central Guayas and drainages just south from the southernmost drainages of western Ecuador. Future research should include biological and ecological data to reinforce the validity of the species. I also suggest looking for signals of adaptive divergence between species or between populations from contrasting habitats like low and high elevations. This study provides a baseline for future studies examining biogeographic relationships of freshwater species in Western Ecuador.

CHAPTER I

Review of the Literature

The following literature review will provide some general knowledge about the biodiversity in Western Ecuador, factors that have threatened it, and the importance of using a phylogeographic approach to mitigate the lack of knowledge of evolutionary relationships of the species in the region. First, I will explore our current understanding of the number of species on earth, the problems they face due to the human activities, and the issue of how and where we should concentrate our efforts as investigators to efficiently contribute to the knowledge of the diversity of life. To this end, I will discuss the high levels of endemism in Western Ecuador with emphasis on its aquatic species, and discuss current knowledge of the freshwater fish genus *Rhoadsia*. I will give some background on the hydrographic systems of Western Ecuador, in which populations of this genus have evolved. Finally, I will talk about the use of phylogeographic methods that employ molecular markers with different rates of evolution to better understand the evolutionary relationships of populations of *Rhoadsia* in Western Ecuador.

The Biodiversity Crisis

Biodiversity can be defined as the variety of life (Manokaran, 1992). This concept not only considers the number of species that inhabit an area (interspecific diversity) but also variation within or among populations of species (intraspecific diversity). One of the biggest aims in biology is to know the number of species on Earth, and although with great uncertainty, there have been serious attempts to estimate the total number of species in the world. Wilson (1988) suggested that by considering extrapolations made of the number of insects (the richest animal group) in

the rain forest the total number of species should be somewhere between 5 to 30 million species. A recent estimation by Mora et al. (2011) mentions that the total number of species is around 8.7 million, and according to the number of already cataloged species (1.2 million), about 86% of species remain to be discovered. The biodiversity on Earth has declined dramatically in the last few years, mainly due to human activities (Chapin et al., 2000; Wilson, 1988) and consequently, many undescribed species may become extinct without our knowing of their existence. For this reason, some authors like Myers et al. (2000) think that biodiversity conservation studies should be focused on specific areas in which positive results would yield the highest impacts. Myers et al. (2000) identified areas around the world with high levels of biodiversity, endemism, and high rates of habitat loss and coined them as biodiversity hotspots.

Among these areas is Western Ecuador, which despite harboring a large proportion of endemic species (Barriga, 2012; Albert and Reis, 2011; Lynch and Duellman, 1997), has been highly impacted by human activities, with large areas of the country suffering from severe habitat degradation and loss. An example of this is the destruction of mangrove forests, which serve as a refuge and breeding habitat for many species, promoted by shrimp companies for aquaculture purposes (Hamilton and Lovette, 2015). Population settlements close to the rivers, the construction of dams, pollution from agriculture, mining and human settlements, and over-fishing have also contributed to the deterioration of aquatic ecosystems (Chapin et al., 2000). Consequently, the biodiversity in the area has been severely declining (Dodson and Gentry, 1991) and not enough is known of the ecology and evolution of most species in the region to prioritize conservation action.

Endemism in Western Ecuador

Western Ecuador, along with the biogeographic region of Chocó (Western Colombia) and the Darien region (Southern Panama), is considered a Biodiversity Hotspot. Western Ecuador shares 418 (of 1,625) endemic vertebrate species with the Darien-Chocó region, as well as 2250 (of 9,000) endemic species of plants (Myers et al., 2000). Among the main factors making these regions prone to high levels of endemism, is their geographic isolation caused by the Andean Mountains, and their regional climates. Hydrographic systems comprised of drainage basins also contribute as an important isolating factor, particularly for freshwater fishes. Barriga (2012) classified the ichthyofauna of Ecuador into Ichthyo-hydrographic zones (hydrographic areas characterized according to the species composition of freshwater fishes) highlighting the high rates of endemism of fishes in western Ecuador. He reported that approximately 218 of the total number of species of freshwater fishes in the country (i.e., 951 species) are endemic to Western Ecuador. Another study by Jimenez-Prado et al. (2015) that defined freshwater fishes more strictly to not include some species that are more estuarine, reported that 43 of 112 freshwater fishes from Western Ecuador were endemic. It should be noted that although endemism is high, species diversity in western Ecuador is known to be low in contrast with the Amazon region east of the Andes (Barriga, 2012). Nonetheless, given the level of habitat degradation mentioned above, more studies focused on these unique species are imperative to help prioritize conservation efforts in this highly impacted region.

The Genus *Rhoadsia*

Rhoadsia is a genus of a freshwater fish endemic to western Ecuador and northwestern Peru. It belongs to the family Characidae known to be the most diverse family of freshwater fishes

in South America with 1,345 species described, followed by the Loricariidae with 973 species, and the Cichlidae with 571 species (Albert and Reis, 2011). *Rhoadsia* is characterized by having a deep body, and a black spot towards the center of its body. It has a pronounced sexual dimorphism that makes it distinct from other characids in the region. Males are bigger with enlarged dorsal and anal fins, have iridescent coloration and bright red and orange colors on the edges of their fins and sometimes on their bodies (Jimenez-Prado et al., 2015). This genus has two recognized species that are morphologically similar: *Rhoadsia altipinna* Fowler, 1911 and *Rhoadsia minor* Eigenmann and Henn, 1914. They mainly differ in their body depth; *Rhoadsia altipinna* has a deeper body while *R. minor* tends to be more streamlined (Fig. 1). Another difference between the species is that *Rhoadsia minor* is known to inhabit high elevations, mostly associated with the Esmeraldas region (Jimenez-Prado et al., 2015; Barriga, 2012), while *R. altipinna* is commonly found at low elevations of the Guayas drainage and rivers farther south to northern Peru (Jimenez-Prado et al., 2015; Barriga, 2012; Ortega et al., 2012). Despite these differences, there is still significant taxonomic uncertainty regarding the validity of these species. For example, Böhlke (1958) noticed that low-elevation Esmeraldas population of *Rhoadsia* resembled R. *altipinna* from the Guayas basin. Géry (1977) suggested that these species might be two geographically isolated subspecies adapted to local conditions at different elevations of mountain streams. A more recent study by Aguirre et al. (2016) examined the morphological variation of *R. altipinna* in a small river in southwestern Ecuador and found significant divergence of body depth (which is also the main trait used to distinguish the two species) related to elevation, suggesting that the trait typically used to distinguish between species also varies within species among populations in different habitats. Malato et al. (2017) expanded on the work of

Aguirre et al. (2016) and found that the body depth of populations of *Rhoadsia* varies continuously along elevational gradients in different rivers and that this difference in body shape overlaps among the drainages in which the two described species occur. Her research showed that high-elevation *R. altipinna* present similar body depth to low-elevation *R. minor* along with little divergence among populations based on the molecular analysis conducted with the mitochondrial Cytochrome oxidase gene (*COI*) and the nuclear gene S7. However, this study was limited in scope for the molecular data. The relationship between *R. altipinna* and *R. minor* is not clear yet and more data are needed to clarify their relationship.

Major Rivers of Western Ecuador

Rivers in Ecuador are organized in hydrographic systems and drainage basins. These are defined as a portion of the area that receives water from a natural source (e.g., precipitation or mountain glaciers) and conveys it to a common point such as a main river, lake, or ocean. Western Ecuador has 24 hydrographic systems comprised of 72 drainage basins (including the Galapagos islands) (National Council of Water Resources, 2002). The largest drainages in western Ecuador are the Esmeraldas (Northwestern) and Guayas (Center) drainages (Illustration 1), and harbor most of the freshwater fishes of the region. Most of the other rivers are quite small and some, particularly those in coastal regions, are seasonal and harbor few fish species. Rivers in western Ecuador form a complex system of isolated drainages that mostly come from the Andean mountains and drain into the Pacific Ocean. These drainages are completely isolated from rivers east of the Andes that ultimately connect to the Amazon drainage. The geographic isolation of the drainages in western Ecuador has resulted in the high levels of endemism in the region.

Below, I will briefly describe drainages that were considered in the sampling area. They were: Sua, Atacames, Esmeraldas, Guayas, Taura, San Pablo, Jubones, and Santa Rosa (Illustration 1). Sua and Atacames are small rivers that belong to the hydrographic system of Muisne in the northwestern part of Ecuador. The Sua drainage covers an area of 69 km^2 while Atacames covers an area of 312 km^2. Both drain north to the Pacific Ocean. The Esmeraldas drainage is one of the largest drainage basins in western Ecuador and is located next to the Atacames River, covering an area of 21,553 km^2. Its tributary rivers are fed from the western side of the Andean mountains, draining into the Esmeraldas River proper, which runs northwest into the Pacific Ocean (Appendix S1). To the south of the Esmeraldas drainage, there is the Guayas drainage, which is the largest drainage basin in western Ecuador, covering an area of 32,218 km^2. Rivers from this basin collect water from the western Andes and the eastern slopes of the Cordillera Chongon-Colonche (coastal mountain range), with these mountain chains funneling water south to the Gulf of Guayaquil and then to the Pacific Ocean. This direction of flow is different to that of most other rivers in western Ecuador, which tend to drain east to west from the Andes Mountains into the Pacific Ocean. Below the Guayas drainage is the Taura drainage, which covers an area of 1,962 km^2 (Appendix S1). Its waters go east to west from the Andes to the Gulf of Guayaquil and then the Pacific Ocean. The next basin sampled to the south is the San Pablo basin, which covers an area of 176 km^2. It is one of several relatively small basins that compose the hydrographic system of Naranjal Pagua. Tributaries from this basin travel east to west directly into the Pacific Ocean. The Jubones drainage, which covers an area of 4,361 km^2, and the Santa Rosa drainage, which covers an area of 309 km^2 are the southernmost rivers sampled (Appendix S2). Both basins drain west directly into the Pacific Ocean (National Council of Water Resources, 2002)(Illustration 1).

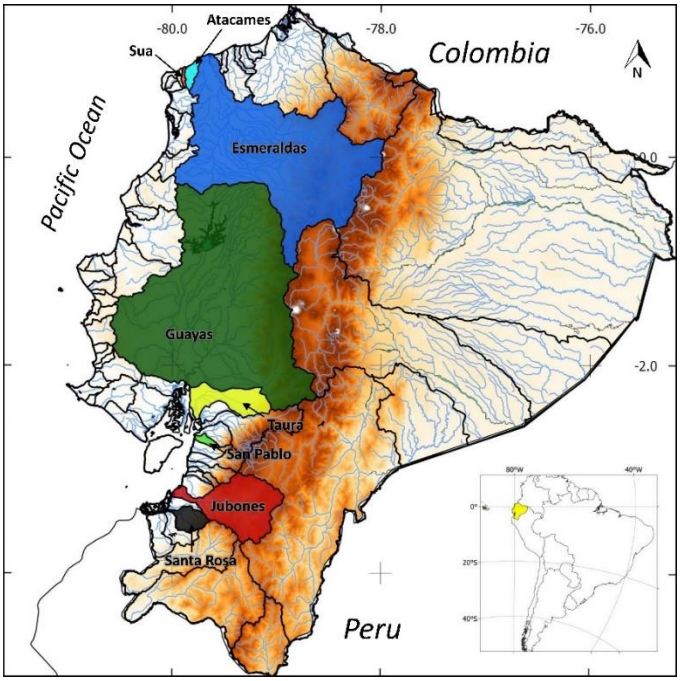


Illustration 1. *Drainage boundaries (black line) of continental Ecuador. Drainages used in this study are represented with different colors.*

Phylogeography

Phylogeography is an area of active research aimed at understanding the evolutionary relationship of populations in a geographic context. It combines a phylogenetic analysis of a species typically resolved using molecular markers with the geographic distribution of its populations (Avise, 2000). Phylogeographic analyses can be used to explain historical biogeographic patterns of species based on multiple co-distributed taxonomic groups (comparative phylogeography) (Arbogast and Kenagy, 2001), as well as to reveal cryptic biodiversity and the processes that promote lineage diversification (Reilly et al., 2015). The highest levels of divergence in populations tend to increase depending on how far they are from the core of the species range (Duncan et al., 2015). The more geographically distant a population is, the more likely it will differ in its genetic composition, which eventually can lead to a speciation

event. The main component required to observe this degree of divergence is the complete isolation of the population, such that individuals are not able to interbreed with members of other populations. Several studies developed in different populations of species have supported this idea. For example, a phylogeographic study of a species of Hummingbird (genus *Adelomyia*) that occurs in the Andean Mountains showed that the strongest sign of divergence occurred between populations that were more geographically distant (Chaves and Smith, 2011). Likewise, another study developed by Cucalón & Bajaña (2015) in a freshwater fish (*Hoplias spp*.) in western Ecuador showed a similar result, also detecting what appeared to be a potential new species in the northernmost part of the sampling area (Santiago-Cayapas drainage) based on the magnitude of genetic divergence of the population. Phylogeographic studies are becoming more common in population genetic studies due to the increasingly easy access to molecular data from worldwide databases and constantly improving technology.

Use of Molecular Markers in Evolutionary Studies

Molecular markers refer to specific DNA sequences with a known location in the genome of an organism (e.g., mitochondrial or nuclear). They have become a basic tool for the study of populations. They provide a powerful framework that allows answering questions related to population genetic topics such as the biogeography and evolutionary history of organisms. Among the most common molecular markers are DNA sequences and microsatellites. However, other types of markers include SNPs, RAPDs, AFLPs, RFLPs, SSCPs among others (Sunnucks, 2000). The choice of the molecular marker typically depends on the type of question. For example, DNA sequences are normally applied in phylogenetic and phylogeographic studies. A portion of a gene is sequenced, allowing the detailed comparison of genetic nucleotide differences among

individuals. This process used to be expensive, however, as technology progresses, this process has become more accessible. Microsatellites are another popular type of molecular marker for phylogeographic studies that are more commonly found in the nuclear genome (nDNA). Microsatellites consist of a series of short sequences of nucleotides (two to four nucleotides) repeated in tandem with high rates of evolution despite being located in the nDNA. Alleles are scored based on differences in the number of repeats inferred from fragment size. Most natural populations exhibit substantial genetic variation at microsatellite loci (Selkoe and Toonen, 2006). Microsatellites are usually used to find genetically structured populations or patterns of differentiation among populations within a species. They are often relatively inexpensive (~$1 US dollar per sample) since they do not require obtaining the specific DNA sequence but simply scoring the size of the alleles, and they can be multiplexed, such that multiple loci can be scored per individual in a single run. Nevertheless, they do require the use of multiple loci (generally more than 10), making the process of collecting microsatellite data relatively labor intensive and time-consuming.

Mitochondrial DNA (mtDNA) genes like cytochrome b (*Cyt-b*) and cytochrome oxidase I (*COI*) are very commonly used in phylogenetic and phylogeographic studies (Zink and Barrowclough, 2008; Avise et al., 1987). The mtDNA is maternally inherited (Illustration 2), so it allows tracking of matrilineal lineages of populations, making it a good candidate for phylogenetic studies conducted at lower taxonomic ranks because of its higher rate of evolution compared to most nuclear genes (Zink and Barrowclough, 2008). Mitochondrial genes vary in their rates of evolution. For example, the *Cyt-b* gene is known to have a higher rate of evolution than the *COI* gene and has been widely used in phylogenetics since it is more useful in the assessment of more

recent relationships (Farias et al., 2001). On the other hand, *COI* due to its more conservative nature has been commonly used as a genetic barcode for species identification (Kress et al., 2015) when comparing against preexisting sequence databases such as GenBank.

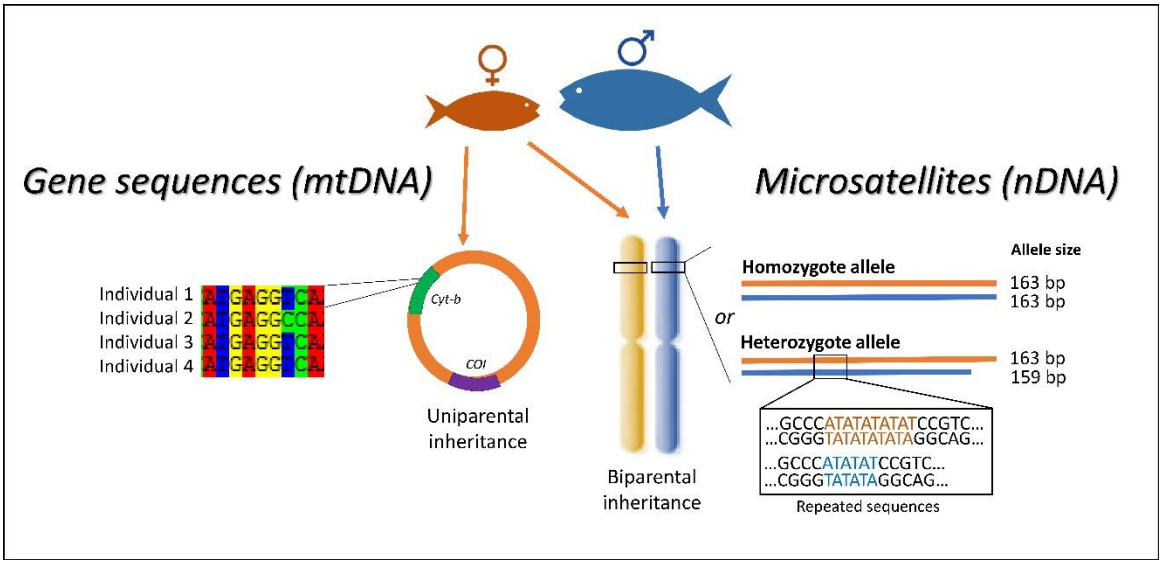


Illustration 2. Simplified representation of the molecular markers implemented in this study.

CHAPTER II

Introduction

Biodiversity on earth has been declining due to human activity (Newbold et al., 2015; Chapin et al., 2000; Wilson, 1988). The loss of biodiversity affects humankind considering the variety of ecosystem services that we rely on (Chivian and Bernstein, 2008). However, how can we address the biodiversity crisis if we do not even know how many species there are or the specific ecological functions they perform? One study estimated the number of species on earth to be about 8.7 million of which 86% have not been described (Mora et al., 2011). Other estimates vary widely. Some species might disappear before even being discovered. Species and threats are distributed heterogeneously around the world and efforts should be focused on strategic areas where results per unit effort can be maximized. Areas that are highly threatened and have high levels of endemic species are known as biodiversity hotspots (Myers et al., 2000). One such biodiversity hotspot occurs in western Ecuador.

Ecuador is a small South American country that harbors a great diversity of animals and plants, many of which are endemic. Rates of endemism are particularly high in western Ecuador because of its heterogeneous topography, variable climates, and its geographic isolation resulting from the rise of the Andes mountains (Jimenez-Prado et al., 2015; Barriga, 2012; Albert and Reis, 2011). Despite the rapid environmental degradation in this region (Ribeiro et al., 2017; Damanik-Ambarita et al., 2016), studies of the evolutionary relationships of species are scarce, especially in aquatic ecosystems.

There are 24 isolated hydrographic systems comprised of 72 drainage basins in western Ecuador that mostly come from the Andean mountains and drain into the Pacific Ocean. The two largest basins in Western Ecuador are the Esmeraldas basin which drains an area of 21,553 km^2 in northwestern Ecuador and the Guayas basin which drains an area of 32,218 km^2 in the central portion of western Ecuador (National Council of Water Resources, 2002). Most of the other drainages are relatively small and some, particularly those in coastal regions, are seasonal and contain few fish species. These drainages are also completely isolated from rivers east of the Andes that connect to the Amazon drainages, allowing their evolutionary divergence.

Precipitation intensity is crucial for climatic variability in Western Ecuador. Precipitation varies seasonally (wet and dry season) and geographically, declining from north to south. This characteristic, together with the influence of oceanic currents, helps to create a variety of regional climates in a relatively small geographic area. Western Ecuador has arid habitats with very low annual precipitation (e.g., near the coast and in southwestern most portion of the country), areas with intermediate precipitation levels and moderate rain seasons (e.g., central and southern regions including the Guayas drainage basin), and areas with high levels of precipitation that are very humid (e.g., Northwestern Ecuador, including the Esmeraldas drainage basin). These features along with the geological history of the coast may have contributed to the divergence in species composition and the formation of the distinctive biogeographic zones in Ecuador (MAE, 2012), which for freshwater fishes appear to coincide with the boundaries formed by the drainages, especially the Esmeraldas and the Guayas basins.

Elevation may also have an important role in the diversification of species in Western Ecuador (Winemiller et al., 2009). The rivers of Western Ecuador originate mostly from Andean

glaciers at approximately 4000 m above the sea. River bottoms at higher altitudes are mainly composed of boulders and cobble, and the flow of water is usually strong due to steep slopes, which make it a challenge for some species to move upstream. Water temperature is also typically lower at higher elevations (MAE, 2012; Winemiller et al., 2008). As elevation declines, water currents become slower, river bottom composition changes to pebbles and sand-like soil, conductivity increases due to the accumulation of minerals, and water temperature rises (MAE, 2012; Winemiller et al., 2008). There are also changes in the composition and functional groups that occur at different elevations, such as in the abundance of large predatory fish, which tend to be more common at lower elevations than higher elevations (Vannote et al., 1980). Changes in environmental conditions associated with elevation likely influence the ecology and may influence the evolution of populations inhabiting the Andean mountain streams (Aguirre et al., 2019; Malato et al., 2017). The variety of environments added to the geographical isolation of the region have resulted in high rates of endemism and significant divergence in species composition over small geographic areas, making this an interesting region for studies of evolution and speciation.

One of the endemic genera found in western Ecuador is the genus *Rhoadsia,* a freshwater fish that belongs to the family Characidae, the most diverse fish family in the Americas. The genus *Rhoadsia* is part of the small subfamily Rhoadsiinae comprised of only four genera: *Carlana*, *Nematocharax*, *Parastremma*, and *Rhoadsia*. This group is characterized by having a single tooth series in the premaxilla when young, and a double series when adults. Two conical teeth comprise the outer series and five multicuspid teeth comprise the inner series (Cardoso, 2003). The longest dorsal fin ray may also reach the caudal fin in adult males (Cardoso, 2003). Two morphologically

differentiated species of *Rhoadsia* are currently recognized, *Rhoadsia altipinna* Fowler, 1911 and *Rhoadsia minor* Eigenmann and Henn, 1914 (Fig. 1). *Rhoadsia minor* has been documented from high elevations in northwestern Ecuador (Mindo, Esmeraldas) (Jimenez-Prado et al., 2015; Barriga, 2012) while *R. altipinna* was originally described from lower elevations in the central region of western Ecuador (Guayas basin), also occurring in northern Peru (Jimenez-Prado et al., 2015; Barriga, 2012; Ortega et al., 2012). Both species are very similar morphologically having a distinctive black spot located almost at the center of their body, and present a deeper body compared to other characids. Although they are not commercially used as a food source due to their small length (100 – 170 mm), it has been suggested that they could be used as ornamental fish due to their bright colors (Jimenez-Prado et al., 2015). The genus *Rhoadsia* exhibits sexual dimorphism; males are much larger (about twice larger), have brighter red and iridescent colors, and have larger dorsal, anal and pectoral fins than females (Jimenez-Prado et al., 2015) (Fig.1).

Although the genus *Rhoadsia* currently includes two species, some authors have questioned their validity due to morphological and geographical inconsistencies (Géry, 1977; Böhlke, 1958). The main distinction between them is body depth. *Rhoadsia minor* tends to be more streamlined, while *R. altipinna* has a deeper body. However, this feature is now known to vary with elevation in both species (Malato et al., 2017; Aguirre et al., 2016). Low-elevation populations from Esmeraldas have been noted to be morphologically similar to *R. altipinna* (Böhlke, 1958). Moreover, Géry (1977) suggested that *R. altipinna* and *R. minor* may constitute a single species separated into morphologically variable populations adapting to local conditions at different elevations. A genetic and morphometric study of *Rhoadsia altipinna* in a few watersheds revealed significant divergence in morphology with elevation, suggesting parallel

patterns of local adaptation in different rivers of western Ecuador. Nevertheless, the genetic analysis based on cytochrome oxidase gene (*COI*) showed little genetic divergence among populations of *Rhoadsia* and no reciprocal monophyly of haplotypes between the two putative species (Malato et al., 2017; Aguirre et al., 2016), suggesting that a more comprehensive study is necessary to verify their validity as two separate species.

The main goal of my study is to examine how populations of *Rhoadsia* in Western Ecuador are evolutionarily related. I expanded the sampling area from an earlier study (Malato et al., 2017) to include all major rivers in western Ecuador known to harbor *Rhoadsia* populations and conducted a phylogeographic analysis based on genetic markers with higher rates of evolution than those used previously: *Cytochrome b* (*Cyt-b*) from the mitochondrial genome (mtDNA) and microsatellites from the nuclear genome (nDNA). Additional data for the *COI* gene were also generated. Within this major goal, there are three main objectives. First, I seek to clarify the taxonomic status of the currently recognized species by determining the most likely number of species according to the inferred population structure. Second, I investigate the possible geographical origin of the genus by comparing phylogenetic relationship with the outgroup and levels of genetic diversity, under the assumption that greater levels of genetic diversity and a closer relationship to the outgroups indicates areas of probably ancestry. Third, I document the evolutionary breaches between drainages and contribute to identifying the major aquatic zoogeographic regions in western Ecuador by examining the patterns of genetic divergence of *Rhoadsia*. This study will contribute to our understanding of the biogeography of aquatic species in western Ecuador and will provide a baseline for future studies of the ecology and evolution of *Rhoadsia* and other freshwater fishes in the region.

Methods

Data Collection

A combination of new and existing samples of *Rhoadsia* and closely related genera were included in this study. Part of the data were obtained from DNA extracted from populations collected in 2008 and 2014 and included in Malato et al. (2017). A set of samples collected in 2012 were provided by the Royal Ontario Museum in Canada (ROM) (Appendix S3). ROM also contributed with tissue samples of a closely related genus in the subfamily Rhoadsiinae, *Parastremma sadina* from Colombia, which was used in the phylogeographic analysis as an outgroup. DNA samples of another closely related genus in the Rhoadsiinae, *Carlana eigenmanni,* were obtained from the Smithsonian Tropical Research Institute in Panama (Appendix S3). These were also used as an outgroup. Finally, samples of up to 24 individuals per population were collected in the summer of 2017 to complement existing collections. Samples were collected from rivers in eight different drainages: Sua, Atacames, Esmeraldas, Guayas, Taura, San Pablo, Jubones, and Santa Rosa (Table 1 and Fig. 2).

Fish collected in this study were obtained from streams located at elevations of 20 m to 1260 m above sea level. The sites selected for sampling were shallow (~500 mm depth) running streams composed mainly of cobble, boulder, gravel, and sand surrounded by vegetation (See appendix S4). A Smith Root model LR24 electrofishing backpack was used for capturing the specimens along with sienes and dip nets. The general outputs set in the equipment were 400 to 850 Volts, a frequency of 25 – 30 Hz and a standard pulse type. The setting options of the equipment were adjusted depending on the environmental conditions of the sampling area. The electrofishing backpack was only used in shallow water with a conductivity lower than

approximately 200 µs. A cast net was used for deeper water or where the electrofishing could not be used due to the high conductivity (Appendix S5). Specimens were euthanized in MS-222 or an ice water bath. Fin clips were collected and stored in 95% ethanol, while the rest of the specimen was fixed in 10% formaldehyde for at least 48 hours, washed in water, and then transferred to 70% ethanol for long term storage.

The Cytochrome b (*Cyt-b*) and Cytochrome Oxidase I (*COI*) genes from the mtDNA and 12 microsatellites isolated by Loh et al. (2013) from the nDNA were used for the phylogeographic and population genetic analyses. I utilized a maximum of 24 specimens per location for the analysis of both *Cyt-b* and the microsatellites. I did not include samples with less than ten individuals for the microsatellite analysis since these would likely give inaccurate estimates of allele frequencies. In the case of the outgroup genera *Carlana and Parastremma*, the genetic distance between them and *Rhoadsia* would have also likely made an analysis with microsatellites uninformative given the high rates of evolution of microsatellite markers. *COI* sequences from Malato et al. (2017) were retrieved from Gene Bank and included in this study (Accession KY440344 - KY440356). I used up to 8 individuals per sample from the new sites collected in this study for the analysis of *COI* (Table 1). Smaller sample sizes were used for *COI* because previous work indicated that this gene varies relatively little among populations of *Rhoadsia* (Malato et al., 2017).

DNA Extraction

DNA extraction was performed using a phenol:chloroform:isoamyl alcohol method following Aguirre et al. (2016). The tissue was digested by a lysis buffer and proteinase k (20 mg

μl^{-1}) and incubated at 55°C overnight for the denaturation of the protein to occur. DNA was isolated the next day using 25:24:1 phenol:chloroform:isoamyl alcohol followed by several ethanol wash steps to eliminate residues from the digestion. Finally, DNA was resuspended in TE (10 mM Tris pH 8, 1 mM EDTA pH8) buffer and stored at -20°C for further analysis. A working stock diluted in water (final concentration 2-25 ng/µl) from the concentrated TE stock, was used in subsequent reactions.

Mitochondrial Genes Amplification (*Cyt-b and COI*)

The polymerase chain reaction (PCR) technique was used to amplify the *Cyt-b* and *COI genes*. Two different set of primers were used for the amplification of *Cyt-b* due to some problems with the quality of the PCR product. The first set was used to amplify *Cyt-b* for some samples of *Rhoadsia spp.*, and all samples of *Carlana eigenmanni*, and *Parastremma sadina*: The primers were: CB1-5 (Forward), 5'-CCATCCAACATCTCAGCATGATGAAA-3' and CB3-3 (Reverse) 5'-GGCAAATAGGAAATATCATTC-3' (Palumbi, 1996). A more specific set of primers for *Rhoadsia spp.* was created and tested using invariant sites of the DNA sequences obtained with the first set of primers as a reference. The second set of primers were Rh_Cytb_F2 (Forward), 5'-GACATTTCTTTAGCTTTTTCC-3' and Rh_Cytb_R1 (Reverse), 5'-GGTTTAATATGAGGTGGTGT-3'. PCR reactions were run in a total volume of 30 µl, which included the following reagents: 1X PCR buffer, 0.2 Mm $MgCl_2$, 12 U BSA, 0.25 µM of each dNTP, 0.35 µM of each primer, and 0.75 U Taq Polymerase. The thermocycler protocol implemented for the amplification was: one cycle of 95°C for 1 min 45 s, 52°C for 45 s, and 72°C for 45 s, 3 cycles of 94°C for 45 s, 52°C for 45 s, and 72°C for 45 s, 29 cycles of 92°C for 30 s, 52°C for 45 s, and 72°C for 45s, and 1 cycle of 72°C for 7 min for the final extension. PCR and thermocycler conditions for *COI* amplification were similar to *Cyt-*

b except for the annealing temperature, which was 50°C. The primers used for *COI* were*:* FISH-BCL (Forward), 5'-TCAACYAATCAYAAAGATATYGGCAC-3' and FISH-BCH (Reverse), 5'-TAAACTTCAGGGTGACCAAA AAATCA-3' (Baldwin et al., 2009). I used 10 µl of the final product to confirm PCR amplification through agarose gel electrophoresis (1.5 % gel). Bands were visualized through UV exposure using SYBR Safe. The remaining 20 µl of the PCR product was sent to the University of Arizona Genetics Core for purification and Sanger sequencing.

Microsatellite Amplification

For the microsatellite amplification, I added a fluorescent tag into the PCR mix following Schuelke (2000). This is an inexpensive method for adding fluorescent tags to PCR products in which universal fluorescent tags can be linked to specific primers, and it is as efficient as the conventional technique normally used for genotyping microsatellite markers. The procedure consists of amplifying the microsatellite locus using standard primers with a nucleotide extension in the 5' end of the forward primer. A universal fluorescent dye label primer is added to the reaction to bind to the 5' extension of the forward primer and is used later to identify alleles by their size. The universal fluorescent dyes used were FAM, HEX, and NED. The specific primers used for the amplification of the microsatellite loci were based on those developed by Loh et al. (2013) of which twelve were used based on the reported level of heterozygosity and their ease of amplification (Appendix S6). PCR reactions were run in 10 µl volumes as follows: 1X PCR buffer, 0.2 mM $MgCl_2$, 4 U BSA, 0.25 µM of each dNTP, 0.1 µM of primer forward, 0.35 µM of primer reverse, 0.35 µM of fluorescent dye label, and 0.35 U Taq Polymerase. The thermocycler protocol implemented was the same for all loci except for the annealing temperature (Appendix S6); one cycle of 95°C for 1 min 45 s, annealing temperature for 45 s, and 72°C for 45 s, 3 cycles of 94°C

for 45 s, annealing temperature for 45 s, and 72°C for 45 s, 21 cycles of 92°C for 30 s, annealing temperature for 45 s, and 72°C for 45s, 8 cycles of 92°C for 30 s, 53°C for 45 s, and 72°C for 45s and 1 cycle of 72°C for 10 min for the final extension. After the PCR products were obtained, they were sent to the University of Arizona's Genetics Core Fragment Analysis Facility for genotyping.

Sequence Editing and Genotyping

DNA Sequence Editing

Editing of the sequences of *Cyt-b* and *COI* was done using Geneious Version 11.0.5 (https://www.geneious.com). The final alignment of the sequences was carried out using the MUSCLE method, which is highly recommended when large samples are being analyzed (more than 100 sequences). The alignment was inspected manually to correct possible errors made by the program. A total of 357 sequences for *Cyt-b* and 217 for *COI* including the outgroups were trimmed from the edges to obtain the same length for all sequences, 555 nucleotides for *Cyt-b* and 588 for *COI*.

Microsatellite Genotyping

Chromatograms were edited and scored in Geneious version 11.0.5 (https://www.geneious.com) using the plug-in for microsatellites. Microsatellites where scored through the program and then manually confirmed to reduce errors. Twenty four samples were randomly selected and analyzed twice to confirm the accuracy of the scoring method.

Phylogeographic and Population Genetics Analyses

Linkage Disequilibrium and Hardy Weinberg Equilibrium

Genepop on the web (Rousset, 2008; Raymond and Rousset, 1995) was used to test the probability of linkage disequilibrium between each pair of loci along with the Hardy Weinberg Equilibrium exact test. The tests were corrected for multiple comparisons using the Bonferroni correction, α/n, where n is the total number of comparisons.

Neutrality Test

I evaluated the values of Tajima's D (Tajima, 1989) and Fu's Fs (Fu, 1997) generated in Arlequin version 3.5.2.2 (Excoffier et al., 2005). This analysis tests the assumption that the mutations are generated randomly in the population. Any deviation from this hypothesis could result from demographic events such as recent population expansions or selective sweeps (reduction of genetic diversity due to fixation of alleles that are under selection) (negative values), and recent population contractions or balancing selection (positive values). Zero is interpreted as no deviation from expected values, hence, evolution is neutral.

Mantel Test

A Mantel test was used to test whether there was a significant correlation between genetic and geographic distances. Pairwise F_{ST} values generated in Arlequin were used for the genetic distance matrix. The pairwise geographic distance matrix was based on straight line distances between sites and was created using the program Geographic Distance Matrix Generator, version 1.2.3 (http://biodiversityinformatics.amnh.org/open_source/gdmg). The Mantel test was performed in Arlequin version 3.5.2.2 (Excoffier et al., 2005).

Genetic Diversity

The genetic diversity indices used included the effective number of alleles (N_e), haplotype diversity (HD), observed heterozygosity (Ho), expected heterozygosity (He), allelic richness (AR), and the number of private haplotypes (PH), which are commonly used in population genetic studies (Excoffier et al., 2005). It is important to notice that the genetic diversity comparisons among sites were not statistically tested, since each genetic diversity value at each site represented a single datum calculated from the overall frequency of haplotypes/alleles at each location. Some drainages such as Sua and Atacames (Su-At), and Guayas and Taura (Gua-Ta) were combined due to their proximity and similar haplotypic composition. Genetic diversity indices based on microsatellites such as allelic richness, the effective number of alleles, observed heterozygosity, and expected heterozygosity were calculated using GeneAlex version 6.5 (Peakall and Smouse, 2012; Peakall and Smouse, 2006). The mean effective number of alleles across loci per population was used as a measure of genetic diversity. Genetic diversity indices based on mitochondrial DNA genes such as haplotype richness, haplotype diversity, the effective number of alleles, and the number of private haplotypes were calculated manually using the information provided in the haplotype network analysis. The effective number of alleles (N_e) based on mtDNA genes of each population was calculated as $1/(\Sigma p_i^2)$, where P_i is the frequency of the i^{th} allele (Frankham et al., 2002). Haplotype diversity (HD) was calculated as $N(1-\Sigma p_i^2)*(N-1)^{-1}$, where P_i is the frequency of the i^{th} allele and N is the number of individuals in the sample (Nei, 1987). In addition, I examined whether genetic diversity was associated with elevation by implementing a general linear regression using R, version 3.4.4 (The R project for statistical computing).

Genetic Differentiation

Arlequin version 3.5.2.2 (Excoffier et al., 2005) was used to conduct an Analysis of Molecular Variance (AMOVA) using the two mitochondrial genes *Cyt-b* and *COI* and the microsatellites. I performed three different AMOVAs where the components of genetic variation were examined as follows: The first AMOVA evaluated the divergence among drainages (i.e., Sua, Atacames, Esmeraldas, Guayas, San Pablo, Jubones, and Santa Rosa), divergence among populations within drainages, and divergence within populations. For the second AMOVA, I used the results from the phylogenetic analysis carried out with the mtDNA markers (i.e. *Cyt-b* and *COI*) to test the divergence between groups of the main monophyletic clades (Northern sites vs. Southern sites), among populations within the clades, and within populations. Finally, the hierarchical levels of the third AMOVA were based on the number of clustered populations estimated in the STRUCTURE analysis (i.e., K=10) (see Appendix S7 for the specific assignment of the sites into the groups for each version of AMOVA). Pairwise F_{ST} was calculated in Arlequin to examine the level of genetic divergence between each pair of populations. F_{ST} values were evaluated following the interpretation suggested by Harlt and Clark (1997), in which values of less than or equal to 0.05 were considered as little genetic differentiation, greater than 0.05 and less than or equal to 0.15 as moderate genetic differentiation, greater than 0.15 and less than 0.25 as great genetic differentiation, and values greater than or equal to 0.25 as very great genetic differentiation.

Phylogenetic Analysis

Phylogenetic relationships of populations of *Rhoadsia* were tested using a maximum likelihood (ML) and a Bayesian approach for the *Cyt-b* and *COI* mitochondrial genes. Duplicated

sequences were removed from the alignments from each sampled population to reduce computational cost. A total of 61 and 64 sequences were used for *Cyt-b* and *COI* respectively, after removal of duplicates. The nucleotide substitution model selected for both the ML and Bayesian methods and both genes (*Cyt-b* and *COI*) was the HKY + G model (Hasegawa et al., 1985) as suggested by the program jModelTest (Posada, 2008) using the Bayesian Information Criterion (BIC) (Bhat and Kumar, 2010). PhyML (Guindon and Gascuel, 2003) implemented in SeaView v4.5.4 (Gouy et al., 2010) was used for the estimation of the ML Tree using the approximate likelihood ratio test, Shimodaira-Hasegawa-Like (aLRT SH-like) (Guindon et al., 2010) for branch support. Default values were used for the rest of the parameters. Bayesian analysis parameters were set using the program Beauti v1.8.4 included in the package of BEAST v1.8.4 (Drummond et al., 2012). The parameters used were: a strict molecular clock which assumes uniform rates across branches, random starting tree and a Markov Chain Monte Carlo (MCMC) length of 50 million iterations with resampling every 5000 samples, discarding 5 million iterations as burn-in. The coalescent model used was Bayesian Skyride (Minin et al., 2008). The effective sample size (ESS) which measures the independence of sampled trees was inspected using the program Tracer v1.6 (Rambaut et al., 2014) and then summarized in TreeAnnotator v1.8.4 (included in Beast Package). The maximum clade credibility (MCC) tree produced was then visualized in FigTree v1.4.2 (Rambaut, 2016).

Haplotype Networks

Haplotype networks for the *Cyt-b* and *COI* genes were created using the program PopART (Leigh and Bryant, 2015). The gene network shows the frequency of the shared and private haplotypes present in the sample, along with the genetic distance measured as the number of

mutations between haplotypes. The networks were inferred using median-joining method. Outgroup sequences were excluded from the total alignments.

Population Structure Analysis

STRUCTURE software version 2.3.4 (Pritchard et al., 2000) was used to investigate the population structure based on the microsatellite markers. This method estimates the most likely number of genetically distinct clusters —known as K— where each individual is assigned to a distinctive cluster based on a Bayesian algorithm. The STRUCTURE analysis provides valuable insight into the number of genetically distinct units of *Rhoadsia* that should be recognized in western Ecuador. Twenty replicates of each possible number of K (1 to 16, the latter being the number of samples included) were run with 1,000,000 MCMC iterations after 250,000 of burn-in. An admixture model was used, which assumes that the populations might have multiple origins and have some genetic admixture. The allele frequencies correlated model was selected. The most likely value of K was estimated according to the mean Log-normal probability of K (LnPK) provided directly by STRUCTURE (Pritchard et al., 2000) and also by calculating ΔK (Evanno et al., 2005) using the program Structure harvester (Earl and VonHoldt, 2012). Cluster Markov Packager Across K (CLUMPAK) was used to analyze the 20 independent runs selected after calculating the "best" value of K. It generates a consensus solution of the clustering using a Markov Clustering algorithm which relies on the similarity of the matrix between replicate runs (Kopelman et al., 2015). A neighbor-joining tree was created in STRUCTURE based on the p-distance computed from the inferred clusters.

A neighbor-Joining tree per population based on the microsatellite data was created using the program PopTree2 (Takezaki et al., 2010) with 1000 bootstrapped replicates. In addition, a

neighbor-joining tree per individual was generated in the program Population 1.2.32 (found in: http://bioinformatics.org/~tryphon/populations/). Both trees were inferred using the Nei's D_A genetic distance method (Nei et al., 1983) usually recommended for microsatellite data (Takezaki and Nei, 1996).

Results

Genetic Divergence Between *Rhoadsia* and the Outgroups

The analysis of the mitochondrial genes was carried out on a total of 357 and 217 successfully sequenced specimens from the *Cyt-b* and *COI* genes, respectively, including from multiple populations of *Rhoadsia* in Western Ecuador and the two related genera from Central America and Colombia used as outgroups. The outgroups included two species of the subfamily Rhoadsiinae: seven specimens of *Carlana eigenmanni* from Costa Rica (1), Nicaragua (2), and Panama (4) and four specimens of *Parastremma sadina* from Colombia. *Rhoadsia* differed substantially from the outgroup genera, forming a monophyletic clade relative to them for both mitochondrial genes (more detail in the phylogenetic analysis section below).

For the *Cyt-b* gene, both outgroups combined showed a divergence of 26% (146 mutations total) when compared to *Rhoadsia* populations in Ecuador (ingroup). Of these, 131 mutations were synonymous substitutions and 15 were non-synonymous, the latter resulting in changes in the amino acid sequence. Samples from *C. eigenmanni* from the Central American countries alone diverged about 21% (114 mutation; 99 synonymous and 15 non-synonymous) from *Rhoadsia* while *P. sadina* from Colombia diverged about 16% (89 mutations; 81 synonymous and 8 non-synonymous).

For the *COI* gene, the total number of mutations for the outgroup was 131 (22% divergence) of which 129 were synonymous and 2 were non-synonymous. More specifically, samples from *C. eigenmanni* diverged about 16% (94 mutations; 93 synonymous and 1 non-

synonymous) while samples from *P. sadina* diverged about 14% (81 mutations; 79 synonymous and 2 non-synonymous).

Analysis of Mitochondrial Genes (*Rhoadsia* populations)

The *Cyt-b* sequence alignment without the outgroups showed 34 polymorphic sites (555 bp total length) with 36 mutations total. Thirty-one of these mutations were synonymous and only 5 resulted in non-synonymous mutations. The latter were found in the following samples: a substitution in the first codon position from G to A at the position 304 (i.e., m. 304G>A) in one individual from the site E2B. The correspondent translation resulted in one amino acid change from alanine to threonine at the position 102 (i.e., p. 102A>T). Another non-synonymous mutation occurred in the first codon positions of an individual from G1 (m. 328A>T; p.110A>T) and the second codon positions of three individuals from SP (m. 494C>A; p. 165A>D). Finally, two non-synonymous mutations were found in all individuals from At and Su, both at the second codon position (m. 464T>C) with the amino acid change p. 155I>T, and m. 329C>T with the amino acid change p. 110A>V. It is worth noting that the latter mutation was also present in all samples of the outgroups.

The *COI* sequence alignment showed 20 polymorphic sites (588 bp total length) and 21 mutations total. All mutations were synonymous, i.e., there was no change in the amino acid sequence.

Neutrality Test

The neutrality test showed no statistically significant deviation from zero for Tajima's D and Fu's Fs in both mitochondrial genes, suggesting that these genes appear to be evolving neutrally in populations of *Rhoadsia* (Appendix S8-S9).

Analysis of Microsatellite Loci

I analyzed a total of 12 loci from 343 specimens corresponding to 16 populations of *Rhoadsia* from Western Ecuador. The percentage of missing data attributable to either amplification failure during PCR or uncertainty during the scoring process varied among loci. Loci RA-20 and RA-08 were the loci with the most missing data: 21% and 19% respectively (Appendix S10). Other loci had 10% or less missing data, with an average of 4% across all loci (0.3% - 10%). Sites J1, J2, and J3 presented the greatest amount of missing data (12% - 14%). The rest the sites presented a percentage missing data under 8% (Appendix S11).

The statistical analysis did not show evidence of linkage disequilibrium between any of the possible pairs of loci after applying Bonferroni correction for multiple test (Chi-squared, $p>0.01$, $\alpha=0.00075$). The exact test for Hardy Weinberg Equilibrium (HWE) showed that At and J2 were the only sites that significantly deviated from HWE, both for the locus RA-09 ($p<0.0001$, $\alpha=0.0003$). Excluding these from the overall analyses did not fundamentally alter the results, so these were left in the analyses with microsatellites.

Genetic Diversity Indices

Genetic diversity varied among sites differing in elevation as well as among drainages. I used the haplotype diversity as a measure of genetic diversity for the *Cyt-b* and *COI* data and the

mean of the effective number of alleles across loci for the microsatellites. A linear regression of genetic diversity on elevation showed no significant trend for the *Cyt-b* (r=-0.18, df=20, p=0.43), *COI* (r=0.09, df=24, p=0.66), or the microsatellite data (r=-0.42, df=14, p=0.11) (Fig. 3). The same non-significant pattern was observed using other measures of genetic diversity like the effective number of alleles for *Cyt-b* and *COI* and expected and observed heterozygosity for microsatellites (data not shown).

In terms of variation among drainages, sites from Esmeraldas, Guayas, San Pablo, and Santa Rosa tended to generally have greater genetic diversity on average than populations from the Sua, Atacames, and Jubones drainages for both *Cyt-b* and microsatellite data (Table 2). On the other hand, *COI* data showed on average that the populations of Jubones had the greatest genetic diversity while the rest of the sites showed little interpretable variation. All the genetic markers examined concur that Atacames and Sua, although being located at low elevations (~30 m above the sea), tended to have the lowest genetic diversity (Fig. 3). Below I describe general patterns of genetic diversity for each marker. Descriptions are qualitative and not based on statistical analyses.

The *Cyt-b* data showed that the Guayas-Taura drainage present the highest number of haplotypes, substantially more than other drainages. Guayas-Taura drainage presented 15 haplotypes, followed distantly by Esmeraldas (5 haplotypes) and Jubones (4 haplotypes) (Fig. 4a, Appendix S12). In regards of the allelic richness per individual site, G1, G4, and G3 from Guayas presented the highest number of haplotypes (5, 5, and 4 haplotypes respectively), followed by E9, E4, and E6 from Esmeraldas, G2 from Guayas, SP from San Pablo, J1, J2 from Jubones and SR

from Santa Rosa, all with three haplotypes each. The rest of samples presented two or fewer haplotypes (Table 2).

The *COI* data showed the largest drainages with more similar number of haplotypes among them, however, the greatest number of haplotypes per site occurred in the Jubones drainage. The number of total haplotypes present in Guayas-Taura, Esmeraldas and Jubones was: 8, 8, and 7 haplotypes respectively. Likewise, haplotype frequency was uneven within the drainages except for Jubones which presented similar frequency among its haplotypes. The rest of drainages presented three or fewer haplotypes (Fig. 5a, Appendix S13). Regarding the individual sites, the highest number of haplotypes was observed in most of the samples from Jubones and one from Guayas. These included J1, J3, and J4 from Jubones and G1 from Guayas, each with four haplotypes. They were followed by E9, E3, and E8 from Esmeraldas, G3 from Guayas, J2 from Jubones and SR from Santa Rosa, all with three haplotypes. The rest of sites presented two or fewer haplotypes (Table 2).

The microsatellite data showed that samples from the Guayas drainage present the greatest allelic richness. The allelic richness was measured as the average number of alleles across loci. Samples G1 and G4 from the Guayas drainage showed the greatest mean allelic richness with 9.8 and 8.6 alleles respectively followed by E4 (8.2), SR (7.7), and E2B (7.6). The rest of samples presented values bellow seven alleles for the average across loci being Atacames with the lowest number of alleles (3 alleles) per loci for this genetic marker (Table 2, Appendix S14).

Private haplotypes were relatively common in all drainages but were more frequent for *Cyt-b* than for *COI* (Fig. 4a). For *Cyt-b*, the Guayas-Taura drainage presented the largest number

of private haplotypes corresponding to 13 of the total number of haplotypes (i.e., 28 haplotypes), followed by four private haplotypes for Esmeraldas and two private haplotypes for each of the remaining drainages. The frequency of the private haplotypes relative to the haplotypes present in each drainage was higher for *Cyt-b*. For example, 100% of the haplotypes found in the Sua-Atacames drainages were private, followed by 86.7% of the Guayas-Taura, 80% of the Esmeraldas, 66.6% each for the San Pablo and Santa Rosa drainages, and 43% for the Jubones drainage (Fig. 4a). It is worth noting that although the number of private haplotypes was considerably large in some of the drainages described, they were not among the most frequent haplotypes within the drainages except for H12 from Sua-Atacames and H11 from Esmeraldas (Fig. 4). The number of private haplotypes per individual site ranged from 33% to 100% and the sites with the highest percentages of private haplotypes tended to be associated with smaller sample sizes. For example, 100% of *Cyt-b* haplotypes from the G5 site were private but this corresponded to 2 of 2 specimens available for the site. On the other hand, for the *COI* data, the greatest number of private haplotypes was present in the Guayas-Taura and Esmeraldas drainages with four haplotypes each, of the total number of haplotypes (i.e., 18 haplotypes). These were followed by three private haplotypes for Jubones and one private haplotype for Sua-Atacames. The rest of drainages corresponding to San Pablo and Santa Rosa did not present private haplotypes for this gene (Fig. 5a). Private haplotypes per individual site ranged from 25% to 50% of samples with less than or equal to 4 haplotypes. Sites with the highest percentages of private haplotypes (50%) were again those with small sample sizes and with only two haplotypes in the population. These were: At (N=7), E7 (N=8), and Ta (N=2) (Table 2).

Genetic Differentiation

Populations of *Rhoadsia* were grouped in three different ways for the AMOVA: by drainage, according to the results of the phylogenetic analysis of the mtDNA genes, and by clusters estimated in the STRUCTURE analysis of the microsatellite data (Appendix S7). The among groups component seemed to be an important source of variation in all analyses of the mitochondrial genes, but not for the microsatellites. The AMOVA with the highest percentage of variation attributed to the among groups component was found after grouping the sites according to the assigned clusters estimated in the STRUCTURE analysis. For the *Cyt-b* and *COI* mitochondrial genes, it was 75% and 53% respectively. In contrast, the percentage of variation for among groups component based on the microsatellite data was always relatively low in all AMOVAs (4.98% - 16.51%). The within population component for microsatellite data always showed the highest percentage of variation, ≈ 80% in all three AMOVAs. This component corresponded to 29.78% - 37.43% for the *COI* data and 14.4% - 19.31% for the *Cyt-b* data (Table 3).

Pairwise F_{ST} values indicated that level of genetic differentiation was related —in most cases— to geographical distances. This was also apparent in the Mantel tests, which were highly significant for all markers (Appendix S15). This was more evident for *COI*, where F_{ST} values were mostly low between sites within drainages. For example, there was very little genetic differentiation (F_{ST}≈0) between populations within the Esmeraldas drainage as well as between the Esmeraldas and geographically neighboring Atacames and Sua drainages, except for the sites E9, E3, and E8 that showed moderate (F_{ST}≈0.15) to very great (F_{ST}>0.25) genetic differentiation compared to neighboring populations within the drainage. In addition, E8 (southernmost site of

Esmeraldas) shared very little genetic differentiation (F_{ST}=0) with populations from the northernmost sites of Guayas: G1, G2, G6, and G7. The latter sites showed very great (F_{ST}>0.25) genetic differentiation with population from the Southern part of the Guayas drainage. This may suggest an apparent north-south divergence within the Guayas drainage —including E8 from Esmeraldas— such that two groups are formed in the drainage. The northern group consisting of E8, G1, G2, G6, and G7, while the southern group consists of G5, G3, and G4. Each group showed very little genetic differentiation among sites (F_{ST}<0.05) but very great differentiation among groups (F_{ST}>0.25). G5 and G3 also differed significantly from the site Ta located at the south next to the Guayas drainage while G4 showed moderate differentiation (F_{ST}=0.09) with this site. Also, G5 and G3 from the southern group of Guayas seem to have some relationship with individuals from the geographically neighboring site SP and the high elevation J4 and J5 sites from Jubones (Table 4). On the other hand, Santa Rosa, the southernmost drainage sampled in this study showed very great genetic differentiation with the rest of the sites except with its closest neighbors J1 and J2, which showed moderate differentiation (F_{ST}≈0.14). Finally, sites J1, J2 and J3 (low elevation) were genetically similar but differed from J4 and J5 (high elevation), forming two genetically divergent groups segregated by elevation within the same river (Table 4).

Cytochrome b F_{ST} values showed considerably higher genetic differentiation among populations between and within the drainages as expected due to its higher mutation rate. For example, pairwise F_{ST} values between populations from the Esmeraldas drainage showed little (F_{ST}<0.05) to mostly moderate (F_{ST}<0.15) genetic differentiation except for E2B and in some cases, E9, which differed substantially (F_{ST}≥0.15) (Table 5). Atacames and Sua differed very little from one another but showed very great genetic differentiation to all other populations. The *Cyt-b*

data also showed some evidence of the formation of two groups in the Guayas drainage distinguishing the northern area from the southern area as observed with the *COI* data. Sites from the northern Guayas group were very similar only to their closest neighbors (i.e. G6≈G7≈G1≈G2 but G6≠G2), contrasting with the *COI* results, for which all sites within this group were very similar to one another. For the southern group of Guayas, G3 showed moderate (F_{ST}=0.13) genetic differentiation with G4 and little differentiation (F_{ST}=0) with Ta. The next drainage south, San Pablo, showed great to very great genetic differentiation with all populations, however, the F_{ST} values with J4 and J5 (sites of high elevation, Jubones) were lower (F_{ST}=0.22 and F_{ST}=0.25 respectively) than for the other sites (F_{ST}≥0.4). As was the case for the *COI* data, the Jubones populations appear to be divided into two groups by elevation, such that J1 exhibits little differentiation with J2 and moderate genetic differentiation with J3 but the three sites together exhibit very great differentiation from J4 and J5, suggesting a possible barrier within this drainage. Finally, the Santa Rosa drainage only presented moderate differentiation from J1 and J2 (F_{ST}=0.08 and F_{ST}=0.13, respectively) and very great differentiation from the rest of the populations (F_{ST}≥0.3) (Table 5), consistent with the *COI* data.

The F_{ST} values for the microsatellites followed a similar pattern to those for the mitochondrial genes. Atacames presented great to very great genetic differentiation when compared with the rest of the sites. Sites within the Esmeraldas drainage showed moderate to great genetic differentiation, except for E2B which was very similar to E4, E8, and G1 (from Guayas) (F_{ST}≤0.05), but moderately different from E6 (F_{ST}=0.14). As observed with *Cyt-b* and *COI*, G1 from north Guayas was very similar to E8 (the southernmost site of Esmeraldas) but also to E2B located at the northernmost part of Esmeraldas. G1 was also similar to all the other sites

from Guayas (i.e., G2, G3, and G4). However, G2 showed moderate genetic differentiation with G3 and G4. On the other hand, SP showed moderate genetic differentiation with G1 and G3 from Guayas, and J4 and J5 from the highest elevation from Jubones. The latter sites were very similar to each other (F_{ST}=0.04). J1, J2, and J3 showed moderate to little genetic differentiation but very great genetic differentiation with J4 and J5. Finally, SR showed moderate genetic differentiation with sites from the lowest elevation of Jubones, J1, J2, and J3 ($0.06<F_{ST}<0.1$) and great to very great genetic difference with J4 and J5 (F_{ST}=0.24 and 0.29 respectively) (Table 6).

Phylogenetic Analysis

Populations of *Rhoadsia* formed a monophyletic group relative to their closely related species, *Carlana eigenmanni* and *Parastremma sadina*. The two outgroups clustered by genus (Fig. 6-7, Appendix S16-S17). Sequences of *C. eigenmanni* formed two well supported clades, one represented by samples from Costa Rica and Nicaragua and the other by samples from Panama. The four sequences of *P. sadina* —all from Colombia— were identical to each other so they were collapsed into one sequence forming a single branch. According to the ML phylogenetic trees for both genes, the outgroup species appear to be more closely related to samples from the Southern region of Western Ecuador than to the northern region (Appendix S16-S17).

Phylogenetic analysis of the mitochondrial genes using Bayesian methods identified two main clades within *Rhoadsia,* clustering sites from northern (Clade I) and Southern (Clade II) Ecuador (Fig. 6-7). Maximum Likelihood analysis showed only Clade I for the northern sites as a well-defined cluster that branch out from the rest of individuals, which fall into multiple subclades from the Southern region (Appendix S16-S17). Topologies inferred using *Cyt-b* and *COI* data yielded similar results. For the *Cyt-b* gene tree, Clade I had three well supported subclades

including *subclade a* clustering exclusively the northernmost sites: At and Su, *subclade b* with samples from the northern sites of the Guayas drainage: G6, G7, and G1, and *Subclade c* with samples from the Esmeraldas drainage: E2B, E4, E9, E8, and E6. Subclade b forms a paraphyletic group with samples present in Clade II (Southern region). Other specimens within Clade I presented negligible support (posterior probability <0.5) including sites from Guayas drainage: G1, G2, G6, and G7, and from the Esmeraldas drainage: E4, E6, E8, and E9. On the other hand, Clade II presented seven subclades. Sites J4 and J5 formed a unique subclade (h) that included some individuals from SP. Sites J1, J2, and J3 also from the Jubones drainage formed two subclades within the Clade II: *Subclade d*, which included individuals from these sites and specimens from the Santa Rosa river, and subclade *g*, which included a few individuals from sites J1 and J2. Other sites like G3, G4, and G5 from Guayas formed the *subclade i*. Individuals from G3 formed another well supported subclade with Ta, identified here as *subclade f.*

For the *COI* gene tree, clade I was also formed by populations mainly belonging to the northern region of Ecuador including samples from Sua, Atacames and Esmeraldas, but also some samples from the northernmost region of Guayas such as G1, G2, G6, and G7. There were two subclades with high support: *subclade a* which comprised some individuals from E9 and E3, and subclade b that included mostly individuals from the Guayas region and 5 individuals from E8. The rest of the clade I individuals included mostly sites from Esmeraldas and a few from the northern Guayas and did not fall into discrete subclades. In addition, they mostly corresponded to haplotype 3, one of the most abundant haplotypes for *COI* (Fig. 7). Clade II did not show strong support (pp=0.38) as observed for the *Cyt-b data*, however, it was also composed of populations from the Southern region except for two samples from the Esmeraldas drainage, E6 and E8.

Within Clade II, there were 5 subclades and some clustering of specimens with negligible support. As observed in clade I, there was not a monophyletic cluster that encloses a site or group of sites within a clade (i.e., sites were dispersed among multiple subclades) (Fig. 7).

Haplotype Network

The haplotype network for *Cyt-b* and *COI* were similar in general topology. Haplotypes from the northern region appear to be separated from the southern haplotypes by three mutations for *Cyt-b* and one mutation for *COI*. For *Cyt-b*, H9, one of the most common haplotypes (freq. 0.29) (Appendix S12), was the only haplotype shared between sites from the northern and Southern regions (Fig. 8, Appendix S18). It was present in almost all sites from Esmeraldas except E2B, and in the four northernmost sites of Guayas: G1, G2, G6, and G7 (Fig. 4). Besides the H9 haplotype, there were four more haplotypes present exclusively in all sites from Esmeraldas: H10, H11, H19, and H21, of which H11 was also one of the most frequent haplotypes (freq. 0.28) (Appendix S12). Atacames and Sua, the northernmost sites sampled, had their own haplotypes (H12 and H22) separated by three mutations from H9. Other haplotypes from the northern Guayas (i.e., G1, G2, G6, and G7) such as H15, H18, and H20 were separated by few mutations from H9, reaffirming the relationship of these sites with the northern region. The southern populations presented 18 haplotypes. Most of them were connected to H5 which was present in the sites J1 and J2 from Jubones and in one individual from Ta. Haplotypes from Southern Guayas (i.e., G3, G4, and G5) were mostly private to this drainage and were present at lower frequencies. Most of them were separated by a few mutations (1 to 5) from H5. Other haplotypes connected to H5 were H6 and H23 from SR, H3 from SP, H8 from J1 and J2, and H1 from J1, J2, J3, and SR. The latter haplotype occurred at one of the highest frequencies (31%) in the Southern region. On

the other hand, J4 and J5 consisted of only one haplotype, H17 which was shared with SP. This haplotype was separated by one mutation from H7, a haplotype comprised by sites from G1 and G2 (Fig. 8).

The haplotype network for the *COI* gene showed H3 as the most common haplotype present in the northern region (Fig. 9, Appendix S19). This included all sites from Esmeraldas, Atacames, Sua and three of the northernmost sites from Guayas, G1, G6, and G7, although it occurred at very low frequencies in the three latter sites. Haplotype H1 was also shared by one site from Esmeraldas (E8) and the northernmost sites from Guayas (i.e., G1, G2, G6, and G7) (Fig. 5a, Appendix S13). Although haplotype H1 was present mostly in populations from Guayas, it was separated by fewer mutation of haplotypes from the northern region, i.e., it was more closely related to haplotypes from the Esmeraldas, Sua, and Atacames drainage than to haplotypes from Guayas or any other drainage from the Southern region (Fig. 9). H4, which was present almost exclusively in southern populations (SR, J1, J2, J3, Ta, G1, and one individual from E6), was the closest southern haplotype to H3, separated by one mutation. H6 was another southern haplotype that was present in one individual from Esmeraldas (E8) along with J1, J4, and SR. Finally, H7 was another very common haplotype from the Southern region and occurred in sites from Guayas like G2, G3, G4, G5, and SP, and the three highest sites from Jubones, J3, J4, and J5 (Fig. 9).

Population Structure Analysis with Microsatellites

The population structure analysis based on the 12 microsatellites suggested the presence of 10 genetic clusters of *Rhoadsia* in western Ecuador. The "true" value of K (i.e. K=10) was supported by both methods used. The Evanno method showed the greatest ΔK (ΔK=66.27) at K

10 (Appendix S20-S21). Accordingly, the value of LnPK from the Pritchard method formed a plateau at K 10 as well (Appendix S22). The clusters estimated by STRUCTURE were as follows: At (Cluster 1), E2B and E4 (Cluster 2), E6 (Cluster 3), E8 (Cluster 4), G1 (Cluster 5), G2 (Cluster 6), G3 and G4 (Cluster 7), SP, J4, and J5 (Cluster 8), J1, J2, and J3 (Cluster 9), and SR (Cluster 10) (Fig. 10). The Structure analysis suggested some degree of population admixture in some sites. The most notorious example was observed between G1 and sites from the Esmeraldas drainage and Guayas drainage. Although, the clustering designation appeared to correlate with the distance among sites in most cases, Cluster 8 (SP, J4, and J5) linked populations from different drainages with no apparent fluvial connection. In contrast, At represented by Cluster 1 seemed to be very distinctive despite its relatively close proximity to its neighboring sites (i.e. E2B and E4) (Fig. 10).

A neighbor-Joining tree based on the 10 clusters estimated by STRUCTURE was created to show the relationships among the clusters. The 10 clusters fell in to three main groups. The first group (I), the northern group, was represented by all the sites from the Esmeraldas and the Atacames drainage (clusters 1-4). The second group (II), the central-south group, comprised all the sites from the Guayas drainage plus SP and the two sites located at the highest altitude of the Jubones drainage, J4 and J5 (clusters 5-8). Finally, the third group (III), the southernmost group, enclosed SR and J1, J2, and J3 from the lowest elevation of the Jubones drainage (clusters 9 and 10). The most divergent clusters were to clusters 1 (At), cluster 3 (E6), and cluster 8 (SP, J1, J2, and J3) (Fig. 10).

The Neighbor-Joining of the populations based on the microsatellite data showed high support for clustering of all northern populations (0.73), excluding E8. The rest of the populations

formed a well-supported clade (0.89) which included the central-south populations (from Guayas to Santa Rosa), where G3 and G4 were closely related to each other and with SP, J4, and J5. Likewise, SR was closely related to J1, J2, and J3 (Fig. 11). The neighbor joining analysis per individuals was consistent with this result (Appendix S23).

Discussion

In this study, I present a description of the phylogeographic relationships of *Rhoadsia* in Western Ecuador, in an attempt to investigate three main objectives. The first objective was to elucidate the validity of the two species currently recognized for this genus (*R. minor and R. altipinna*). The phylogeographic analyses based on both mtDNA genes (*Cyt-b* and *COI*) (Fig. 6-7) and the STRUCTURE analysis with microsatellites appear to confirm the existence of two species, with introgression of genetic variation of *R. minor* into *R. altipinna* in the northern Guayas drainage (Fig. 4-5, and Fig. 10). Southward, the rest of sites sampled showed genetic characteristics consistent with the presence of one species exhibiting geographic population structuring, *R. altipinna*. The microsatellite data showed an overall finer structuring pattern of *Rhoadsia* populations in western Ecuador, with the identification of ten genetic clusters divided into three groups (Fig. 10), although the mtDNA data supported these general patterns. Some of the genetic clusters identified appear to be very distinctive and I explore their relevance for future conservation efforts.

The second objective was to investigate the most likely geographical origin of *Rhoadsia* in Western Ecuador based on the patterns of genetic diversity of the populations of *Rhoadsia* and the phylogenetic distances with two closely related genera (*Parastremma* (Colombia) and *Carlana* (Central America)). Although the results were somewhat inconclusive, I hypothesized that the most likely geographical origin of *Rhoadsia* is somewhere in the Guayas drainage. Some sites from the Guayas drainage exhibited the highest genetic diversity (Table 2). Likewise, the outgroup appeared to show a closer relation to sites from the Guayas drainages according to the ML phylogenetic tree (Appendix S16-S17).

The third objective was to document potential evolutionary breaches among drainage basins in Western Ecuador based on patterns of genetic divergence in *Rhoadsia*. The microsatellite data showed that *Rhoadsia* may be divided into three main groups with two main breaches: one located between the Esmeraldas and the Guayas drainage and the second one in the Southern region between the high-elevation and the low-elevation sites of the Jubones drainage (Fig. 10). However, I discuss the possibility of permeability in the first breach (Esmeraldas-Guayas) that might lead to the observed introgression of *R. minor* from the Esmeraldas drainage into the northern Guayas drainage. Also, I suggest a potential explanation for the appearance of an important breach for *Rhoadsia* located somewhere in the middle of the Guayas drainage where there is no apparent physical barrier.

Species Delimitation and Evolutionary Relationship of *Rhoadsia* in Ecuador

Delimiting species is a difficult task, considering the large number of species concept that exist. All concepts (e.g., biological, phylogenetic, morphological, ecological, etc.) have limitations, preventing their implementation to all groups of organisms (Aldhebiani, 2018). To overcome this issue, De Queiroz (2007) proposed the general lineage concept (GLC) of species. It encloses the idea of a unified species concept, where all species concepts are treated as properties of the speciation process. Although this approach has significant merit, different species concepts continue to be employed in practice (Aldhebiani, 2018; Bagley et al., 2016; Costa-Silva et al., 2015).

Historically, species of fishes has been described and delimited using morphological data, applying meristic and morphometric analysis (Stauffer et al., 2002a; Stauffer et al., 2002b).

Molecular methods are now commonly implemented such as the use of DNA barcodes like *COI* as a method to distinguish species (Kress et al., 2015; Barrett and Hebert, 2005; Hebert et al., 2003). Although some authors have argued that this method may be unreliable in some situations (Hebert and Barrett, 2005; Prendini, 2005), the use of DNA barcodes to delimit species has garnered significant support and is commonly used in combination with other genetic markers (e.g., nDNA genes, microsatellites, etc.) (Bagley et al., 2016; Costa-Silva et al., 2015; Sunnucks, 2000) or with morphological and behavioral data (Vernygora et al., 2018; Stauffer et al., 2002a).

The present study leaned towards a genetic-based approach where more than one genetic data-type (i.e., mtDNA and microsatellites(nDNA)) will be used to delimit the species. It is worth noting, this study did not rely on a particular species concept (e.g., phylogenetic species concept) for delimiting species. Phylogenetic relationships of species based on single mtDNA genes like *COI* and *Cyt-b* can cause biases that could mislead the interpretation of the evolutionary history of species (Knowles and Carstens, 2007; Hickerson et al., 2006; Shaw, 2002). Although reciprocal monophyly appear to be a self-evident indicator to delimit species, the lack thereof does not necessarily mean absence of species divergence (Knowles and Carstens, 2007). Fortunately, comparative analysis on multiples types of genetic data help to overcome these biases in delimitation (Huang and Knowles, 2016; Shaw, 2002). In this study, I take advantage of both mtDNA and the 12 microsatellites to infer the history of *Rhoadsia*. More specifically, I implemented a phylogeographic analysis to determine the number of valid species of *Rhoadsia* based on the recognition of independently evolving metapopulations (i.e., multiple populations of a species that coalesce into a common ancestor) (Reeves and Richards, 2011; De Queiroz,

2007). This approach has been quite useful when dealing with genetically cryptic species (Schield et al., 2018; Arteaga et al., 2016; Fennessy et al., 2016; Bagley et al., 2016; Unmack et al., 2012).

Genetic markers are fundamental for phylogeographic studies as well as for the resolution of phylogenetic hypotheses (Avise and Riddle, 2009; Sunnucks, 2000; Avise et al., 1987). Genetic markers from the mtDNA and nDNA are both widely used for evolutionary and systematics studies (Kress et al., 2015; Zink and Barrowclough, 2008; Selkoe and Toonen, 2006; Balloux et al., 2002; Farias et al., 2001; Avise et al., 1987). However, due to differences in their intrinsic characteristics (e.g., microsatellites from nDNA with a faster mutation rate than mtDNA genes, mtDNA is uniparental while nDNA is biparental), some authors have reported different results between analyses conducted on mitochondrial and nuclear genes, also known as mito-nuclear discordance (Toews and Brelsford, 2012). Although these kind discordances can be problematic for systematics and taxonomy (Ivanov et al., 2018; Shaw, 2002), they can be useful to detect phylogeographic processes such as recent or ancient hybridization events (Tóth et al., 2017). In this study, the mtDNA *Cyt-b* and *COI* genes appear to show evidence of introgressive hybridization between the currently recognized species, with haplotypes common in the northern region (*R. minor*) also present in the Southern region (*R. altipinna*) (Fig. 4-5), which complicates the delimitation of the species. Although the microsatellite data split populations in a matter that was more concordant with the traditionally recognized species ranges, they also showed evidence of admixture in the locations where the introgression of the mtDNA was observed (Fig. 10) (see below). The microsatellite data also resulted in a more complex pattern of population genetic structuring than the mtDNA data, indicating the existence of ten populations divided into three clusters that generally segregated geographically (Fig. 10).

Populations of *Rhoadsia* in Ecuador were split into two main lineages inhabiting the northern and a Southern region of western Ecuador (Fig. 6-9). The two phylogroups seem to generally coincide with the two species of *Rhoadsia* currently recognized in Ecuador, i.e., *Rhoadsia minor* (north) *and R. altipinna* (south), with some admixture near the limits of the species ranges. The distinctive clade formation was more apparent with *Cyt-b than with COI*. The latter presented considerably lower branch support particularly in the southern clade (Fig. 7). This negligible support in the *COI* data could be attributed to the lower rate of evolution of this barcoding gene (Kress et al., 2015), preventing a clear detection of more recently diverged populations. The STRUCTURE analysis of the microsatellite data on the other hand presented a finer subdivision of the populations of *Rhoadsia* forming ten clusters organized into three groups (I, II, and III). Group I represented by *R. minor* and both Group II and III represented by populations of *R. altipinna* (see below). It showed evidence consistent with genetic admixture occurring at the G1 site (Northern Guayas) between species (Fig. 10). Given the current separation of the drainages (Esmeraldas and Guayas) this pattern may suggest a possible historical interaction between the two species.

Diversification of *Rhoadsia* in Western Ecuador

Divergence pattern of *Rhoadsia* appears to be too complex to simply delimit two species as a solution to the problem (*R. minor* and *R. altipinna*). Species categories may be difficult to determine due to subjective viewpoints that can obscure some aspects of the diversification process (Willis, 2017). Systematic corrections are continuously being proposed regarding the number of species and their distribution, as better approaches are implemented (Fennessy et al.,

2016; Bagley et al., 2016). However, when differences among populations are not discrete, lower-level denominations for species (e.g., subspecies, variety, morph, evolutionary significance units and management units) may be useful for conservation purposes (Palsbøll et al., 2007; Waples, 1995; Moritz, 1994; Ryder, 1986), with the caveat that they are not considered part of the formal hierarchy of biological classification (except the subspecies category) (The International Commission on Zoological Nomenclature, 2000) and are often assigned arbitrarily (Gippoliti and Amori, 2007; Ryder, 1986; Wilson and Brown, 1953). The term evolutionary significant units or biological units is applied to populations that are genetically distinctive enough (e.g., presence of rare alleles due to a complete isolation or substantial genetic differentiation) to be considered separate populations (Palsbøll et al., 2007; Moritz, 1994; Ryder, 1986). When the isolated population seems rare and/or endangered, the evolutionary significant unit could also be treated as a management unit in which conservation efforts can be enforced (Palsbøll et al., 2007). *Rhoadsia* appears to present multiple genetically isolated populations that may be considered as evolutionary significant units or even management units within each named species.

A Bayesian-based analysis of the microsatellites showed that the species are subdivided into at least ten genetically structured populations divided into three main groups (I, II, and III) (Fig. 10). Group I was comprised by the northern populations (*R. minor*) At (Cluster 1), E2B and E4 (Cluster 2), E6 (Cluster 3), and E8 (Cluster 4). Group II contained G1 (Cluster 5), G2 (Cluster 6), G3 and G4 (Cluster 7), SP, J4, and J5 (Cluster 8). Group III was comprised of J1, J2, and J3 (Cluster 9), and SR (Cluster 10). Genetic similarities found in the populations from the groups II and III suggest that they may be subdivisions of *R. altipinna*, although with significant genetic

differentiation. The ten structured populations were supported by the AMOVA which showed the greatest percentage of variation for the among-group component (when groups were organized by the K=10 populations), according to both mtDNA genes (75% and 54% for *Cyt-b* and *COI* respectively), but not the microsatellite (Table 3). The populations and major clusters also seemed to align with the geographical history of the drainages in western Ecuador. That is, the ten genetically distinct populations seem to be formed following the natural barriers found among the rivers within and among the drainages (Fig. 10). Some of the most isolated populations were evident with the mtDNA data as well, shown as subclades in the phylogenetic analysis (Fig. 6-7) and distribution of haplotypes (Fig. 4-5).

Contrary to what the mtDNA data showed, microsatellite data suggest that most of the genetic variation (80%) is evenly distributed within populations, with much smaller components of genetic variation segregating among populations within drainages (4%) and among drainages (17%) (Table 3). A possible explanation for this inconsistency among genetic markers could be the susceptibility of microsatellite data to show homoplasy as consequence of saturation of mutations (for highly divergent populations) and/or the type of mutation model (e.g. IAM, KAM, SMM, or GSM) leading to biased estimates (Jarne and Lagoda, 1996). F_{ST}-based analyses (e.g., AMOVA) can be affected by this type of phenomena, which is why genetic differentiation results based on microsatellites are often taken with caution. This also would explain the overall low values in the pairwise F_{ST} matrix based on the microsatellites data compared to the mtDNA markers. R_{ST} (Slatkin, 1995), an alternative measure for genetic differentiation (relative to F_{ST}) often recommended for microsatellites (Balloux et al., 2002) was also used. However, it was

discarded due to erratic results that were incompatible with the geographic distribution of the sites or the mitochondrial data (data not shown).

Evolutionary Significant Units of *Rhoadsia minor*

Sua (Su) and Atacames (At) form one of the most distinctive genetic clusters (cluster 1). They are located in northwestern Ecuador just north of the mouth of the Esmeraldas drainage. Although they inhabit two different drainage basins, they shared a similar genetic makeup, probably due to the existence of a passage that still allows for gene flow between the populations (perhaps during sporadic flooding) or a recent separation of the drainages. Su and At are a strongly supported subclade (*a*) within the northern clade and significantly divergent from other northern clade populations according to the *Cyt-b* and microsatellite data (Fig. 6; Appendix S23). The Atacames drainage shares its eastern border with the Esmeraldas drainage. However, the Atacames and Esmeraldas populations seem to be very different despite their geographic proximity according to *Cyt-b* data (F_{ST}>0.7) and microsatellite data (F_{ST}>0.18). At and Su also presented two non-synonymous mutations in *Cyt-b* that were not found in any other populations of *Rhoadsia* suggesting substantial time since divergence or a genetic bottleneck that may have allowed for the rapid fixation of new amino acids. However, there was no significant difference between the Su-At subclade and the Esmeraldas populations for the more conserved *COI* (F_{ST}=0). Haplotype H3 for *COI* was dominant and shared between Su, At and all the sites from Esmeraldas, suggesting that these populations belong to the same species.

The Su-At subclade was among the lowest in terms of genetic diversity and allelic richness (Table 2, Fig. 3), perhaps due to the small size of the drainage basins they inhabit or genetic drift.

The high level of differentiation, low genetic diversity and rareness of the alleles presented in these sites should be sufficient to assign this genetic cluster as an evolutionarily significant unit. Given the small size of the drainage basins and the environmental degradation of freshwater ecosystems in the region, it is also likely a priority for conservation studies.

Esmeraldas, mostly comprised of panmictic populations of *Rhoadsia minor* with some genetically distinctive populations

The overall genetic makeup among the samples from the Esmeraldas drainage seems to be shared among populations, suggesting sustained gene flow or very recent divergence. Most of the sites exhibited similar haplotypes but at different frequencies (Fig. 4-5). Genetically divergent sites were geographically distant from one another. The most divergent sites were also often located at high elevations. This is consistent with a pattern of isolation by distance in which populations are more likely to exchange genes with neighboring populations than with distant ones. Likewise, populations located at different elevations are influenced by distinctive biotic and abiotic factors for which they may be better adapted (Malato et al., 2017; Vannote et al., 1980), affecting the frequency of migrants among populations, hence, limiting gene flow. Two of the highest sites Mindo (E6) (Cluster 3) and Transito (E8) (Cluster 4), located at 1260 meters above the sea level (m.a.s.l.) and 1093 m.a.s.l., respectively, were significantly divergent according to the Structure analysis conducted on the microsatellites (Fig. 10). In contrast, Teaone (E2B) and Hermoso (E4) were similar enough to be part of the same genetic cluster, Cluster 2 (Fig. 10). Assuming no physical barriers between the rivers neighboring clusters 2, 3 and 4, it would be likely that the surrounding sites (analyzed with mtDNA genes) would form part of one of the three main clusters present in the Esmeraldas drainage basin depicted in Fig. 10 in accordance

with their geographic location. Consistent with its lower rate of evolution, *COI* showed very low genetic differentiation (F_{ST}=0) for most sites from Esmeraldas except E8, E9, and E3 (Table 4). It is not clear why E9 and E3 showed such high differentiation since they appear to be relatively close and are located at similar elevations, although it is possible that the small sample size (N=8) may have influenced the result.

Evolutionary Significant Units of *Rhoadsia altipinna*

Northern Guayas populations form a transition zone between *R. minor and R. altipinna*

Individuals of *Rhoadsia* from the northernmost part of Guayas, here represented by the sites (north to south) Chiguilpe (G6), Otongo (G7), Palenque (G1), and Jaunache (G2), displayed haplotypes characteristic of both *R. minor* and *R. altipinna*. Most individuals from these sites (89% *Cyt-b*, 84% *COI*) showed mitochondrial haplotypes that were the same or closely related to haplotypes also found in the Esmeraldas drainage (Fig. 4-5, 8-9). The mtDNA thus appears to indicate that genes characteristic of *R. minor* are most abundant in the northern Guayas. The remaining individuals corresponded to haplotypes more similar to haplotypes found in the southern group (Fig. 8-9). The microsatellite data on the other hand, showed that cluster 5 represented a greater proportion of individuals from G1 —the northernmost site in the Guayas drainage analyzed for this genetic marker— with 66% of membership. Eighteen percent of the microsatellite allele makeup of Cluster 5 corresponded to that of sites from the northern region, with the greatest membership in site E4 (12%). The remaining 16% of Cluster 5 was evenly distributed across the southern samples. G1 (besides containing 66% of cluster 5) also contained

about 14% of the genetic makeup corresponding to alleles with characteristics of *R. minor* (northern region), especially from site E8 (Cluster 4) with 8.7% of membership. The remaining 20% of membership in G1 was represented by the southern clusters specially cluster 6 (G2) with 9.2% followed by cluster 7 (G3 and G4) with 4.5% (Fig. 10, Appendix S24). This cluster also grouped more closely with other clusters of *R. altipinna*. The microsatellite data thus indicate that individuals from northern Guayas contain alleles and mtDNA haplotypes characteristic of both species. It is possible that both species diverged in isolation a long time ago forming northern and southern species and there has been introgressive hybridization in this area where the ranges meet (i.e., northern Guayas). The individuals in this region may thus be descendants of hybrids with the mtDNA haplotypes from the northern species having risen in frequency to be the dominant haplotypes, while the nuclear genome remains more similar to that of *R. altipinna* (Fig. 10-11). Mitochondrial and nuclear DNA discordance is commonly observed in populations of animals that have gone through a secondary contact (Toews and Brelsford, 2012). However, the discordance can be a product of variety of other processes like adaptive introgression of mtDNA, sex-biased asymmetries, and demographic disparities (Toews and Brelsford, 2012). A more detailed study is required to investigate the causes of the apparent introgression documented.

Southern Guayas, genetically distinct from northern Guayas

Sites from the northern Guayas seem to genetically differ from sites from the Southern Guayas (i.e., Caluma (G5), Chague Grande (G3), and Chimbo (G4)) according to all genetic markers, although specially for the mtDNA genes (Table 4-5). The STRUCTURE analysis with microsatellite data (G5 excluded due to sample size, N=2) identified these sites under cluster

seven (Fig. 10), with 91% and 78% of membership for G3 and G4 respectively (Appendix 20). The level of genetic differentiation may be due to the presence of a natural barrier caused by any number of factors. The Guayas river drains in an estuarine ecosystem before arriving to the sea (river mouth of the Gulf of Guayaquil). It harbors many estuarine species well adapted to such environment (Twilley et al., 2001), and serves as an outlet for many tributary rivers coming from different parts of the Guayas drainage. Unfortunately, the Guayas river is also known to have very low water quality in some of its larger branches, which is associated with intense agriculture, proximity to large human populations, and general habitat degradation (Damanik-Ambarita et al., 2016). The role of the biotic and abiotic factors are crucial in the distribution of communities of freshwater fishes (Jackson et al., 2001). It is possible that these conditions might serve as a stressor to some species like *Rhoadsia* leading to its isolation in upstream rivers, hence, reduction of gene flow. Another hypothesis could be that the separation of this northern and southern populations in Guayas is associated with phenological patterns (influence of seasonal climate on plant and animal life) (Chuine, 2010) in Western Ecuador (MAE, 2012). This distribution pattern of *Rhoadsia* may be reflecting similar patterns in other species of freshwater fishes in the Guayas drainage. If so, the formation of an Ichthyo-hydrographic breach within the Guayas drainage could be plausible (See below).

Taura, genetically similar to neighboring sites

A sample of two individuals from the Taura basin showed similar haplotypes with sites from Esmeraldas, Guayas and low-elevation Jubones. Taura showed overlapping haplotypes with the sites G3 (Haplotype 4) and low-elevation sites from Jubones Huizo (J1) and Cascay (J2) (Haplotype 5) according to the *Cyt-b* data, and with E6 from Esmeraldas, G1 from northern

Guayas, and J1, J2, and Mollopongo (J3) from low-elevation of Jubones according to *COI* (Haplotype 4). Taura also presented one private haplotype (H18) for *COI.* Although the small sample size impedes making conclusive inferences, the genetic composition of this site seems to be consistent with that of neighboring populations, suggesting close relations to other Southern Guayas populations.

San Pablo and high-elevation Jubones are closely related

Continuing south from the Taura drainage, there are various small basins that drain east to west into the Pacific Ocean. Despite their current isolation, historical connections may have allowed for gene flow among these neighboring basins, given their similarities in species communities (Barriga, 2012). However, a population from the San Pablo (SP) basin separated by about 4 basins from Guayas, showed signs of genetic isolation from its northern neighbors except with G5 and G3 (Table 4). Likewise, the STRUCTURE analysis showed that about 4.5% of the genetic membership of SP is similar to cluster 7 that correspond to the sites G3 and G4 from the Southern Guayas (Appendix S24). This suggest evidence of an ancient connection from the Southern Guayas into the SP population or shared ancestry. On the other hand, SP seems to be particularly similar to the sites Minas (J4) and Mondur (J5) according to *Cyt-b* and the microsatellite data (Table 5 and 6). This is interesting since J4 and J5 are located about six basins apart to the south at 909 m and 1095 m of elevation respectively in the Jubones drainage. However, the genetic composition of SP, J4, and J5 is very similar, and these populations show great genetic differentiation from other populations. The structure analysis placed these three sites into Cluster 8 (Fig. 10), suggesting they may be part of a genetically distinctive population that split relatively recently. I suggest that Cluster 8 should be considered as an evolutionary

significant unit due to its genetic divergence, although it is expected that the obvious geographic separation between SP and the high-elevation sites from Jubones will allow those populations to take independent evolutionary paths in the future.

Low-elevation and high-elevation Jubones, ancient divergence or independent origins?

The Jubones drainage located at the Southern Ecuador is segregated into two distinct genetic clusters based on elevation. The three lowest sites sampled (i.e., J1, J2, and J3) formed their own cluster (Cluster 9) according to the STRUCTURE analysis of the microsatellite data. Likewise, as mentioned above the two highest sites (i.e., J4 and J5) formed a separate cluster with the site SP (Cluster 8) (Fig. 10). This pattern was also observed with the *Cyt-b* data (Fig. 4) but was less obvious with the *COI* data, for which some haplotypes seem to be shared between the sites J3, J4, and J5 (Fig. 5a). It is worth noting that there was a relatively large gap in elevation between the sites J3 (251 m.a.s.l.) and J4 (909 m.a.s.l.). The sampling area between these sites was difficult to access or when it was possible, there was no *Rhoadsia* present. It is possible that the separation of the populations may be influenced by natural barriers like the presence of waterfalls or steep slopes with rapid water. These kind of barriers are known to have a crucial effect in the segregation of fish assemblage (Silva et al., 2016). The level of differentiation between the two clusters is remarkable (Table 4-6), given that they inhabit the same relatively small river basin (Appendix S2). The phylogenetic analysis reveals an apparent independent divergence of the two populations. Current gene flow between the low and high elevation samples seem unlikely due to the magnitude of the genetic divergence in the *Cyt-b* and microsatellite data (Fig. 4 and Fig. 10). The *COI* data, however, did show some overlapping

haplotypes among some individuals between the two clusters (Fig. 5a). Given the slower evolution rate of *COI*, the presence of similar haplotypes is expected even among distantly related populations. An example of this was observed with the haplotype H4 present in multiple sites from different distant drainages such as E6 (Esmeraldas), G1 (Guayas), Ta (Taura), J1, J2, J3 (Jubones), and SR (Santa Rosa). Another example was the haplotype H6 present in sites E8 (Esmeraldas), J1 and J4 (Jubones) (Fig. 5a). The genetic distinctiveness of the low-elevation and high-elevation populations in the Jubones drainage may be worth further investigation regarding potential evolutionary adaptations to their local environments.

Santa Rosa, closely related to low-elevation populations of Jubones but independent enough to be considered an evolutionary significant unit

The Santa Rosa drainage (SR), located just south of the Jubones drainage, presented overlapping haplotypes with the lowest sites of Jubones. However, according to the microsatellite data, this site appears to form its own genetic cluster (cluster 10) having only a 5.8% of membership with Cluster 9 that correspond sites J1, J2, and J3. Despite the presence of shared haplotypes with the low-elevation population of Jubones for both mtDNA genes, the SR drainage appears to have its own genetic identity and may deserve to be considered as an evolutionary significant unit. Significant divergence in haplotype frequencies reflecting genetic isolation of this population was previously documented by Aguirre et al. (2016) ,emphasizing the potential of the SR population for local adaptation.

Possible Geographic Origin of the Genus *Rhoadsia*

The second main objective of this study was to investigate the most likely geographical origin of the genus *Rhoadsia* by using two different sources of evidence. The first involved measuring levels of genetic diversity among populations under the assumption that older, ancestral populations would be more diverse genetically than the daughter populations they gave rise to. The justification for this is that ancient populations would have a longer genetic history, hence, more alleles compared to more recent populations. An example of this would be the origin of humans in Africa, with African populations typically being more genetically diverse than populations in other regions in the world (Ingman et al., 2000). This pattern is also a reflection of the core-periphery hypothesis where more recent populations are expected to be found at peripherical areas with less genetic variability compared to core populations because they originate from relatively small numbers of individuals carrying little genetic diversity with them (Duncan et al., 2015). It is worth nothing that high diversity could be also attributed to a large population size which is usually more likely to be found in larger areas (Preston, 1962). For this reason, a higher genetic diversity is not necessarily direct evidence of geographical origin although it is consistent with it. The second source of evidence involved analyzing phylogenetic relationships relative to the outgroups (other members of the subfamily Rhoadsiinae: *Carlana eigenmanni* and *Parastremma sadina*). The population of *Rhoadsia* closest to the outgroups could be interpreted as the population with more ancestral characteristics, hence, more likely to resemble the original population from which the genus originated.

Guayas the most diverse and the most likely place of origin of *Rhoadsia*

The Guayas drainage exhibited the greatest allelic variation with high levels of haplotype richness mostly represented by rare alleles according to the *Cyt-b* data (Fig.5a). Thirteen of the 15 haplotypes were exclusive to the Guayas-Taura drainage (Taura was combined with Guayas due to its proximity and low sample size (N=2); one of the 15 haplotypes was private to Taura). That is, 87% of *Cyt-b* haplotypes only occur in the Guayas-Taura drainage. Despite most of them occurring at low frequency (mean 0.05 ±0.01(SE)), their presence could be seen as evidence of a population that is old enough to have generated numerous alleles. Allelic richness (as a measure of genetic diversity) unlike heterozygosity, takes into account rare alleles, which is important for appropriate inferences of population genetics (Greenbaum et al., 2014). Likewise, microsatellite data showed that sites from Guayas had the greatest allele richness, particularly G1 and G4 (Appendix S14). It is worth noting that the genetic diversity values of these sites were closely followed by the sites from other drainages like E4, E2B (from Esmeraldas) and SR (from Santa Rosa). Other genetic diversity indices such as haplotype diversity for *Cyt-b* and the mean effective number of alleles across loci for the microsatellite data also showed some sites from Guayas with the greatest genetic diversity (Table 2). Similar to allelic richness, these values were closely followed by other sites outside of the Guayas drainage (e.g., E2B, E9, E4, SP, SR) (Table 2). Also, these sites mostly correspond to a range of low to mid-level elevations between 60 to 300 m. Sites at these lower elevations are expected to harbor greater genetic diversity compared to sites at higher elevations following the species-area relationship (Preston, 1962) and the downstream bias in the movement of alleles.

The *COI* data, on the other hand, yielded different results with the greatest level of genetic diversity found in sites from the Jubones drainage. However, it is possible that the result has been obscured due to the small sample sizes analyzed (N≈8) for the gene and its lower rate of evolution. For the reasons indicated above, it seems like the *Cyt-b* and microsatellite data provide more reliable estimates of genetic diversity.

The highest levels of genetic diversity found in some sites from Guayas were not that much higher than levels seen in other populations. They could also be a byproduct of the Guayas drainage being the largest drainage in western Ecuador, draining an area of 32,218 km^2 (National Council of Water Resources, 2002), and thus facilitating the accumulation of genetic diversity over geological time. The introgression of alleles from *R. minor* in the northern Guayas may have also resulted in artificially elevated estimates of genetic diversity in the Guayas drainage. Nonetheless, using the genetic diversity criterion, the Guayas drainage seems like the most likely place of origin of *Rhoadsia* based on the *Cyt-b* and microsatellite data, although this conclusion must be taken as highly tentative at this stage since no formal statistical test was implemented to support this hypothesis.

Evolutionary relationship of *Rhoadsia* to the neighboring genera *Carlana* and *Parastremma* in the subfamily Rhoadsiinae.

The phylogenetic analyses appear to indicate that populations from the southern region (*R. altipinna*) are genetically closer with the outgroups than individuals from the northern (*R. minor*) region (Appendix S16-S17). The genus *Rhoadsia* is a member of the subfamily Rhoadsiinae along with three other nominal genera: *Nematocharax* (East Brazil – samples not available), *Parastremma* (Colombia) *and Carlana* (Central America) (Mirande, 2010). Other studies have

corroborated the inclusion of these genera in the subfamily Rhoadsiinae (Cardoso, 2003; Géry, 1977), although phylogenetic relationships within this group have rarely been studied. *Rhoadsia* exhibited fewer nucleotide differences with the Colombian *P. sadina* (16% for *Cyt-b,* 14% for *COI*) than with the Central American *C. eigenmanni* (21% for *Cyt-b,* 16% for *COI*). This is consistent with the geographical distance between *Rhoadsia* and these two genera. Although the direction of evolution is not clear, *Rhoadsia* appears more closely related phylogenetically to the genus *Parastremma* based on my data.

Furthermore, the topology depicted in the ML gene tree (Appendix S16-S17) is consistent with the genetic diversity data presented above in suggesting that *Rhoadsia* expanded from a population that originated in the Guayas drainage. This was more notorious with the *Cyt-b* data which showed at the base of the ML tree some individuals from the Southern Guayas (G3, G4, and G5), followed by a supported clade that includes the rest of the sites from the Southern drainages (Guayas, Taura, San Pablo, Jubones, and Santa Rosa) and a clade that clustered all the samples from the northern drainages (Sua, Atacames, and Esmeraldas) including most of the individuals from the northern Guayas (G1, G2, G6, and G7) (Appendix S16). This also may suggest that populations from the southern drainages like San Pablo, Jubones and Santa Rosa may have derived from the same ancestral lineage in Guayas some time ago. The northern populations of *R. minor,* on the other hand, could have derived even earlier denotating a more conspicuous genetic separation, producing a strongly supported clade (bootstrap ≥ 70) that branches out from the Southern region (Appendix S16-S17). The *COI* data presented a similar topology. However, the base of the tree included not only sites from Guayas but other regions from the southern

drainages. The clade from the northern region with the populations from the northern Guayas was also present (Appendix S17).

Zoogeographic Regions and Potential Evolutionary Breaches of Freshwater Fishes in Western Ecuador

Fluctuations in the topography have contributed to the isolation and diversification of many species in western Ecuador (Rodríguez Tribaldos et al., 2017; Albert and Reis, 2011). For example, the lift of the Andean mountain approximately 20 million years ago (Albert and Reis, 2011) isolated the freshwater fishes on opposite sides such that almost no freshwater fish species are common to both sides of the Andes, and those that are may be cryptic species pairs (e.g., *Hoplias malabaricus,* which is listed from the Amazon region in eastern Ecuador and the Esmeraldas and Santiago-Cayapas basins in northwestern Ecuador, and most likely constitutes distinct species). Regional climate and the distribution of the hydrographic systems also may play an important role in the isolation of populations of organisms, especially of freshwater fishes due to their limited ability to disperse. Patterns of diversification caused by major vicariant events can be detected with the help of comparative phylogeographic studies (Avise et al., 2016; Gutiérrez-García and Vázquez-Domínguez, 2011). These types of studies are rare in Ecuador. Some of the organisms studied include amphibians and reptiles from the northwestern Ecuador (Arteaga et al., 2016), and hummingbirds from the Andean mountains (Chaves and Smith, 2011; Chaves et al., 2007). Phylogeographic studies on aquatic species are among the most underrepresented. Some examples include studies on genetic divergence of the genus *Hoplias* between rivers and impoundments in Western Ecuador (Cucalón and Bajaña, 2015; Aguirre et al.,

2013), and morphological and genetic variation of *Rhoadsia* along elevational gradients (Malato et al., 2017; Aguirre et al., 2016). However, besides an unpublished thesis by Cucalón and Bajaña (2015), no study has focused on describing the evolutionary relationship of populations of fishes throughout western Ecuador.

The Ministry of Environment of Ecuador (MAE, acronym in Spanish) (2012) released a report describing a methodology for the cartographic representation of the ecosystems in continental Ecuador. In this report, they presented a map of the biography of Ecuador where they divided Western Ecuador into two biogeographic provinces based on distribution patterns of plant species and bioclimatic factors. The two biogeographic provinces are Choco (Esmeraldas), and Tumbes-Guayaquil (Guayas and southward). On the other hand, Barriga (2012), in his list of fishes of Ecuador, divided western Ecuador into five Ichthyo-hydrographic zones. These Ichthyo-hydrographic zones were mainly assigned following the structure of the Ecuadorian hydrographic systems. They are Santiago-Cayapas, Esmeraldas, Guayas, Catamayo, and the Intertidal zone. It presents a greater division (five zones) than MAE (2012) presumably due to the inability of freshwater fishes to move between drainages. The Esmeraldas and the Guayas Ichthyo-hydrographic zone are the largest. The Esmeraldas Ichthyo-hydrographic zone is represented by the Esmeraldas drainage, while the Guayas Ichthyo-hydrographic zone is comprised of multiple basins with similar species compositions. Consequently, the Guayas Ichthyo-hydrographic zone extends from the Guayas drainage to the Jubones drainage. According to Barriga (2012), *R. minor* is endemic to the Esmeraldas Ichthyo-hydrographic zone while *R. altipinna* is native to the Guayas Ichthyo-hydrographic zone (it is not endemic because it also appears in northern Peru). The biogeographic provinces recognized by MAE (2012) and the

Ichthyo-hydrographic zones of Barriga (2012) share the same breach between the Esmeraldas and Guayas drainages (Jimenez-Prado et al., 2015), suggesting the presence of an important barrier there that could have served to facilitate allopatric speciation between these basins.

Divergence patterns of *Rhoadsia* varied depending on the genetic marker used (i.e., mtDNA or nuclear genome). The microsatellite data seem consistent with the breach between the Guayas and Esmeraldas basins identified by Barriga (2012) and MAE (2012). *Rhoadsia altipinna* from the northern Guayas was more closely related to populations from the southern region (Group II and III, Fig. 10) than to populations from Esmeraldas (Group I, Fig. 10). There was evidence of some admixture of *R. minor* (Southern Esmeraldas) *and R. altipinna* (Northern Guayas) according to the STRUCTURE analysis (Fig. 10) and the neighbor-joining analysis conducted by sampling site (Appendix S23). On the other hand, the mtDNA data mainly highlighted the introgression of *R. minor* into the northern Guayas, showing that the strongest breach between the two species might not be between the Guayas and the Esmeraldas drainages as previously thought (Jimenez-Prado et al., 2015; Barriga, 2012), but somewhere within the Guayas drainage. This was evidenced by the distribution of haplotypes from both mtDNA markers (Fig. 4-9) and the FST values (Table 4-5). It is worth noting that the Guayas and the Esmeraldas drainages almost meet near the city of Santo Domingo. The river Toachi in this locality has been reported to be the place where both Esmeraldas and Guayas drainages start northward and southward, respectively (MAE, 2012). Moreover, local news in Ecuador have previously reported overflow of the Toachi river that has affected nearby villages (Guerrero, 2016). It is possible that some fishes occurring in this area including *R. minor* may have crossed over through this passage into the Guayas drainage. However, this introgression was more strongly evidenced in the

northern Guayas than in the Esmeraldas basin, suggesting that the flow of alleles occurred mostly from Esmeraldas into Guayas. Introgression was also not apparent in the southern Guayas, suggesting the presence of a barrier somewhere lower in the Guayas drainage. There are no reports of divergence of any other fish species at this particular location (i.e., middle Guayas drainage), although this has rarely been studied. A similar phylogeographic study on *Hoplias microlepis* (a low-elevation fish species) showed that most of the populations sampled within the Guayas drainage shared very similar genetic makeup with no major divergence pattern (Cucalón and Bajaña, 2015; Aguirre et al., 2013). This may suggest that the genetic differences observed between the populations of *Rhoadsia spp.* from the Northern and Southern Guayas could be mainly influenced by intrinsic attributes of the species since other species in the area like *H. microlepis* do not exhibit this pattern.

The distribution of *Rhoadsia* in Western Ecuador is also consistent with patterns of climatic variation. Phenology patterns (relation of seasonal climate with plant and animal life) can be a crucial driver of species distributions (Chuine, 2010). The seasonal oscillation of the climate seems to have accentuated the distribution range of different types of forest in Western Ecuador, and with them, potentially the distribution of many species (MAE, 2012). For instance, some species of reptiles and amphibians seem to be associated with vegetation zones in northwestern Ecuador (Arteaga et al., 2016). The Northern and the Southern regions of Ecuador differ significantly on the levels of annual precipitation which consequently affect the type of vegetation (evergreen forest and semi-evergreen forest, respectively). It is possible that some aspects of the climatic oscillation and phenology of the region (MAE, 2012) affect the distribution of *Rhoadsia*. This pattern of variation in climatic conditions and forests virtually coincides with

the distribution of *Rhoadsia* in Ecuador. More specifically, it is consistent with the distribution of *R. minor* and might potentially explain the appearance of this species in the Northern Guayas, a region that is more similar to the ecosystems found in Esmeraldas (Evergreen forest) than the Southern Guayas (deciduous forest) (MAE, 2012). This hypothesis, if true, could have potential repercussions regarding the distribution of *Rhoadsia spp.* and other aquatic species considering how the climate is changing (Ahas and Aasa, 2006).

Another potential breach seems to be located in the southernmost drainages, where Populations of *Rhoadsia* exhibit genetic distinctiveness compared to Guayas. This isolation seems to be caused by the relative isolation of the multiple small basins that exist south of the Guayas drainage, preventing the passage of populations of fishes located farther south (Illustration I). Interestingly, the breach seems to go through the middle of the Jubones drainage, separating the high-elevation populations from the low-elevation populations of Jubones. The high-elevation populations were genetically similar to San Pablo (a low elevation site located six basins to the north), and according to microsatellite data they are included within the group II along with Guayas. The low-elevation portion of Jubones (J1, J2, and J3) although genetically distinguishable from populations of the nearby Santa Rosa drainage, cluster with them in group III (Fig. 10). The formation of these clusters seems to suggest the presence of a breach located somewhere between the groups II and III (Fig. 10). Barriga (2012) included the Santa Rosa drainage as part of the Ichthyo-hydrographic zone Catamayo (the southernmost zone) but not the Jubones drainage. Due to the similarities presented between Santa Rosa and the low-elevation sites of the Jubones drainage, it could be argued that the low-elevation portion of Jubones should be considered as part of the Catamayo Ichthyo-hydrographic zone. Regardless of the precise

location of the breach, my data are consistent with the occurrence of the Catamayo Ichthyo-hydrographic zone and a breach separating it from the Guayas and neighboring small southern drainages.

Conclusions

Based on the phylogeographic analysis, I conclude that *R. minor* and *R. altipinna* should be retained as closely related but valid species. *Rhoadsia spp*. are known to differ in body shape (Malato et al., 2017; Aguirre et al., 2016) and vertebral number (Aguirre et al., 2019) with elevation in both species, although *R. minor* is more streamlined than *R. altipinna* at high elevations (Malato et al., 2017; Jimenez-Prado et al., 2015). However, it is still not clear whether these morphological changes are driven by an adaptive genetic component or phenotypic plasticity. This is something that should be addressed in future studies. If adaptive evolution is at play, I suggest the use of next-generation approaches like Radseq (Davey and Blaxter, 2010) to investigate signals of adaptation at different elevations.

Gene flow between high and low elevation population occurred within drainages as evidenced by clustering of populations by drainage and the sharing of the same haplotypes. Nevertheless, there were significant differences in the frequencies of alleles and haplotypes between populations occurring at low and high elevations within the same drainage, indicating that natural barriers associated with the topography of the hydrographic system are limiting gene flow between high and low elevation populations of *Rhoadsia* (e.g., Jubones: cluster 8 and 9, Fig. 10), which may be facilitating their adaptive divergence.

Although I conclude that the two described species of *Rhoadsia* are valid, there are some caveats to this conclusion. A rather complex pattern suggesting the introgression of *R. minor* alleles into Northern Guayas *R. altipinna* populations was observed in both the mtDNA (Fig. 6-7) and microsatellite data (Fig. 10). This suggests that the two species may interbreed, which would conflict with the intrinsic aspect of the biological species concept (Mayr, 1942). According to some authors, this would no longer impede the delimitation of species if it is treated as only one of many features to consider (De Queiroz, 2007; De Queiroz, 2005). Nevertheless, the biology and ecology of the species are important aspects that should be considered when delimiting species (Dayrat, 2005). Unfortunately, virtually no information exists on the ecology of these species, something that should be corrected in the near future.

Each species of *Rhoadsia* was comprised of multiple evolutionary significant units, represented here as the 10 clusters estimated in the STRUCTURE analysis with the microsatellite data. Some of these appear to be of particular concern from a conservation perspective given the level of environmental degradation in the region (Ribeiro et al., 2017; Damanik-Ambarita et al., 2016). For instance, populations of *R. minor* from Atacames and Sua presented characteristics worth considered as a potential management unit. They were one of the most genetically distinctive populations and exhibited the lowest genetic diversity according to all genetic markers (Table 2, Fig. 3). The International Union for Conservation of Nature (IUCN) in the red list of threatened species placed *R. minor* in the “least concern” category (*R. altipinna* not listed), considering it as a stable population due to their commonness in the zone (Esmeraldas) (IUCN, 2018). This assessment appears to agree with the genetic diversity analysis performed in this

study, except for Atacames and Sua drainages (West to Esmeraldas), which for the reasons indicated above may warrant special protection.

The most likely geographical origin of the genus appears to be somewhere in the Guayas drainage according to the patterns of genetic diversity and phylogenetic comparison with the outgroup (Table 2, Fig. 6-7). However, this conclusion is still tentative, since this hypothesis was not formally tested. The greatest genetic diversity was found in sites from Guayas. However, it is possible that this result is due to the large size of this drainage basin (Barriga, 2012). On the other hand, the ML phylogenetic relationships showed the southern region (*R. altipinna*) as the closest related to the outgroups. The *Cyt-b* data was more specific showing some individuals from the Southern Guayas at the base of the ingroup, which seems to support the hypothesis of Guayas as potential geographical origin of the genus.

Three main zoogeographical zones were identified with two important zoogeographical breaches (Fig. 10). The first zoogeographic zone included Su, At, and all sites from Esmeraldas (Group I), the second zone included populations from Guayas, San Pablo and the high-elevation portion of the Jubones drainage (Group II), and the third zone included the low-elevation portion of the Jubones drainage and Santa Rosa (Group III). The first breach was between the Esmeraldas and the Guayas drainage. However, a pattern of genetic introgression into the population of the Northern Guayas suggested the existence of a passage located somewhere at the border of these two drainages. The introgression seemed to occur primarily from Esmeraldas to the Northern Guayas. Interestingly, the *R. minor* haplotypes did not pass into Southern Guayas, suggesting the existence of a potential natural barrier somewhere in the middle of the Guayas drainage. It was

argued that this barrier may be taxon specific, since the only other species studied (*H. microlepis*) did not show this pattern of divergence. The second main breach was found on the southernmost drainages with the boundary located between populations of the high (J4 and J5) and low (J1, J2, and J3) elevation of the Jubones drainage. This second breach was generally consistent with the occurrence of the Catamayo ictio-hydrographic zone identified by Barriga (2012).

This study contributes to the knowledge of the evolutionary relationships of populations of freshwater fishes in Western Ecuador, a region known for having high levels of endemic species while facing grave environmental degradation caused by human activity (Ribeiro et al., 2017; Damanik-Ambarita et al., 2016). To date, there is no other molecular study of this magnitude conducted on the freshwater fishes in this region. This study is intended to be used as a baseline for future research related (but no limited) to the distribution, conservation, and evolution of freshwater fishes, highlighting the unique biodiversity of Western Ecuador.

Literature cited

Aguirre, W. E., Shervette, V. R., Navarrete, R., Calle, P. and Agorastos, S. (**2013**). Morphological and Genetic Divergence of *Hoplias microlepis* (Characiformes: Erythrinidae) in Rivers and Artificial Impoundments of Western Ecuador. *Copeia* **2013**, 312–323.

Aguirre, W. E., Navarrete, R., Malato, G., Calle, P., Loh, M. K., Vital, W. F., Valadez, G., Vu, V., Shervette, V. R. and Granda, J. C. (**2016**). Body Shape Variation and Population Genetic Structure of *Rhoadsia altipinna* (Characidae: Rhoadsiinae) in Southwestern Ecuador. *Copeia* **104**, 554–569.

Aguirre, W. E., Young, A., Navarrete-Amaya, R., Valdiviezo-Rivera, J., Jiménez-Prado, P., Cucalón, R. V., Nugra-Salazar, F., Calle-Delgado, P., Borders, T. and Shervette, V. R. (**2019**). Vertebral number covaries with body form and elevation along the western slopes of the Ecuadorian Andes in the Neotropical fish genus *Rhoadsia* (Teleostei : Characidae). *Biol. J. Linn. Soc.* 1–15. *in press*.

Ahas, R. and Aasa, A. (**2006**). The effects of climate change on the phenology of selected Estonian plant, bird and fish populations. *Int. J. Biometeorol.* **51**, 17–26.

Albert, J. and Reis, R. (**2011**). *Historical Biogeography of Neotropical Freshwater Fishes*. Los Angeles: University of California Press.

Aldhebiani, A. Y. (**2018**). Species concept and speciation. *Saudi J. Biol. Sci.* **25**, 437–440.

Arbogast, B. S. and Kenagy, G. J. (**2001**). Comparative phylogeography as an integrative approach to historical biogeography. *J. Biogeogr.* **28**, 819–825.

Arroyo-Santos, A., Olson, M. E. and Vergara-Silva, F. (**2014**). The phylogeography debate and the epistemology of model-based evolutionary biology. *Biol. Philos.* **29**, 833–850.

Arteaga, A., Pyron, R. A., Peñafiel, N., Romero-Barreto, P., Culebras, J., Bustamante, L., Yánez-Muñoz, M. H. and Guayasamin, J. M. (**2016**). Comparative phylogeography reveals cryptic diversity and repeated patterns of cladogenesis for amphibians and reptiles in northwestern Ecuador. *PLoS One* **11**,.

Avise, J. C. (**2000**). *Phylogeography: The History and Formation of Species*. Harvard University Press, Cambridge, MA.

Avise, J. C. and Riddle, B. (**2009**). Phylogeography: Retrospect and Prospect. *J. Biogeogr.* **36**, 3–15.

Avise, J. C., Arnold, J., Ball, R. M., Bermingham, E., Lamb, T., Neigel, J. E., Reeb, C. A. and Saunders, N. C. (**1987**). Intraspecific Phylogeography: The Mitochondrial DNA Bridge Between Population Genetics and Systematics. *Annu. Rev. Ecol. Syst.* **18**, 489–522.

Avise, J. C., Bowen, B. W. and Ayala, F. J. (**2016**). In the light of evolution X: Comparative phylogeography. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 7957–7961.

Bagley, J. C., Matamoros, W. A., Mcmahan, C. D., Tobler, M., Chakrabarty, P. and Johnson, J. B. (**2016**). Phylogeography and species delimitation in convict cichlids (Cichlidae : Amatitlania): Implications for taxonomy and Plio-Pleistocene evolutionary history in central America. *Biol. J. Linn. Soc.* **120**, 155–170.

Baldwin, C. C., Mounts, J. H., Smith, D. G. and Weigt, L. A. (**2009**). Genetic identification and color descriptions of early life-history stages of Belizean *Phaeoptyx* and *Astrapogon* (Teleostei:

Apogonidae) with comments on identification of adult *Phaeoptyx*. *Zootaxa* **2008**, 1–22.

Balloux, F., Lugon-Moulin, N., Article, R., Balloux, F. and Lugon-Moulin, N. (**2002**). The estimation of population differentiation with microsatellite markers. *Mol. Ecol.* **11**, 155–165.

Barrett, R. D. H. and Hebert, P. D. N. (**2005**). Identifying spiders through DNA barcodes. *Can. J. Zool.* **83**, 481–491.

Barriga, R. S. (**2012**). Lista de peces de agua dulce e intermareales del ecuador. *Rev. Politec.* **30**, 83–119.

Bhat, H. and Kumar, N. (**2010**). *On the derivation of the Bayesian Information Criterion. Technical report. http://nscs00.ucmerced.edu/~nkumar4/BhatKumarBIC.pdf*.

Böhlke, J. (**1958**). Studies on Fishes of the Family Characidae . No . 14 . A Report on Several Extensive Recent Collections from Ecuador. *Proc. Acad. Nat. Sci. Philadelphia* **110**, 1–121.

Cardoso, A. R. (**2003**). Subfamily Rhoadsiinae. In *Check list of the freshwater fishes of South and Central America* (eds.Reis, R., Kullande, S. & Ferraris, C.), p. Edipucrs.

Chapin, F. S., Zavaleta, E. S., Eviner, V. T., Naylor, R. L., Vitousek, P. M., Reynolds, H. L., Hooper, D. U., Lavorel, S., Sala, O. E., Hobbie, S. E., et al. (**2000**). Consequences of changing biodiversity. *Nature* **405**, 234–42.

Chaves, J. A. and Smith, T. B. (**2011**). Evolutionary patterns of diversification in the Andean hummingbird genus *Adelomyia*. *Mol. Phylogenet. Evol.* **60**, 207–218.

Chaves, J. A., Pollinger, J. P., Smith, T. B. and LeBuhn, G. (**2007**). The role of geography and ecology in shaping the phylogeography of the speckled hummingbird (*Adelomyia melanogenys*) in Ecuador. *Mol. Phylogenet. Evol.* **43**, 795–807.

Chivian, E. and Bernstein, A. (**2008**). *Sustaining life: How human health depends on biodiversity*. New York: Center for Health and the Global Environment. Oxford University Press.

Chuine, I. (**2010**). Why does phenology drive species distribution? *Philos. Trans. R. Soc. B Biol. Sci.* **365**, 3149–3160.

Costa-Silva, G. J., Rodriguez, M. S., Roxo, F. F., Foresti, F. and Oliveira, C. (**2015**). Using different methods to access the difficult task of delimiting species in a complex neotropical hyperdiverse group. *PLoS One* **10**, 1–12.

Cucalón, R. V. and Bajaña, L. (**2015**). Filogeografía molecular del guanchiche *Hoplias spp.* (Characiformes: Erythrinidae) de la costa ecuatoriana (Unpublished thesis). Escuela Superior Politécnica del Litoral, Guayaquil, Ecuador.

Damanik-Ambarita, M. N., Lock, K., Boets, P., Everaert, G., Nguyen, T. H. T., Forio, M. A. E., Musonge, P. L. S., Suhareva, N., Bennetsen, E., Landuyt, D., et al. (**2016**). Ecological water quality analysis of the Guayas river basin (Ecuador) based on macroinvertebrates indices. *Limnologica* **57**, 27–59.

Davey, J. L. and Blaxter, M. W. (**2010**). RADseq: Next-generation population genetics. *Brief. Funct. Genomics* **9**, 416–423.

Dayrat, B. (**2005**). Towards integrative taxonomy. *Biol. J. Linn. Soc.* **85**, 407–415.

De Queiroz, K. (**2005**). Ernst Mayr and the modern concept of species. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 6600–6607.

De Queiroz, K. (**2007**). Species Concepts and Species Delimitation. *Syst. Biol.* **56**, 879–886.

Dodson, C. H. and Gentry, A. H. (**1991**). Biological Extinction in Western Ecuador. *Ann. Missouri Bot. Gard.* **78**, 273–295.

Drummond, A. J., Suchard, M. A., Xie, D. and Rambaut, A. (**2012**). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* **29**, 1969–1973.

Duncan, S. I., Crespi, E. J., Mattheus, N. M. and Rissler, L. J. (**2015**). History matters more when explaining genetic diversity within the context of the core-periphery hypothesis. *Mol. Ecol.* **24**, 4323–4336.

Earl, D. A. and VonHoldt, B. M. (**2012**). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* **4**, 359–361.

Eigenmann, C. H. and Henn, A. W. (**1914**). On new species of fishes from Colombia, Ecuador, and Brazil. *Contrib. from Zool. Lab. Indiana Univ.* **140**, 231–234.

Evanno, G., Regnaut, S. and Goudet, J. (**2005**). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Mol. Ecol.* **14**, 2611–2620.

Excoffier, L., Laval, G. and Schneider, S. (**2005**). Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evol. Bioinform. Online* **1**, 47–50.

Farias, I. P., Ortí, G., Sampaio, I., Schneider, H. and Meyer, A. (**2001**). The cytochrome b gene as a phylogenetic marker: the limits of resolution for analyzing relationships among cichlid fishes. *J. Mol. Evol.* **53**, 89–103.

Fennessy, J., Bidon, T., Reuss, F., Vamberger, M., Fritz, U., Janke Correspondence, A., Kumar, V., Elkan, P., Nilsson, M. A. and Janke, A. (**2016**). Multi-locus Analyses Reveal Four Giraffe Species Instead of One. *Curr. Biol.* **26**, 1–7.

Fowler, H. W. (**1911**). New fresh-water from western Ecuador. *Proc. Acad. Nat. Sci. Philadelphia* **63**, 493–520.

Frankham, R., Ballou, J. D., Briscoe, D. A. and McInnes, K. H. (**2002**). *Introduction to Conservation Genetics*. Cambridge: Cambridge University Press.

Fu, Y.-X. (**1997**). Statistical Test of Neutrality of Mutations Agains Population Growth, Hitchhiking and Background Selection. *Genetics* **147**, 917–925.

Geneious Prime 2018 11.0.5 (https://www.geneious.com).

Géry, J. (**1977**). *Characoids of the world*. Neptune City, NJ: T.F.H: Publications.

Gippoliti, S. and Amori, G. (**2007**). The problem of subspecies and biased taxonomy in conservation lists: The case of mammals. *Folia Zool.* **56**, 113–117.

Gouy, M., Guindon, S. and Gascuel, O. (**2010**). Sea view version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol. Biol. Evol.* **27**, 221–224.

Greenbaum, G., Templeton, A. R., Zarmi, Y. and Bar-David, S. (**2014**). Allelic richness following population founding events - A stochastic modeling framework incorporating gene flow and genetic drift. *PLoS One* **9**, 1–23.

Guerrero, G. (**2016**). Desbordamiento de río en Santo Domingo alcanzó al menos 2 metros de altura. *El Universo*.

Guindon, S. and Gascuel, O. (**2003**). Society of Systematic Biologists A Simple, Fast, and Accurate Algorithm to Estimate Large Phylogenies by Maximum A Simple, Fast, and Accurate Algorithm to Estimate Large Phylogenies by Maximum Likelihood. *Source Syst. Biol. Syst. Biol* **52**, 696–704.

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O., Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., et al. (**2010**). Society of Systematic Biologists New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0 New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Source Syst. Biol. Syst. Biol* **59**, 307–321.

Gutiérrez-García, T. A. and Vázquez-Domínguez, E. (**2011**). Comparative Phylogeography: Designing Studies while Surviving the Process. *Bioscience* **61**, 857–868.

Hamilton, S. E. and Lovette, J. (**2015**). Ecuador's mangrove forest carbon stocks: A spatiotemporal analysis of living carbon holdings and their depletion since the advent of commercial aquaculture. *PLoS One* **10**, 1–14.

Harlt, D. L. and Clark, G. C. (**1997**). *Principles of Population Genetics*. Sunderland: Sinauer Associates.

Hasegawa, M., Kishino, H. and Yano, T. aki (**1985**). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* **22**, 160–174.

Hebert, P. D. N. and Barrett, R. D. H. (**2005**). Reply to the comment by L. Prendini on "Identifying spiders through DNA barcodes." *Can. J. Zool.* **83**, 505–506.

Hebert, P. D., Cywinska, A., Ball, S. L. and DeWaard, J. R. (**2003**). Biological identification through DNA barcodes. *Proc. R. Soc. / Biol. Sci.* **207**, 313–321.

Hickerson, M. J., Meyer, C. P. and Moritz, C. (**2006**). DNA barcoding will often fail to discover new animal species over broad parameter space. *Syst. Biol.* **55**, 729–739.

Huang, J. P. and Knowles, L. L. (**2016**). The species versus subspecies conundrum: Quantitative delimitation from integrating multiple data types within a single Bayesian approach in hercules beetles. *Syst. Biol.* **65**, 685–699.

Ingman, M., Kaessmann, H., Bo, S. and Gyllensten, U. (**2000**). Mitochondrial genome variation and the origin of modern humans. *Nature* **408**, 708–713.

IUCN (**2018**). The IUCN Red List of Threatened Species. Version 2018-2. http://www.iucnredlist.org. Downloaded on 10 February 2018.

Ivanov, V., Lee, K. M. and Mutanen, M. (**2018**). Mitonuclear discordance in wolf spiders: Genomic evidence for species integrity and introgression. *Mol. Ecol.* **27**, 1681–1695.

Jackson, D. A., Peres-Neto, P. R. and Olden, J. D. (**2001**). What controls who is where in freshwater fish communities – the roles of biotic, abiotic, and spatial factors. *Can. J. Fish. Aquat. Sci.* **58**, 157–170.

Jarne, P. and Lagoda, P. J. L. (**1996**). Microsatellites, from molecules to populations and back. *Trends Ecol. Evol.* **11**, 424–429.

Jimenez-Prado, P., Aguirre, W., Laaz-Moncayo, E., Navarrete-Amaya, R., Nugra-Salazar, F., Rebolledo-Monsalve, E., Zárate-Hugo, E., Torres-Noboa, A. and Valdiviezo-Rivera, J. (**2015**). *Guia de peces para aguas continentales en la vertiente occidental del Ecuador*. 1st ed. Esmeraldas, Ecuador.

Knowles, L. L. and Carstens, B. C. (**2007**). Delimiting Species without Monophyletic Gene Trees. *Soc. Syst. Biol.* **56**, 887–895.

Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A. and Mayrose, I. (**2015**). CLUMPAK: a program for identifying clustering modes and packaging population structure inferences across K. *Mol. Ecol. Resour.* **15**, 1179–1191.

Kress, J. W., García-Robledo, C., Uriarte, M. and Erickson, D. L. (**2015**). DNA barcodes for ecology, evolution, and conservation. *Trends Ecol. Evol.* **30**, 25–35.

Leigh, J. W. and Bryant, D. (**2015**). POPART: full-feature software for haplotype network construction. *Methods Ecol. Evol.* **6**, 1110–1116.

Loh, M., Vital, W. F., Vu, V., Navarrete, R., Calle, P., Shervette, V. R., Torres, A. and Aguirre, W. E. (**2013**). Isolation of sixteen microsatellite loci for *Rhoadsia altipinna* (Characiformes: Characidae) from an impacted river basin in western Ecuador. *Conserv. Genet. Resour.* **6**, 229–231.

Lynch, J. D. and Duellman, W. E. (**1997**). *Frogs of the Genus Eleutherodactylus in Western Ecuador*. (eds.Trueb, L. & Collins, J. T.) Lawrence: Natural History Museum.

Malato, G., Shervette, V. R., Navarrete Amaya, R., Valdiviezo Rivera, J., Nugra Salazar, F., Calle Delgado, P., Karpan, K. C. and Aguirre, W. E. (**2017**). Parallel body shape divergence in the Neotropical fish genus *Rhoadsia* (Teleostei: Characidae) along elevational gradients of the western slopes of the Ecuadorian Andes. *PLoS One* **12**, 1–27.

Manokaran, N. (**1992**). an Overview of Biodiversity in Malaysia. *J. Trop. For. Sci.* **5**, 271–290.

Mayr, E. (**1942**). *Systematics and the origin of species*. New York: Columbia University Press.

Minin, V., Bloomquist, E. and Suchard, M. (**2008**). Skyride coalescent. *Mol. Biol. Evol.* **25**, 1459–1471.

Ministerio de Ambiente del Ecuador (MAE) (**2012**). *Metodología para la Representación Cartográfica de los Ecosistemas del Ecuador Continental*. Quito.

Mirande, J. M. (**2010**). Phylogeny of the family characidae (teleostei: Characiformes): From characters to taxonomy. *Neotrop. Ichthyol.* **8**, 385–568.

Mora, C., Tittensor, D. P., Adl, S., Simpson, A. G. B. and Worm, B. (**2011**). How many species are there on earth and in the ocean? *PLoS Biol.* **9**, 1–8.

Moritz, C. (**1994**). Defining “Evolutionarily Significant Units” for conservation. *Trends Ecol. Evol.* **9**, 373–375.

Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A. B. and Kent, J. (**2000**). Biodiversity hotspots for conservation priorities. *Nature* **403**, 853–858.

National Council of Water Resources (**2002**). *División Hidrográfica del Ecuador*. Ecuador.

Nei, M. (**1987**). *Molecular Evolutionary Genetics*. New York: Columbia University Press.

Nei, M., Tajima, F. and Tateno, Y. (**1983**). Accuracy of estimated phylogenetic trees from molecular data

- I. Distantly Related Species. *J. Mol. Evol.* **19**, 153–170.

Newbold, T., Hudson, L. N., Hill, S. L., Contu, S., Lysenko, I., Senior, R. a, Börger, L., Bennett, D. J., Choimes, A., Collen, B., et al. (**2015**). Global effects of land use on local terrestrial biodiversity. *Nature* **520**, 45-.

Ortega, H., Hidalgo, M., Trevejo, G., Correa, E., Cortijo, A. M., Meza, V. and Espino, J. (**2012**). *Lista anotada de los peces neotropicales de aguas continentales del Perú*.

Palsbøll, P. J., Bérubé, M. and Allendorf, F. W. (**2007**). Identification of management units using population genetic data. *Trends Ecol. Evol.* **22**, 11–16.

Palumbi, S. R. (**1996**). *Moleculas Systematics*. Second Edi. (eds.Hillis, D. M., Moritz, C. & Mable, B. K.) Sunderland: Sinauer Associates.

Peakall, R. and Smouse, P. E. (**2006**). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol. Ecol. Notes* **6**, 288–295.

Peakall, R. and Smouse, P. E. (**2012**). GenALEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics* **28**, 2537–2539.

Posada, D. (**2008**). jModelTest: Phylogenetic model averaging. *Mol. Biol. Evol.* **25**, 1253–1256.

Prendini, L. (**2005**). Comment on "Identifying spiders through DNA barcodes." *Can. J. Zool.* **83**, 498–504.

Preston, F. (**1962**). The canonical distribution of commonness and rarity: Part I. *Ecology* **43**, 185–215.

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

Rambaut, A. (**2016**). FigTree v1.4.3. http://tree.bio.ed.ac.uk/software/figtree/.

Rambaut, A., Suchard, M., Xie, D. and AJ, D. (**2014**). Tracer v1.6. http://beast.bio.ed.ac. uk/Tracer.

Raymond, M. and Rousset, F. (**1995**). GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *J. Hered.* **86**, 248–249.

Reeves, P. A. and Richards, C. M. (**2011**). Species delimitation under the general lineage concept: An empirical example using wild North American hops (Cannabaceae: *Humulus lupulus*). *Syst. Biol.* **60**, 45–59.

Reilly, S. B., Corl, A. and Wake, D. B. (**2015**). An integrative approach to phylogeography: investigating the effects of ancient seaways, climate, and historical geology on multi-locus phylogeographic boundaries of the Arboreal Salamander (*Aneides lugubris*). *BMC Evol. Biol.* **15**, 241.

Ribeiro, L., Pindo, J. C. and Dominguez-Granda, L. (**2017**). Assessment of groundwater vulnerability in the Daule aquifer, Ecuador, using the susceptibility index method. *Sci. Total Environ.* **574**, 1674–1683.

Rodríguez Tribaldos, V., White, N. J., Roberts, G. G. and Hoggard, M. J. (**2017**). Spatial and temporal uplift history of South America from calibrated drainage analysis. *Geochemistry, Geophys. Geosystems* **18**, 2321–2353.

Rousset, F. (**2008**). Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. *Mol. Ecol. Resour.* **8**, 103–106.

Ryder, O. A. (**1986**). Species conservation and systematics: the dilemma of subspecies. *Trends Ecol. Evol.* **1**, 9–10.

Schield, D. R., Adams, R. H., Card, D. C., Corbin, A. B., Jezkova, T., Hales, N. R., Meik, J. M., Perry, B. W., Spencer, C. L., Smith, L. L., et al. (**2018**). Cryptic genetic diversity, population structure, and gene flow in the Mojave rattlesnake (*Crotalus scutulatus*). *Mol. Phylogenet. Evol.* **127**, 669–681.

Schuelke, M. (**2000**). An economic method for the fluorescent labeling of PCR fragments A poor man ' s approach to genotyping for research and high-throughput diagnostics . *Nat. Biotechnol.* **18**, 233–234.

Selkoe, K. A. and Toonen, R. J. (**2006**). Microsatellites for ecologists: A practical guide to using and evaluating microsatellite markers. *Ecol. Lett.* **9**, 615–629.

Shaw, K. L. (**2002**). Conflict between nuclear and mitochondrial DNA phylogenies of a recent species radiation: What mtDNA reveals and conceals about modes of speciation in Hawaiian crickets. *Proc. Natl. Acad. Sci.* **99**, 16122–16127.

Silva, J. C., Gubiani, É. A., Piana, P. A. and Delariva, R. L. (**2016**). Effects of a small natural barrier on the spatial distribution of the fish assemblage in the Verde River, Upper Paraná River Basin,Brazil. *Brazilian J. Biol.* **76**, 851–863.

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

Stauffer, J. R., Mckaye, K. R. and Konings, A. F. (**2002a**). Behaviour: an important diagnostic tool for Lake Malawi cichlids. *Fish Fish.* **3**, 213–214.

Stauffer, J. R., Kocovsky, P. M. and Ruffing, R. A. (**2002b**). Species concepts and speciation of fishes: Concluding remarks. *Fish Fish.* **3**, 230–232.

Sunnucks, P. (**2000**). Efficient genetic markers for population biology. *Trends Ecol. Evol.* **15**, 199–203.

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

Takezaki, N. and Nei, M. (**1996**). Genetic Distances and Reconstruction of Phylogenetic Trees From Microsatellite DNA. *Genetics* **144**, 389–399.

Takezaki, N., Nei, M. and Tamura, K. (**2010**). POPTREE2: Software for constructing population trees from allele frequency data and computing other population statistics with windows interface. *Mol. Biol. Evol.* **27**, 747–752.

The International Commission on Zoological Nomenclature (**2000**). Retrieved from: http://iczn.org/iczn/index.jsp. *Fourth Ed.*

Toews, D. P. L. and Brelsford, A. (**2012**). The biogeography of mitochondrial and nuclear discordance in animals. *Mol. Ecol.* **21**, 3907–3930.

Tóth, J. P., Varga, Z., Verovnik, R., Wahlberg, N., Váradi, A. and Bereczki, J. (**2017**). Mito-nuclear discordance helps to reveal the phylogeographic patterns of *Melitaea ornata* (Lepidoptera: Nymphalidae). *Biol. J. Linn. Soc.* **121**, 267–281.

Twilley, R. R., Cardenas, W., Rivera-monroy, V. H., Espinoza, J., Suescum, R., Armijos, M. M. and

Solorzano, L. (**2001**). The Gulf of Guayaquil and the Guayas river estuary, Ecuador. In *Coastal Marine Ecosystems of Latin America*, pp. 245–263. Springer, Berlin, Heidelberg.

Unmack, P. J., Bagley, J. C., Adams, M., Hammer, M. P. and Johnson, J. B. (**2012**). Molecular phylogeny and phylogeography of the australian freshwater fish genus *Galaxiella*, with an emphasis on dwarf galaxias (*G. pusilla*). *PLoS One* **7**, e38433.

Vannote, R. L., Minshall, G. W., Cummins, K. W., Seddell, J. R. and Cushing, C. E. (**1980**). The river continuum concept. *Can. J. Fish. Aquat. Sci.* **37**, 130–137.

Vernygora, O. V., Davis, C. S., Murray, A. M. and Sperling, F. A. H. (**2018**). Delimitation of *Alosa* species (Teleostei: Clupeiformes) from the Sea of Azov: integrating morphological and molecular approaches. *J. Fish Biol.* **93**, 1216–1228.

Waples, R. S. (**1995**). Evolutionarily Significant Units and the Conservation of Biological Diversity under the Endangered Species Act. *Am. Fish. Soc. Symp.* **17**, 8–27.

Willis, S. C. (**2017**). One species or four? Yes!...and, no. Or, arbitrary assignment of lineages to species obscures the diversification processes of Neotropical fishes. *PLoS One* **12**, e0172349.

Wilson, E. O. (**1988**). *Biodiversity*. 1st ed. (eds.Wilson, E. O. & Frances, M. P.) Washington: National Academy of Sciences.

Wilson, E. 0 and Brown, W. L. (**1953**). The Subspecies Concept and Its Taxonomic Application. *Syst. Zool.* **2**, 97–111.

Winemiller, K. O., Agostinho, A. A. and Caramaschi, E. P. (**2009**). Fish Ecology in Tropical Streams. In *Tropical Stream Ecology* (ed.Dundgeon, D.), pp. 107–146. Academic Press.

Zink, R. M. and Barrowclough, G. F. (**2008**). Mitochondrial DNA under siege in avian phylogeography. *Mol. Ecol.* **17**, 2107–2121.

Tables

Table 1. Sampling sites and sample sizes for each molecular marker. *N (total number of individuals), COI (Cytochrome Oxidase), Cyt-b (Cytochrome b), Mstl (Microsatellites) ROM (Royal Ontario Museum), STRI (Smithsonian Tropical Research Institute Panama), CR (Costa Rica), Nic (Nicaragua), Pan (Panama). ** Parastremma sadina (Par) and Carlana eigenmanni (Car), species used as an outgroup for the phylogeographic analysis. * Sequences retrieved from Gene Bank. Further details of samples from museums provided in Appendix S3.*

Drainage	Site	Code	Lat	Lon	Elevation (m)	year	N	*Cyt-b*	*COI*	Mstl	Source
Atacames	Atacames	At	0.75888889	-79.85055556	29	2017	24	24	7	24	This Study
Sua	Sua	Su	0.7517634	-79.8966713	~50	2017	3	3	2	-	This Study
Esmeraldas	Afuera	E1	0.584138889	-79.61080583	50	2014	-	-	6*	-	Malato et al., 2017
	Teaone	E2B	0.71166667	-79.69083333	64	2017	24	24	8	24	This Study
	Silanchi	E9	0.1589667	-79.24345	133	2012	8	8	8	-	ROM
	Bravo	E3	-0.03875	-79.34408333	174	2014	-	-	10*		Malato et al., 2017
	Hermoso	E4	-0.07888889	-79.285	282	2014	24	24	8*	24	Malato et al., 2017
	Meme Chico	E7	-0.223916833	-79.07216683	668	2014	-	-	8*	-	Malato et al., 2017
	Blanco	E5	-0.006388889	-78.90291667	810	2014	-	-	8*	-	Malato et al., 2017
	Transito	E8	-0.31138889	-78.90444444	1093	2014	21	21	8*	21	Malato et al., 2017
	Mindo	E6	-0.05805556	-78.77388889	1260	2017	2	2	2	-	This Study
	Mindo	E6	-0.05805556	-78.77388889	1260	2014	24	24	13*	24	Malato et al., 2017
Guayas	Jauneche	G2	-1.23833333	-79.6725	~50	2008	15	15	15*	15	Malato et al., 2017
	Chimbo	G4	-2.18333333	-79.35083333	62	2017	19	19	7	19	This Study
	Palenque	G1	-0.57388889	-79.36194444	~100	2008	24	24	16*	24	Malato et al., 2017
	Caluma	G5	-1.61305556	-79.31055556	178	2017	2	2	2	-	This Study
	Otongo	G7	-0.3519167	-79.2145333	382	2012	4	4	4	-	ROM
	Chague Grande	G3	-2.14444444	-79.1225	401	2017	13	13	8	13	This Study
	Chiguilpe	G6	-0.3225	-79.2166	422	2012	2	2	2	-	ROM
Taura	Cutahay	Ta	-2.30345	-79.1733167	291	2012	2	2	2	-	ROM
San Pablo	Estero Mina	SP	-2.7175	-79.63472222	50	2017	15	15	6	15	This Study
Jubones	Huizo	J1	-3.32777778	-79.74166667	69	2014	24	24	8*	24	Malato et al., 2017
	Cascay	J2	-3.3300556	-79.7113611	136	2014	24	24	7	24	Malato et al., 2017
	Mollopongo	J3	-3.3153889	-79.65825	251	2014	24	24	8*	24	Malato et al., 2017
	Minas	J4	-3.3465278	-79.3815278	909	2014	20	20	8*	20	Malato et al., 2017
	Mondur	J5	-3.31416667	-79.28	1095	2014	24	24	10*	24	Malato et al., 2017
Santa Rosa	Santa Rosa	SR	-3.55861111	-79.94611111	86	2008	24	24	16*	24	Malato et al., 2017
Calima	Colombia	Par **	-	-	-	-	4	4	4	-	ROM
-	CR/Nic/Pan	Car**	-	-	-	-	7	7	7	-	STRI

Table 2. Genetic diversity indices of populations of *Rhoadsia* in western Ecuador based on mitochondrial genes cytochrome b (*Cyt-b*), cytochrome oxidase I (*COI*), and nuclear microsatellites (MstI). *N (number of individuals); mN (mean number of individuals analyzed); H (number of haplotypes); Na (mean number of alleles); Ne (effective number of alleles); mNe (mean of effective number of alleles across loci); HD (haplotype diversity); He (expected heterozygosity); %PA (percentage of private alleles). The grey colored bar within the cells represents the value in each cell relative to the maximum for the column. * Appendix S14 lists the number of alleles per locus for the microsatellites.*

Markers		Cyt-b					COI					MstI				
Drainage	**Site**	**N**	**H**	**Ne**	**HD**	**%PA**	**N**	**H**	**Ne**	**HD**	**%PA**	**mN**	**Na***	**mNe**	**Ho**	**He**
Sua	Su	3	1	1.00	0.00	0.00	2	1	1.00	0.00	0.00	-	-	-	-	-
Atacames	At	23	2	1.09	0.09	50.00	7	2	1.32	0.29	50.00	23.33	3.00	2.02	0.47	0.42
Esmeraldas	E2B	18	2	1.12	0.11	50.00	7	1	1.00	0.00	0.00	22.08	7.58	4.77	0.69	0.70
	E1	-	-	-	-	-	6	1	1.00	0.00	0.00	-	-	-	-	-
	E9	8	3	2.46	0.68	0.00	8	3	2.67	0.71	33.33	-	-	-	-	-
	E3	-	-	-	-	-	10	3	1.85	0.51	33.33	-	-	-	-	-
	E4	22	3	2.26	0.58	0.00	8	1	1.00	0.00	0.00	22.33	8.17	4.65	0.63	0.63
	E5	-	-	-	-	-	8	1	1.00	0.00	0.00	-	-	-	-	-
	E6	25	3	1.84	0.48	33.33	15	2	1.14	0.13	0.00	22.92	4.42	2.40	0.60	0.54
	E7	-	-	-	-	-	8	2	1.28	0.25	50.00	-	-	-	-	-
	E8	21	2	1.96	0.51	0.00	8	3	2.13	0.61	0.00	19.58	5.25	3.43	0.72	0.66
Guayas	G6	2	2	2.00	1.00	0.00	2	2	2.00	1.00	0.00	-	-	-	-	-
	G7	4	2	1.60	0.50	0.00	4	2	1.60	0.50	0.00	-	-	-	-	-
	G1	24	5	1.71	0.43	40.00	16	4	1.94	0.52	25.00	23.58	9.83	6.19	0.80	0.77
	G2	15	3	2.10	0.56	33.33	15	2	1.30	0.25	0.00	13.83	5.92	3.43	0.68	0.64
	G5	2	2	2.00	1.00	100.00	2	1	1.00	0.00	0.00	-	-	-	-	-
	G3	11	4	2.69	0.69	0.00	8	3	2.13	0.61	33.33	12.83	6.25	4.32	0.73	0.71
	G4	18	5	3.12	0.72	40.00	7	2	1.96	0.57	0.00	17.83	8.58	5.32	0.73	0.72
Taura	Ta	2	2	2.00	1.00	0.00	2	2	2.00	1.00	50.00	-	-	-	-	-
San Pablo	SP	11	3	2.47	0.65	66.67	6	1	1.00	0.00	0.00	14.92	5.08	3.11	0.70	0.61
Jubones	J1	24	3	1.95	0.51	0.00	9	4	3.24	0.78	0.00	21.00	5.67	3.01	0.60	0.62
	J2	24	3	2.32	0.59	0.00	8	3	2.46	0.68	0.00	20.75	5.75	3.05	0.55	0.60
	J3	24	1	1.00	0.00	0.00	6	4	3.60	0.87	0.00	20.83	5.33	2.76	0.55	0.58
	J4	20	1	1.00	0.00	0.00	8	4	3.20	0.79	25.00	18.42	3.83	2.51	0.59	0.56
	J5	24	1	1.00	0.00	0.00	10	2	2.00	0.56	0.00	23.25	3.67	2.44	0.59	0.53
Santa Rosa	SR	21	3	1.89	0.50	0.00	16	3	1.86	0.49	0.00	22.42	7.67	4.31	0.71	0.70
	TOTAL	346					206									

Table 3. AMOVA of populations of *Rhoadsia* in western Ecuador grouped by drainage, clades inferred from mtDNA genes, and clusters estimated in STRUCTURE analysis (K=10) for the mitochondrial gene cytochrome b (*Cyt-b)*, and Cytochrome Oxidase I (*COI)*, and nuclear microsatellites (*MstI*). See Appendix S7 for the specific assignment of the sites into the groups of each version of AMOVA per molecular marker. df (degrees of freedom). P-value was very significant (<0.0001) for all analysis otherwise specified. * p<0.01.

Cyt-b	**By drainage**				**By clade**				**By cluster (K=10)**			
Source of variation	**df**	**Sum of squares**	**% of variation**	**Fixation indices**	**df**	**Sum of squares**	**% of variation**	**Fixation indices**	**df**	**Sum of squares**	**% of variation**	**Fixation indices**
Among groups	7	431.23	55.09	FCT: 0.55091	1	355.96	62.46	FCT: 0.62460	9	564.42	75.37	FCT: 0.75369
Among populations within groups	14	157.16	26.73	FSC: 0.59517	20	232.42	23.14	FSC: 0.61632	12	23.963	5.32	FSC: 0.21595
Within populations	324	147.46	18.18	FST: 0.8182	324	147.46	14.4	FST: 0.85597	324	147.46	19.31	FST: 0.80688
Total	345	735.84			345	735.84			345	735.84		
COI												
Source of variation	**df**	**Sum of squares**	**% of variation**	**Fixation indices**	**df**	**Sum of squares**	**% of variation**	**Fixation indices**	**df**	**Sum of squares**	**% of variation**	**Fixation indices**
Among groups	7	95.103	35.45	FCT: 0.35448	1	77.729	46.81	FCT: 0.46809	9	139.73	54.18	FCT: 0.54182
Among populations within groups	18	63.552	28.81	FSC: 0.4463	24	80.926	23.41	FSC: 0.44006	16	18.924	8.46	FSC: 0.18465*
Within populations	180	85.437	35.74	FST: 0.35448	180	85.437	29.78	FST: 0.70216	180	85.437	37.36	FST: 0.62642
Total	205	244.09			205	244.09			205	244.09		
MstI												
Source of variation	**df**	**Sum of squares**	**% of variation**	**Fixation indices**	**df**	**Sum of squares**	**% of variation**	**Fixation indices**	**df**	**Sum of squares**	**% of variation**	**Fixation indices**
Among groups	5	301	7.63	FCT: 0.07635*	1	106.84	4.98	FCT: 0.04977*	9	508.9	16.51	FCT: 0.16512
Among populations within groups	10	266.36	12.72	FSC: 0.13769	14	460.51	16.09	FSC: 0.16935	6	58.451	3.56	FSC: 0.04266
Within populations	670	2272.2	79.65	FST: 0.20353	670	2272.2	78.93	FST: 0.21069	670	2272.2	79.93	FST: 0.20074
Total	685	2839.6			685	2839.6			685	2839.6		

Table 4. Pairwise F_{ST} of populations of *Rhoadsia* in western Ecuador based on mitochondrial gene cytochrome oxidase I. *F_{ST} values are below the diagonal. Significance values are above the diagonal (significance level=0.05): + (significant), - (not significant). F_{ST} values were colored to different shades of grey to represent the level of genetic differentiation (see legend at the bottom of the table).*

Drainages		Sua	Atacames	Esmeraldas									Guayas							Taura	San Pablo	Jubones					Santa Rosa
	Sites	Su	At	E2B	E1	E9	E3	E4	E5	E6	E7	E8	G6	G7	G1	G2	G5	G3	G4	Ta	SP	J1	J2	J3	J4	J5	SR
Sua	Su	*	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	+	-	+	+	+
Atacames	At	0.00	*	-	-	+	-	-	-	-	-	+	-	+	+	+	+	+	+	-	+	+	+	+	+	+	+
Esmeraldas	E2B	0.00	0.00	*	-	+	-	-	-	-	-	+	-	+	+	+	+	+	+	-	+	+	+	+	+	+	+
	E1	0.00	0.00	0.00	*	-	-	-	-	-	-	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	E9	0.00	0.23	0.26	0.23	*	-	+	-	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	E3	0.00	0.10	0.11	0.09	0.06	*	-	-	+	-	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	E4	0.00	0.02	0.00	0.00	0.29	0.13	*	-	-	-	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	E5	0.00	0.02	0.00	0.00	0.29	0.13	0.00	*	-	-	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	E6	0.00	0.03	0.00	0.00	0.37	0.18	0.00	0.00	*	-	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	E7	0.00	0.00	0.00	0.00	0.25	0.11	0.00	0.00	0.02	*	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	E8	0.17	0.39	0.42	0.39	0.38	0.36	0.44	0.44	0.52	0.41	*	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+
Guayas	G6	0.00	0.40	0.59	0.54	0.21	0.25	0.63	0.63	0.63	0.44	0.00	*	-	-	-	-	-	-	-	+	+	+	+	+	+	+
	G7	0.53	0.67	0.76	0.74	0.49	0.56	0.78	0.78	0.79	0.69	0.00	0.00	*	-	-	-	+	+	-	+	+	+	+	+	+	+
	G1	0.37	0.49	0.51	0.49	0.50	0.47	0.53	0.53	0.58	0.51	0.00	0.00	0.00	*	-	+	+	+	+	+	+	+	+	+	+	+
	G2	0.52	0.61	0.63	0.61	0.58	0.58	0.64	0.64	0.69	0.62	0.00	0.00	0.00	0.00	*	+	+	+	+	+	+	+	+	+	+	+
	G5	1.00	0.92	1.00	1.00	0.70	0.83	1.00	1.00	0.96	0.93	0.64	0.75	0.84	0.73	0.75	*	-	-	-	-	-	+	-	-	-	+
	G3	0.36	0.54	0.56	0.54	0.51	0.56	0.58	0.58	0.65	0.56	0.49	0.40	0.56	0.59	0.62	0.00	*	-	-	-	+	+	-	-	+	+
	G4	0.33	0.54	0.56	0.53	0.51	0.56	0.58	0.58	0.66	0.56	0.49	0.36	0.53	0.60	0.63	0.05	0.02	*	-	-	+	+	-	-	+	+
Taura	Ta	0.25	0.65	0.72	0.68	0.53	0.65	0.75	0.75	0.81	0.70	0.50	0.20	0.55	0.63	0.68	0.40	0.27	0.09	*	+	-	-	-	+	+	+
San Pablo	SP	1.00	0.95	1.00	1.00	0.80	0.88	1.00	1.00	0.97	0.95	0.75	0.91	0.92	0.79	0.80	0.00	0.12	0.30	0.76	*	+	+	-	+	-	+
Jubones	J1	0.50	0.63	0.66	0.64	0.57	0.62	0.68	0.68	0.71	0.64	0.53	0.53	0.66	0.63	0.68	0.52	0.30	0.36	0.45	0.65	*	-	-	+	+	-
	J2	0.58	0.67	0.72	0.70	0.56	0.62	0.74	0.74	0.74	0.68	0.53	0.59	0.71	0.60	0.68	0.74	0.34	0.39	0.46	0.82	0.12	*	-	+	+	-
	J3	0.40	0.59	0.63	0.60	0.51	0.58	0.65	0.65	0.70	0.61	0.48	0.43	0.60	0.59	0.63	0.02	0.00	0.10	0.32	0.30	0.18	0.25	*	-	+	+
	J4	0.59	0.70	0.73	0.71	0.64	0.70	0.75	0.75	0.79	0.72	0.60	0.60	0.70	0.69	0.71	0.05	0.11	0.22	0.52	0.29	0.30	0.51	0.05	*	-	+
	J5	0.86	0.88	0.91	0.90	0.79	0.84	0.91	0.91	0.91	0.88	0.76	0.83	0.86	0.79	0.80	0.17	0.24	0.35	0.76	0.36	0.67	0.78	0.33	0.22	*	+
Santa Rosa	SR	0.74	0.77	0.80	0.79	0.70	0.74	0.81	0.81	0.81	0.78	0.66	0.75	0.81	0.71	0.76	0.82	0.53	0.56	0.68	0.86	0.13	0.15	0.45	0.60	0.83	*

Genetic differentiation level

Very great	Great	Moderate	Little
≥0.25	0.16 - 0.24	0.06 - 0.15	≤0.05

Table 5. Pairwise F_{ST} of populations of *Rhoadsia* in western Ecuador based on mitochondrial gene cytochrome b. *F_{ST} values are below the diagonal. Significance values are above the diagonal (significance level=0.05): + (significant), - (not significant). F_{ST} values were colored to different shades of grey to represent the level of genetic differentiation (see legend at the bottom of the table).*

Drainages		Sua	Atacames	Esmeraldas					Guayas							Taura	San Pablo	Jubones					Santa Rosa
	Sites	Su	At	E2B	E9	E4	E6	E8	G6	G7	G1	G2	G5	G3	G4	Ta	SP	J1	J2	J3	J4	J5	SR
Sua	Su	*	-	+	+	+	+	+	-	+	+	+	-	+	+	-	+	+	+	+	+	+	+
Atacames	At	0.00	*	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Esmeraldas	E2B	0.98	0.98	*	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	E9	0.84	0.93	0.74	*	+	+	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	E4	0.82	0.89	0.28	0.15	*	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	E6	0.80	0.88	0.22	0.20	0.00	*	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	E8	0.77	0.86	0.51	0.05	0.07	0.08	*	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Guayas	G6	0.68	0.93	0.89	0.49	0.62	0.62	0.51	*	-	-	-	-	+	+	-	+	+	+	+	+	+	+
	G7	0.66	0.90	0.83	0.27	0.48	0.48	0.29	0.00	*	-	-	-	+	+	-	+	+	+	+	+	+	+
	G1	0.60	0.74	0.59	0.18	0.34	0.35	0.16	0.22	0.03	*	-	+	+	+	-	+	+	+	+	+	+	+
	G2	0.75	0.86	0.77	0.29	0.46	0.47	0.26	0.41	0.14	0.05	*	+	+	+	+	+	+	+	+	+	+	+
	G5	0.97	0.99	0.98	0.90	0.88	0.87	0.86	0.70	0.76	0.72	0.85	*	+	+	-	+	+	+	+	+	+	+
	G3	0.77	0.90	0.85	0.71	0.76	0.76	0.73	0.63	0.62	0.54	0.67	0.49	*	+	-	+	+	+	+	+	+	+
	G4	0.76	0.87	0.81	0.70	0.75	0.75	0.73	0.66	0.65	0.59	0.68	0.38	0.13	*	-	+	+	+	+	+	+	+
Taura	Ta	0.95	0.98	0.97	0.83	0.83	0.82	0.79	0.54	0.61	0.51	0.73	0.75	0.00	0.16	*	+	-	-	+	+	+	-
San Pablo	SP	0.87	0.94	0.91	0.83	0.83	0.82	0.81	0.78	0.77	0.68	0.79	0.66	0.42	0.40	0.48	*	+	+	+	+	+	+
Jubones	J1	0.89	0.93	0.92	0.84	0.85	0.85	0.83	0.83	0.80	0.68	0.79	0.79	0.40	0.45	0.33	0.63	*	-	+	+	+	+
	J2	0.88	0.93	0.91	0.83	0.84	0.84	0.82	0.81	0.79	0.66	0.78	0.76	0.31	0.38	0.15	0.58	0.02	*	+	+	+	+
	J3	1.00	0.99	0.99	0.97	0.93	0.92	0.92	0.97	0.95	0.79	0.91	0.99	0.68	0.63	0.97	0.85	0.19	0.41	*	+	+	+
	J4	1.00	0.99	0.99	0.97	0.92	0.91	0.91	0.97	0.95	0.80	0.92	0.98	0.71	0.61	0.98	0.22	0.83	0.80	1.00	*	-	+
	J5	1.00	1.00	0.99	0.97	0.93	0.92	0.92	0.97	0.96	0.81	0.93	0.99	0.74	0.63	0.98	0.25	0.84	0.82	1.00	0.00	*	+
Santa Rosa	SR	0.89	0.93	0.91	0.83	0.85	0.84	0.83	0.82	0.79	0.68	0.79	0.77	0.41	0.45	0.34	0.62	0.08	0.13	0.30	0.83	0.84	*

Genetic differentiation level

Very great	Great	Moderate	Little
≥0.25	0.16 - 0.24	0.06 - 0.15	≤0.05

Table 6. Pairwise F_{ST} of populations of *Rhoadsia* in western Ecuador based on microsatellite data. *F_{ST} values are below the diagonal. Significance values are above the diagonal (significance level=0.05): + (significant), - (not significant). F_{ST} values were colored to different shades of grey to represent the level of genetic differentiation (see legend at the bottom of the table).*

Drainages		Atacames	Esmeraldas				Guayas				San Pablo	Jubones					Santa Rosa
		At	E2B	E4	E6	E8	G1	G2	G3	G4	SP	J1	J2	J3	J4	J5	SR
Atacames	At	*	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Esmeraldas	E2B	0.18	*	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	E4	0.27	0.02	*	+	+	+	+	+	+	+	+	+	+	+	+	+
	E6	0.37	0.14	0.13	*	+	+	+	+	+	+	+	+	+	+	+	+
	E8	0.27	0.05	0.13	0.19	*	+	+	+	+	+	+	+	+	+	+	+
Guayas	G1	0.24	0.03	0.09	0.18	0.05	*	+	+	+	+	+	+	+	+	+	+
	G2	0.34	0.08	0.14	0.21	0.10	0.05	*	+	+	+	+	+	+	+	+	+
	G3	0.37	0.08	0.12	0.22	0.12	0.05	0.13	*	-	+	+	+	+	+	+	+
	G4	0.33	0.06	0.11	0.21	0.12	0.05	0.11	0.01	*	+	+	+	+	+	+	+
San Pablo	SP	0.43	0.18	0.23	0.30	0.21	0.15	0.21	0.14	0.16	*	+	+	+	+	+	+
Jubones	J1	0.32	0.15	0.21	0.24	0.15	0.09	0.21	0.16	0.18	0.23	*	-	+	+	+	+
	J2	0.30	0.10	0.20	0.26	0.16	0.09	0.14	0.15	0.17	0.21	0.00	*	+	+	+	+
	J3	0.37	0.14	0.24	0.29	0.17	0.10	0.20	0.18	0.17	0.25	0.05	0.07	*	+	+	+
	J4	0.45	0.19	0.25	0.34	0.22	0.16	0.22	0.18	0.18	0.12	0.27	0.26	0.27	*	+	+
	J5	0.49	0.23	0.27	0.35	0.27	0.21	0.27	0.20	0.20	0.13	0.32	0.31	0.32	0.04	*	+
Santa Rosa	SR	0.26	0.07	0.16	0.22	0.12	0.07	0.14	0.13	0.14	0.21	0.06	0.07	0.10	0.24	0.29	*

Genetic differentiation level

Very great	Great	Moderate	Little
≥0.25	0.16 - 0.24	0.06 - 0.15	≤0.05

Figures

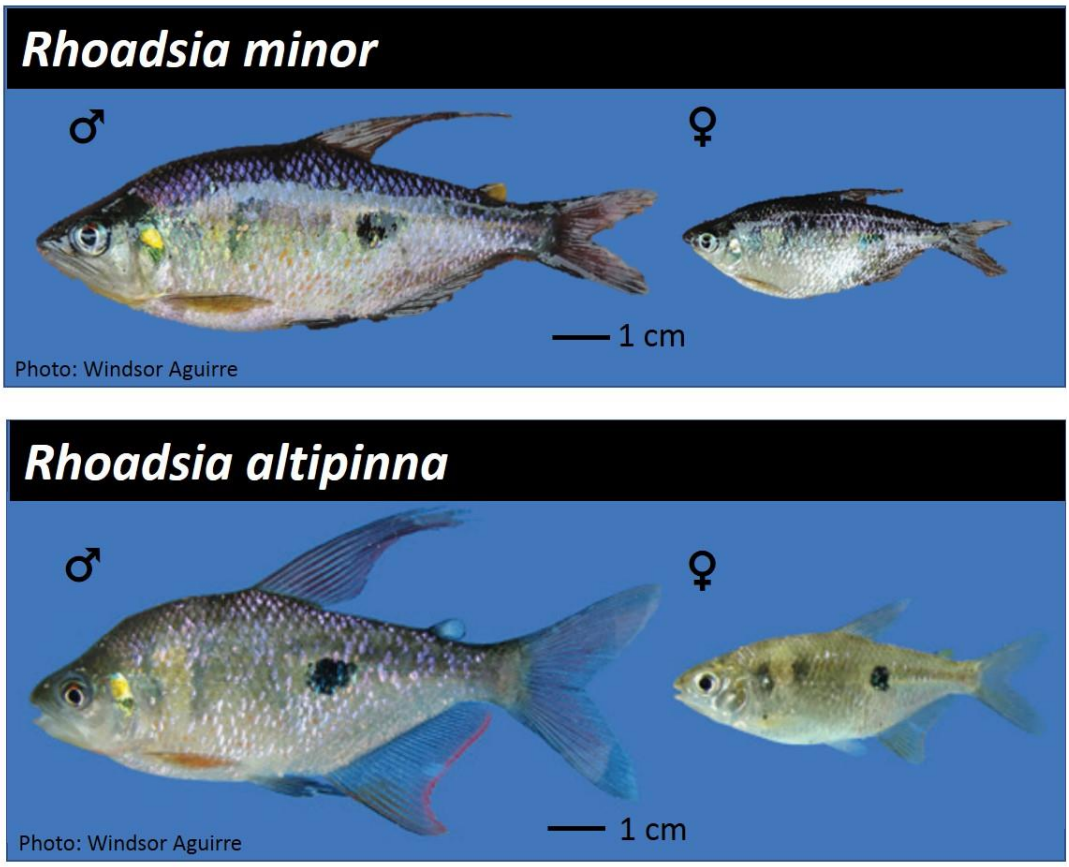


Figure 1. The two species of *Rhoadsia* recognized in western Ecuador. *Pictures taken from* Jimenez-Prado et al. (2015)

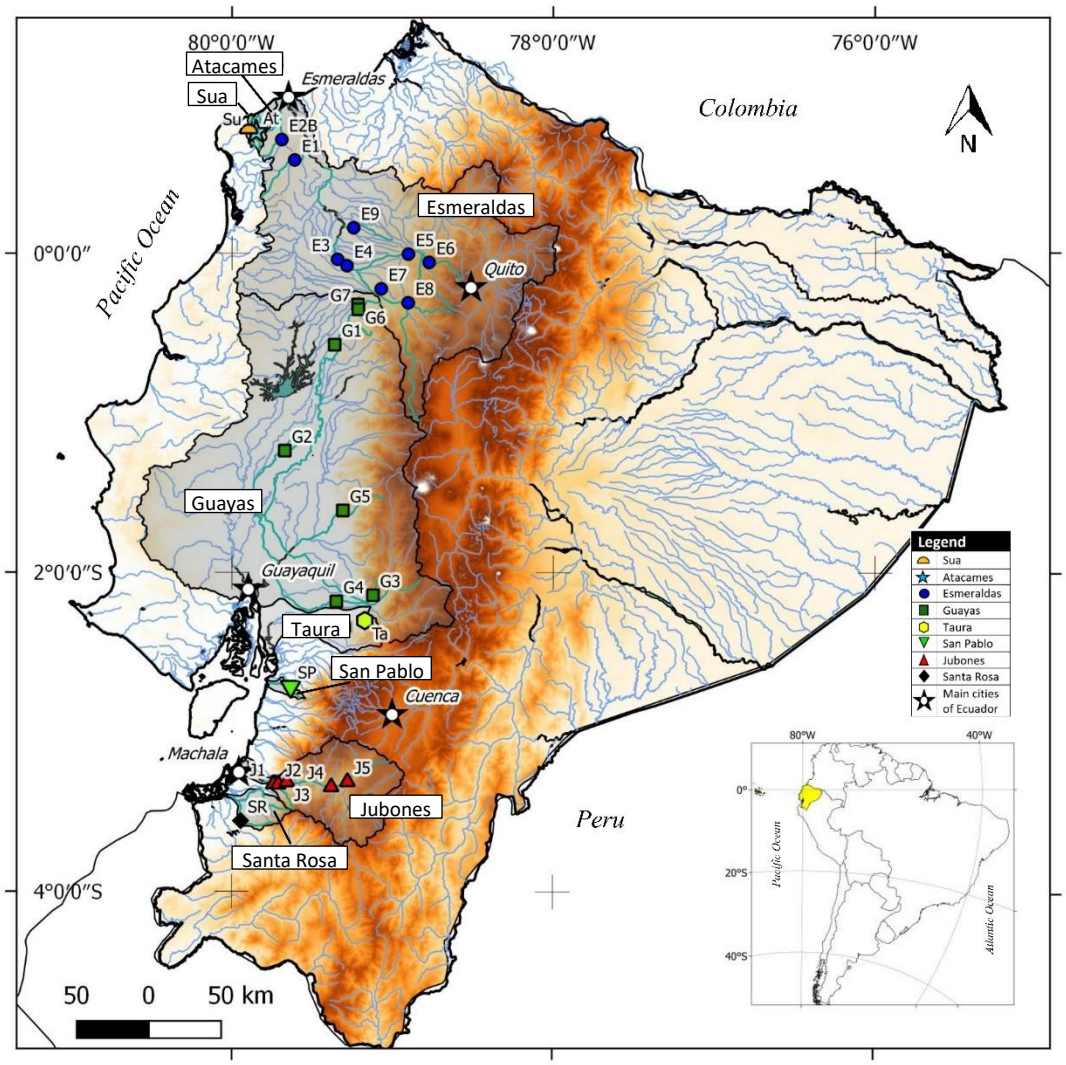


Figure 2. Map of Ecuador showing the sites and drainages in western Ecuador sampled for this study. The drainages are colored in grey and labeled in bold. The sites are represented with different symbols referring to their respective drainage. Streams related to the sites sampled are highlighted. Meaning of code of sites can be seen in Table 1. Scale of map, 1:390000

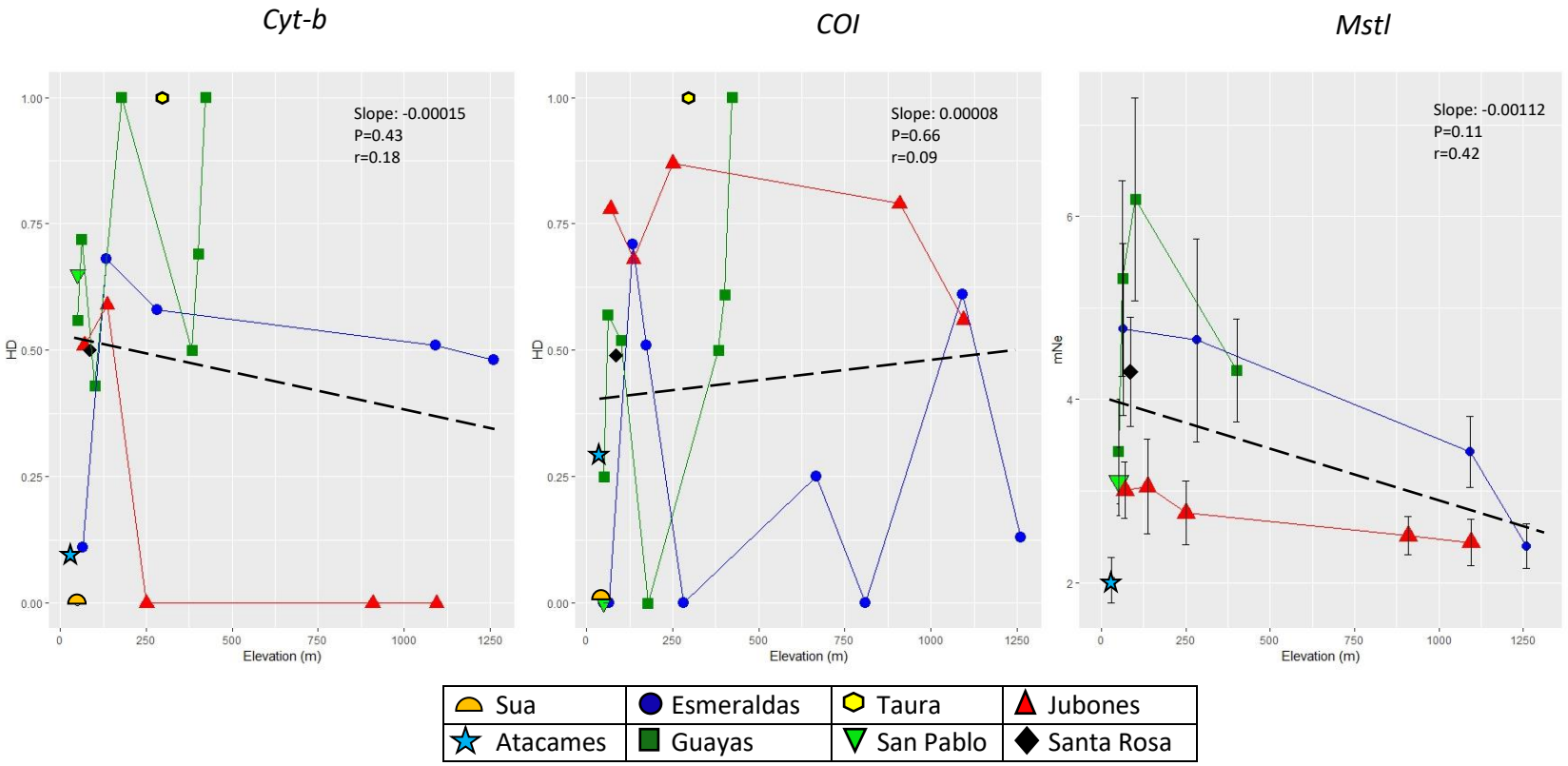


Figure 3. Linear regression of the genetic diversity as a function of elevation based on the haplotype diversity (HD) for the mitochondrial genes cytochrome b (*Cyt-b*) and Cytochrome oxidase I (*COI*) and the mean effective number of alleles (mNe) for the microsatellites (error bars represent the standard deviation of the mNe).

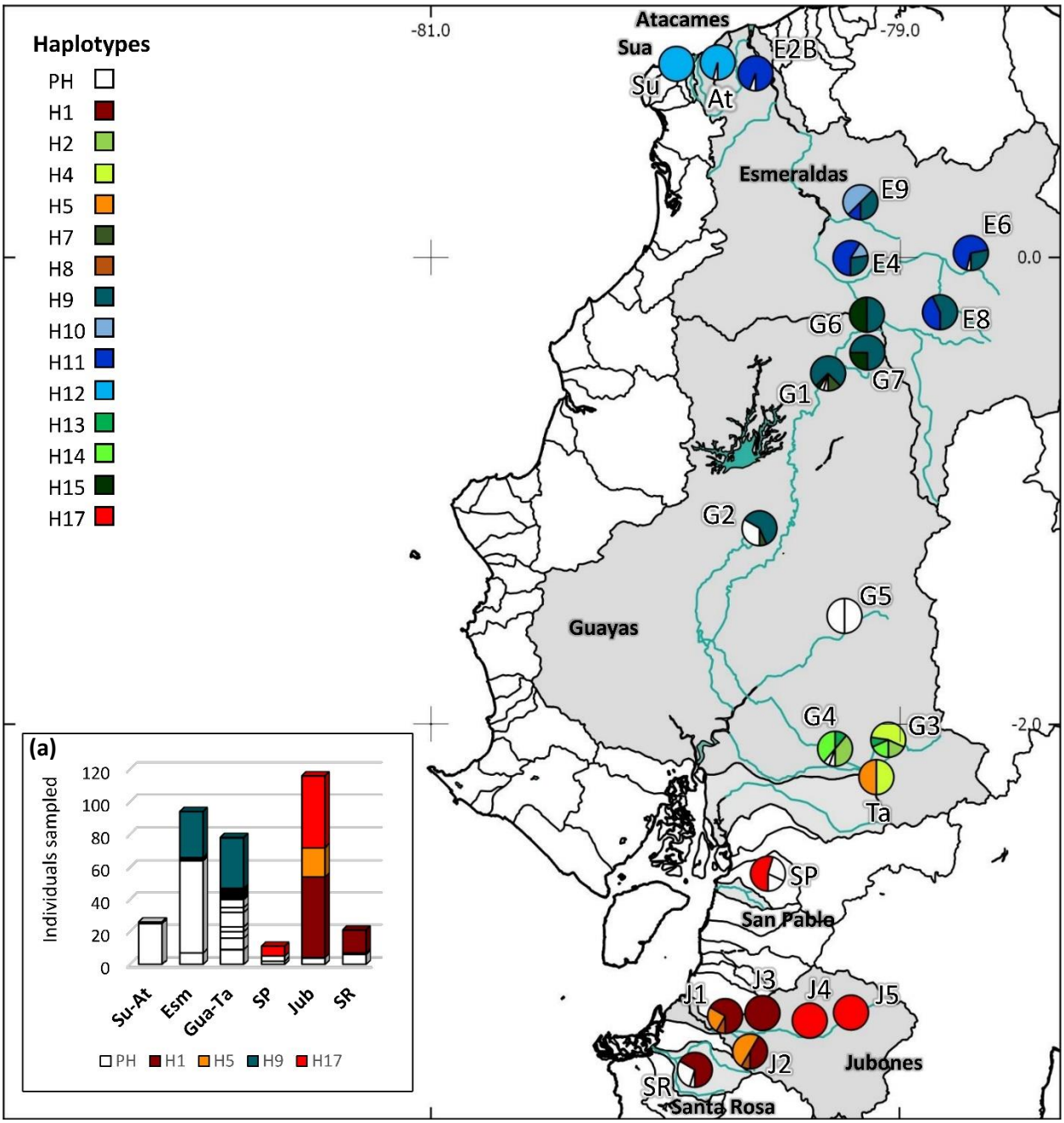


Figure 4. Frequency of Cytochrome b haplotypes per site. *Haplotypes colored in white correspond to private haplotypes (PH) (i.e. haplotypes that are unique to a particular site). The bottom left corner graph (a) presents the haplotype frequency showing the number of individuals sampled in each drainage. Su-At (Sua and Atacames combined), Esm (Esmeraldas), Gua-Ta (Guayas and Taura combined), SP (San Pablo), Jub (Jubones), and SR (Santa Rosa).*

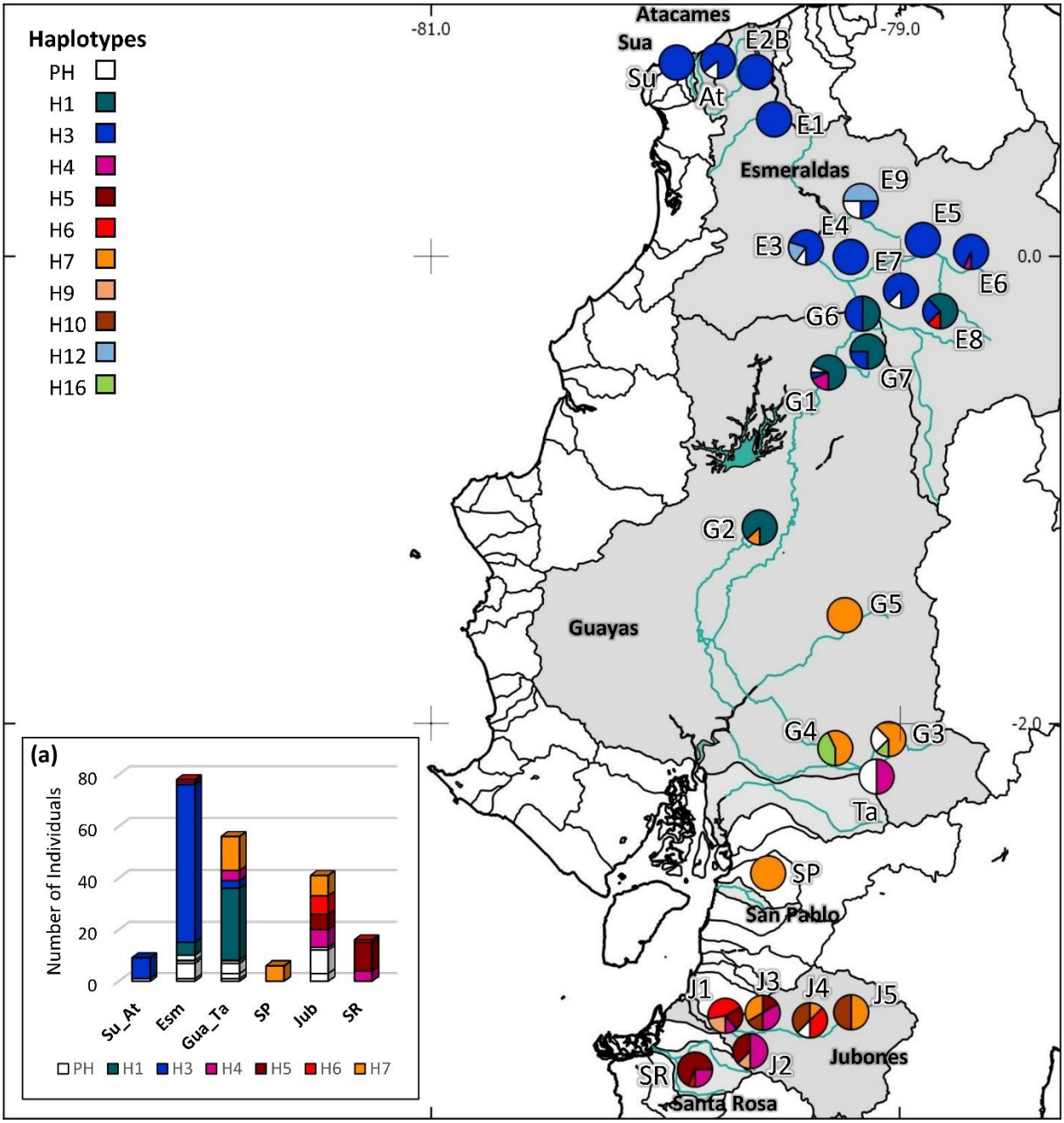


Figure 5. Frequency of Cytochrome oxidase I haplotypes per site. *Haplotypes colored in white correspond to private haplotypes (PH) (i.e. haplotypes that are unique to a particular site). The bottom left corner graph (a) presents the haplotype frequency showing the number of individuals sampled in each drainage. Su-At (Sua and Atacames combined), Esm (Esmeraldas), Gua-Ta (Guayas and Taura combined), SP (San Pablo), Jub (Jubones), and SR (Santa Rosa)*

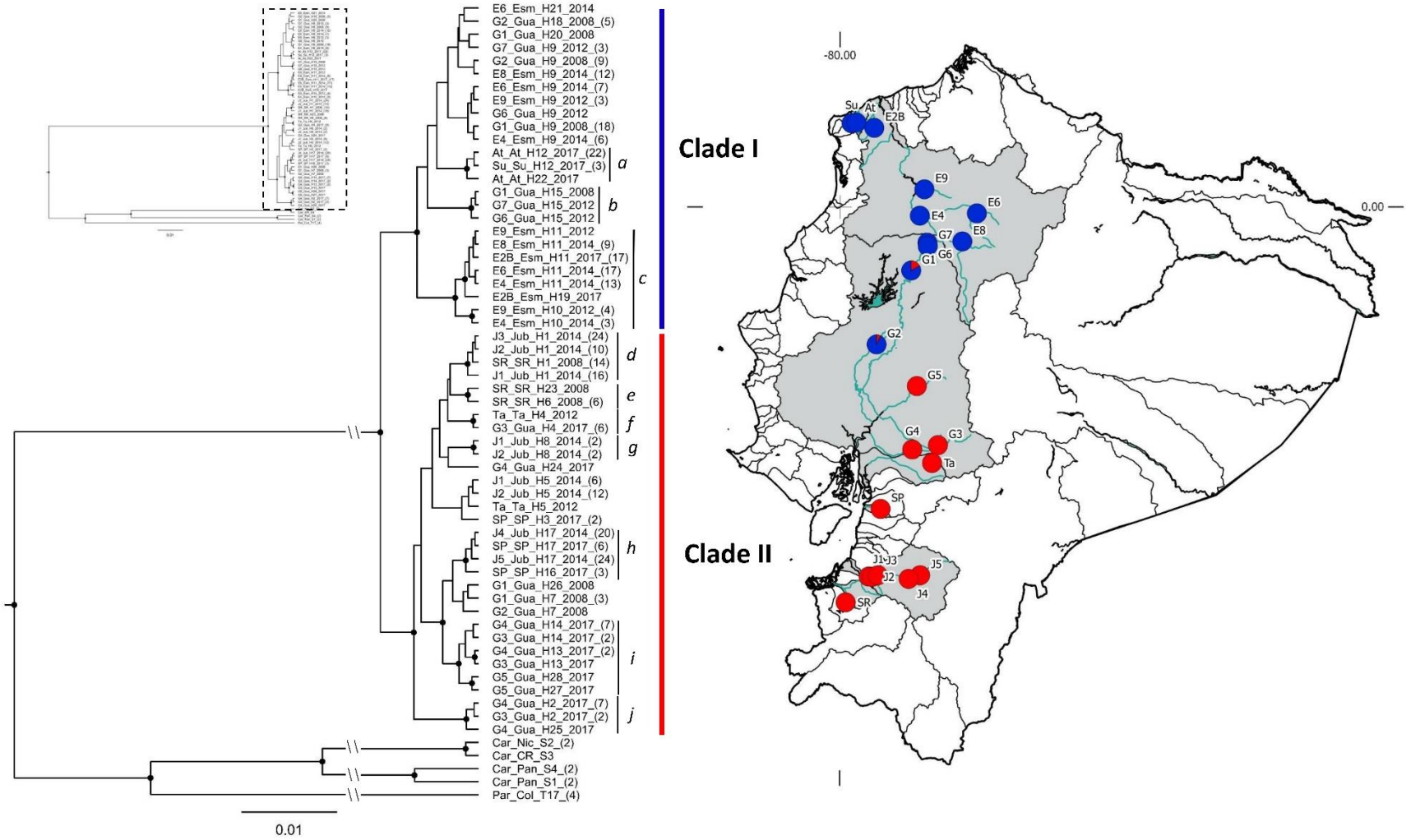


Figure 6. Bayesian phylogenetic tree based on the mitochondrial gene Cytochrome b. *Nodes supported by posterior probabilities greater or equal to 0.8 are displayed with black dots. Clade I (blue) comprised populations from the northern region while Clade II (red) is represented by populations from the southern region. The format for the label of the in-group is: Site-code_Drainage-code_Haplotype-number_Year_(number of individuals with identical haplotype). For the out-group, it is: Species-code_Country-code_Sample-ID_(Number of individuals with identical haplotype).*

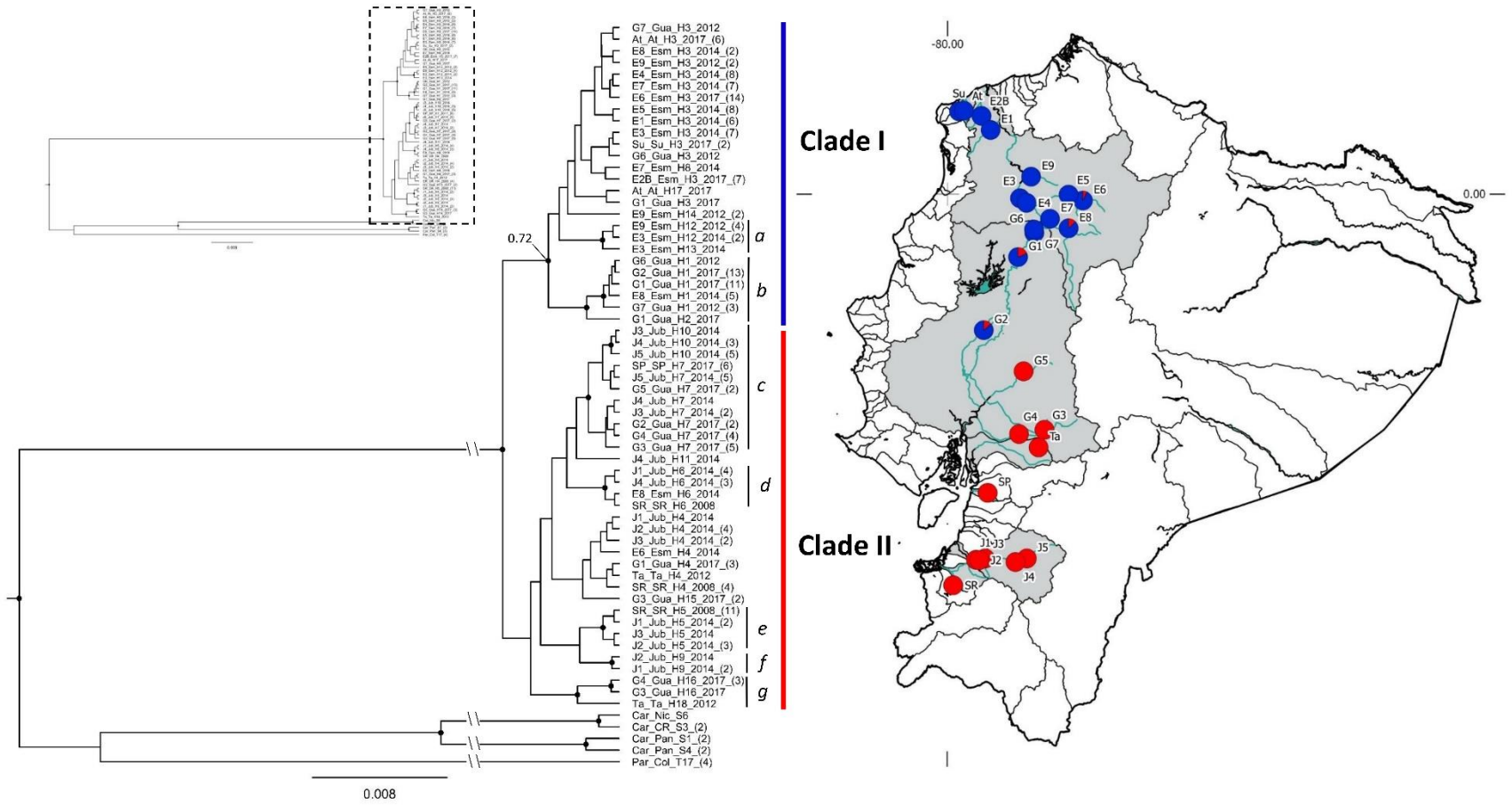


Figure 7. Bayesian phylogenetic tree based on the mitochondrial gene cytochrome oxidase I. *Nodes supported by posterior probabilities greater or equal to 0.8 are displayed with black dots unless specified. Clade I (blue) comprised populations from the northern region while Clade II (red) is represented by populations from the southern region. The format for the label of the in-group is: Site-code_Drainage-code_Haplotype-number_Year_(number of individuals with identical haplotype). For the out-group, it is: Species-code_Country-code_Sample-ID_(Number of individuals with identical haplotype).*

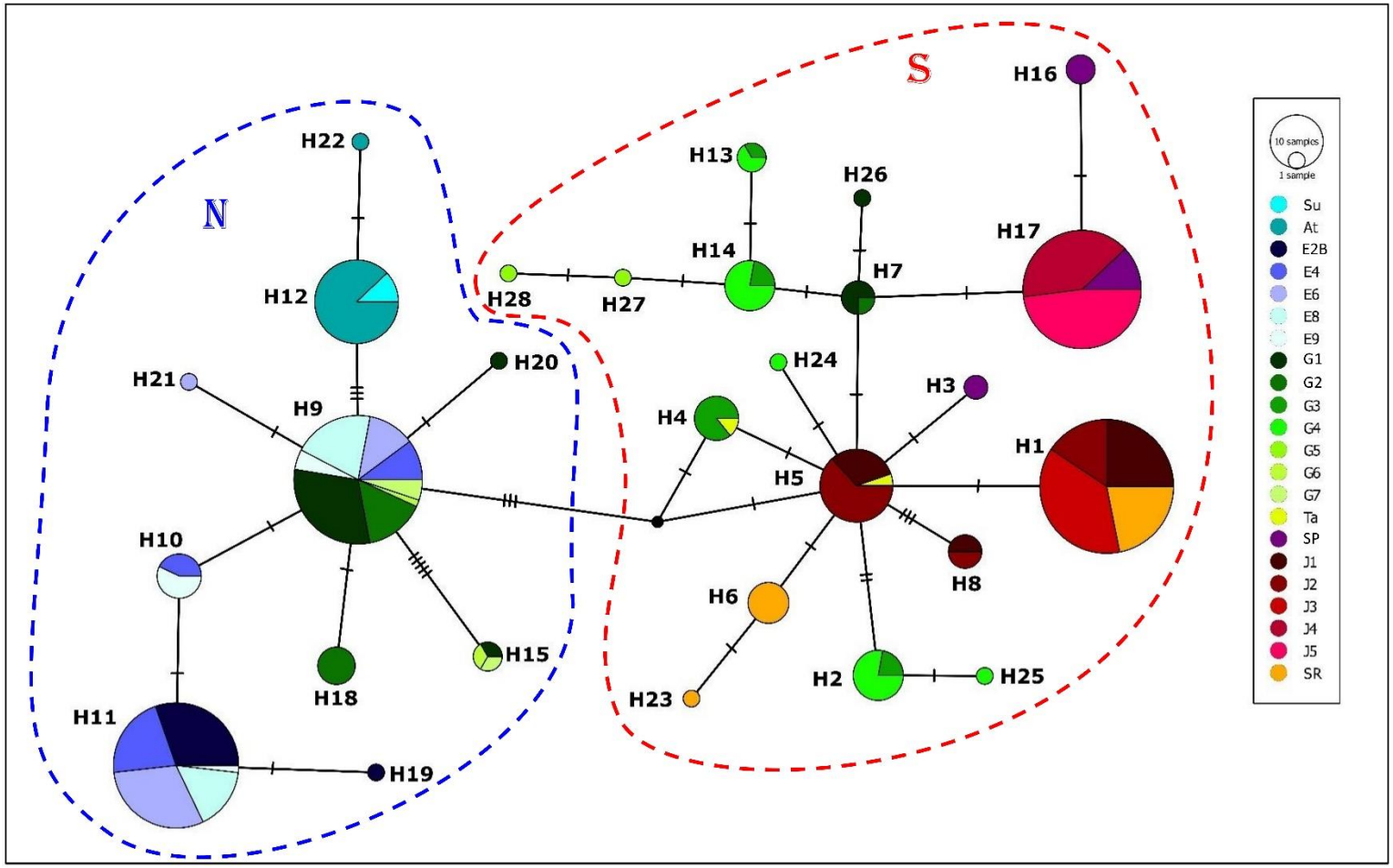


Figure 8. Haplotype network for the mitochondrial gene cytochrome b color coded by site. *The size of the circles is proportional to the haplotype frequency. The number of mutations between the haplotypes are represented by hatch marks. The populations within drainages are represented by different color shades (See legend). Haplotypes fall into two groups: the northern (N, enclosed by blue dashed line) and the southern group (S, enclosed by red dashed line).*

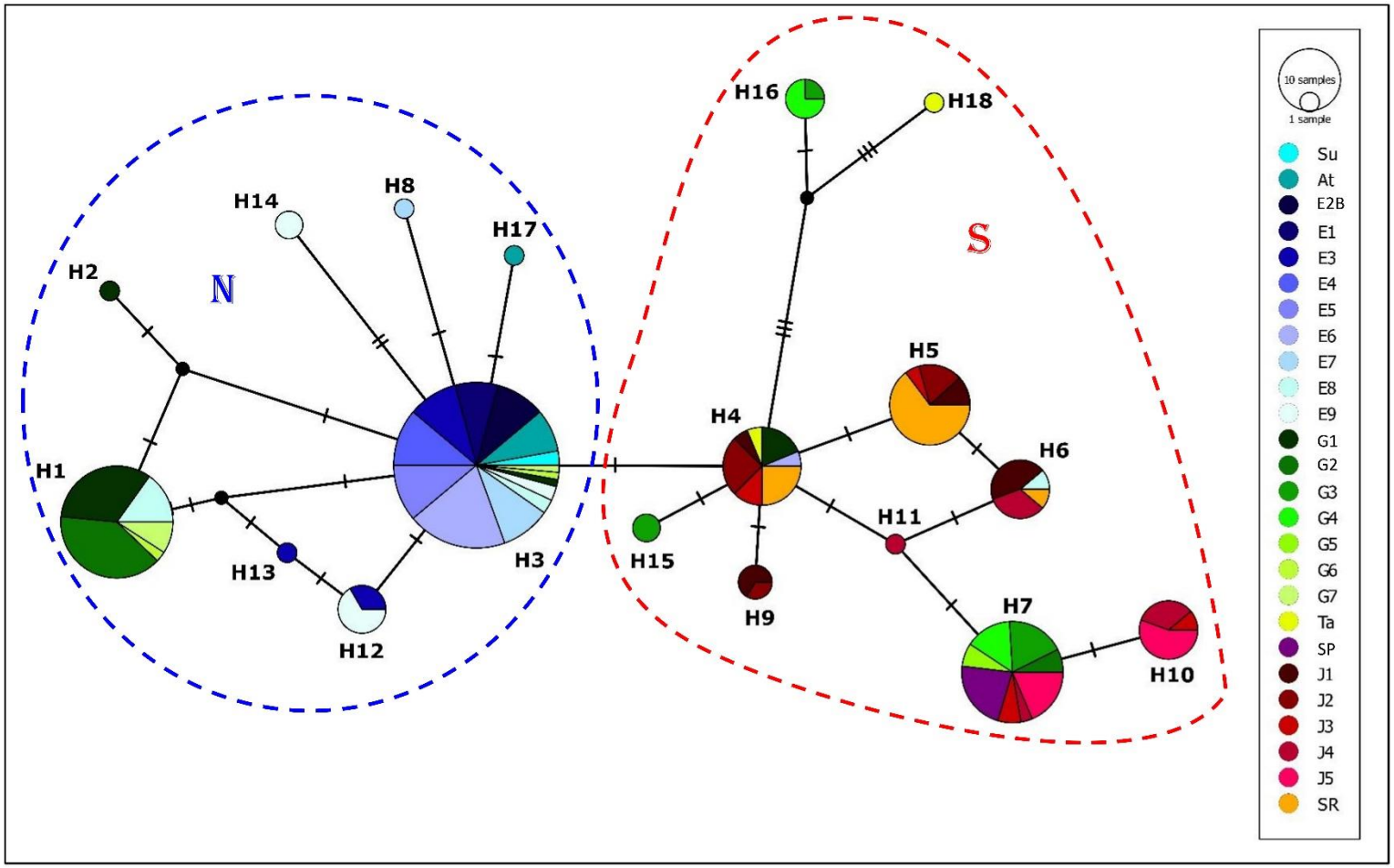


Figure 9. Haplotype network for the mitochondrial gene cytochrome oxidase I color coded by site. *The size of the circles is proportional to the haplotype frequency. The number of mutations between the haplotypes are represented by hatch marks. The populations within drainages are represented by different color shades (See legend). Haplotypes fall into two groups: the northern (N, enclosed by blue dashed line) and the southern group (S, enclosed by red dashed line).*

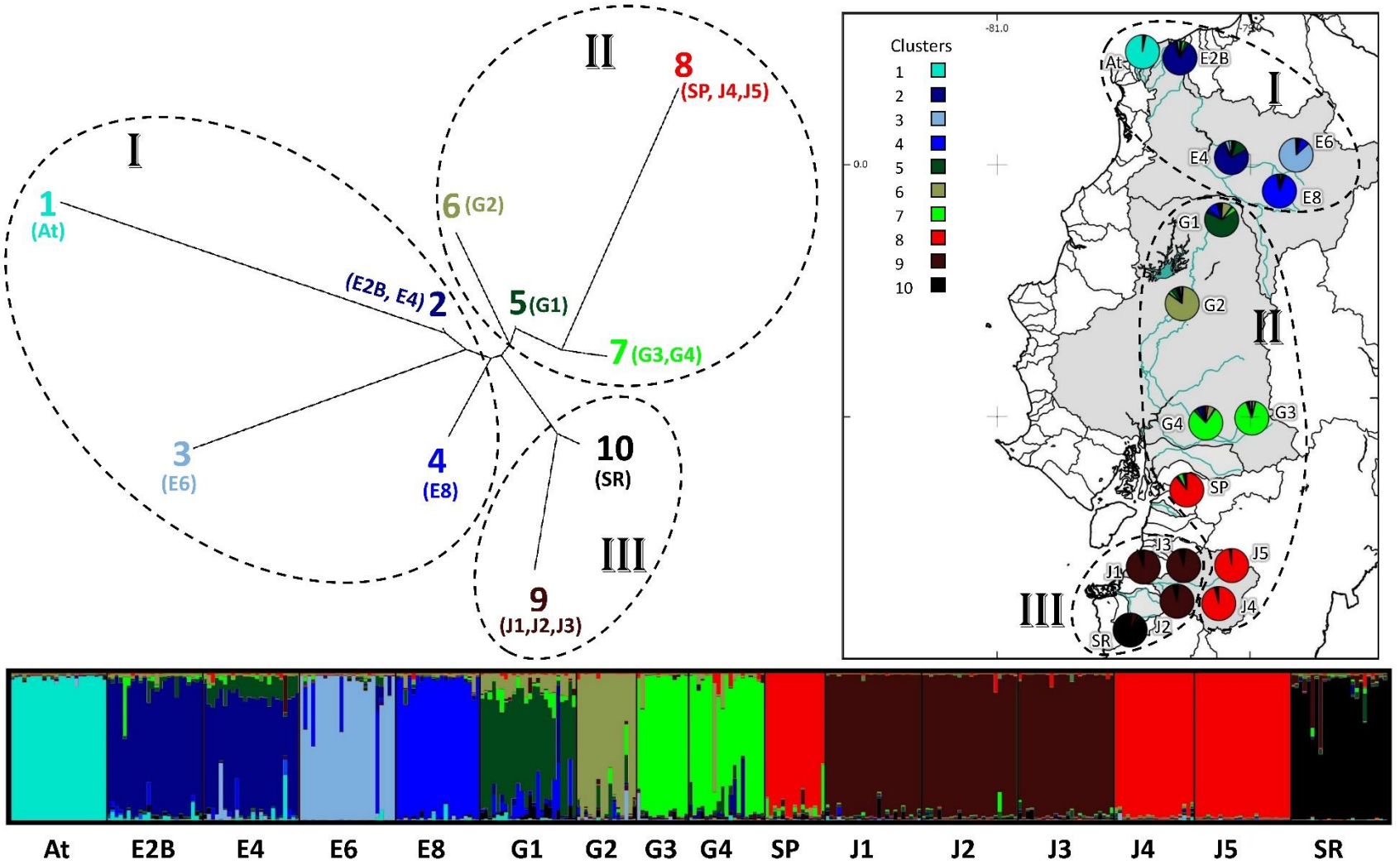


Figure 10. Structure analysis of populations of *Rhoadsia* in Ecuador based on 12 microsatellites. Neighbor-Joining of clusters (K=10) inferred in STRUCTURE (top left). Proportion of membership of each cluster per site (top right). Proportion of membership of each cluster assigned to each individual (Bottom). *Clusters fell into three main groups (I, II, and III) based on the distance matrix generated in STRUCTURE.*

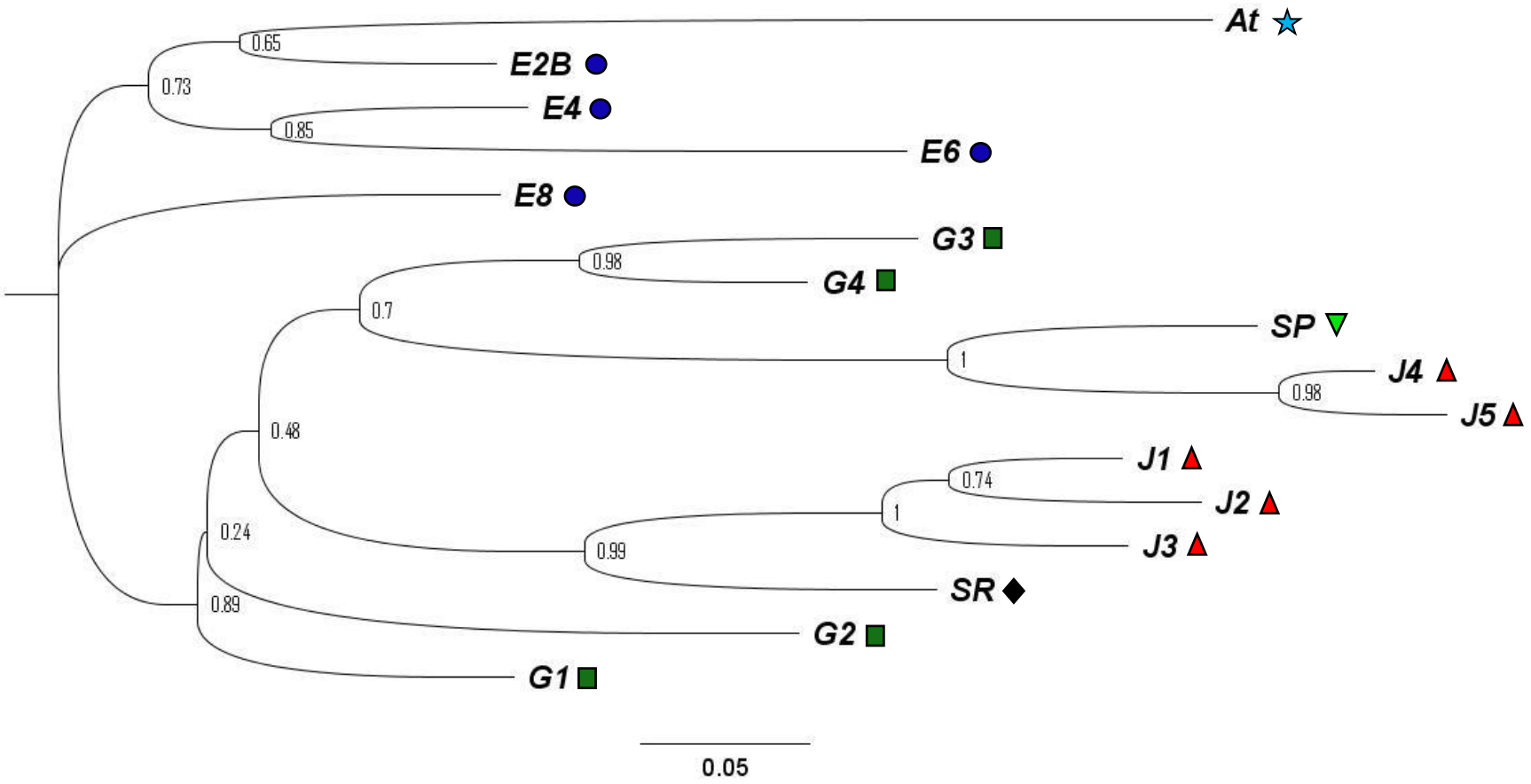
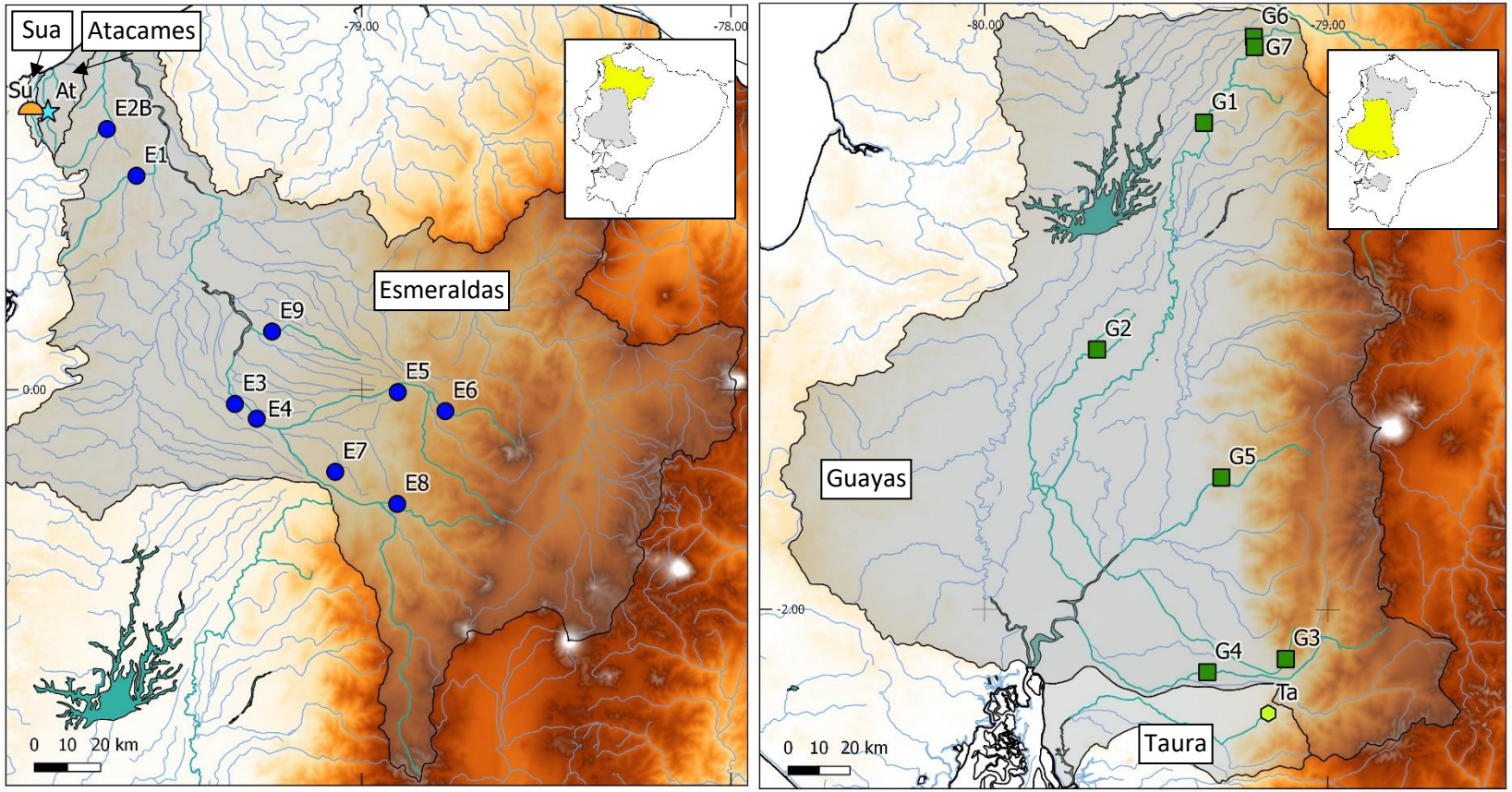


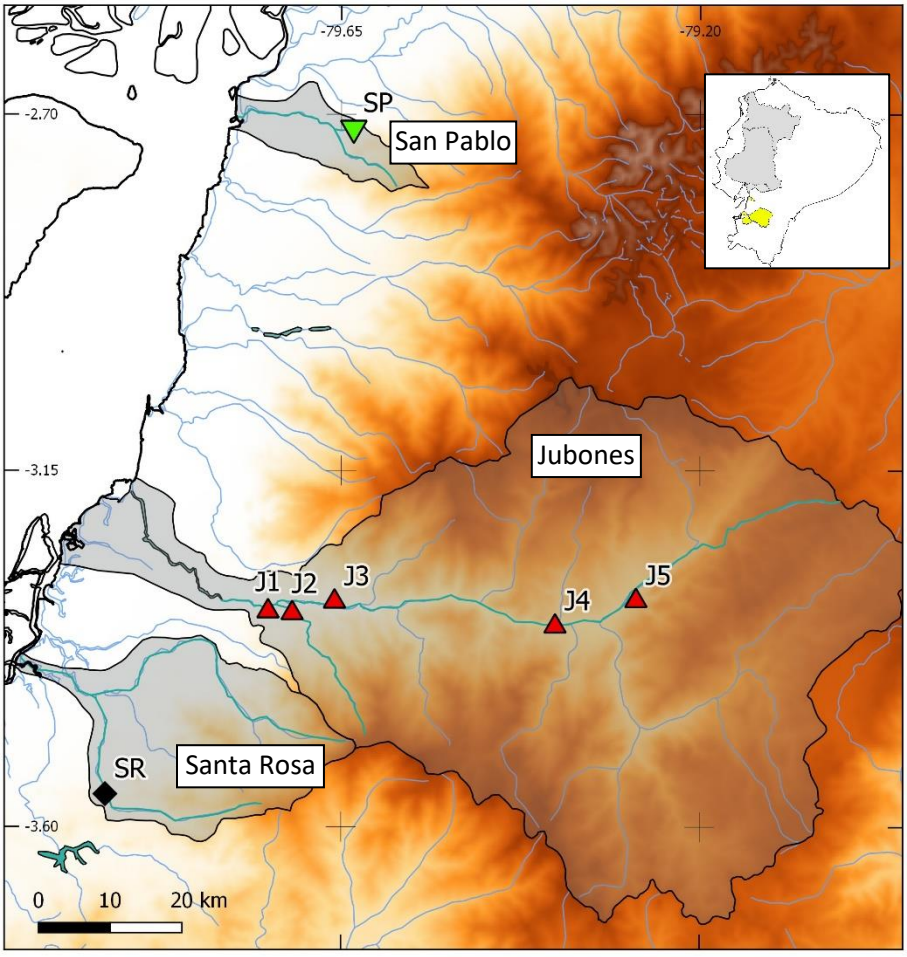
Figure 11. Neighbor-Joining tree of the populations sampled based on the microsatellite data. Branch support values were calculated after 1000 bootstrap replications In POPTREE2.

Appendix

Appendix S1. Close up of drainages Sua, Atacames, Esmeraldas (left), Guayas, and Taura (right).



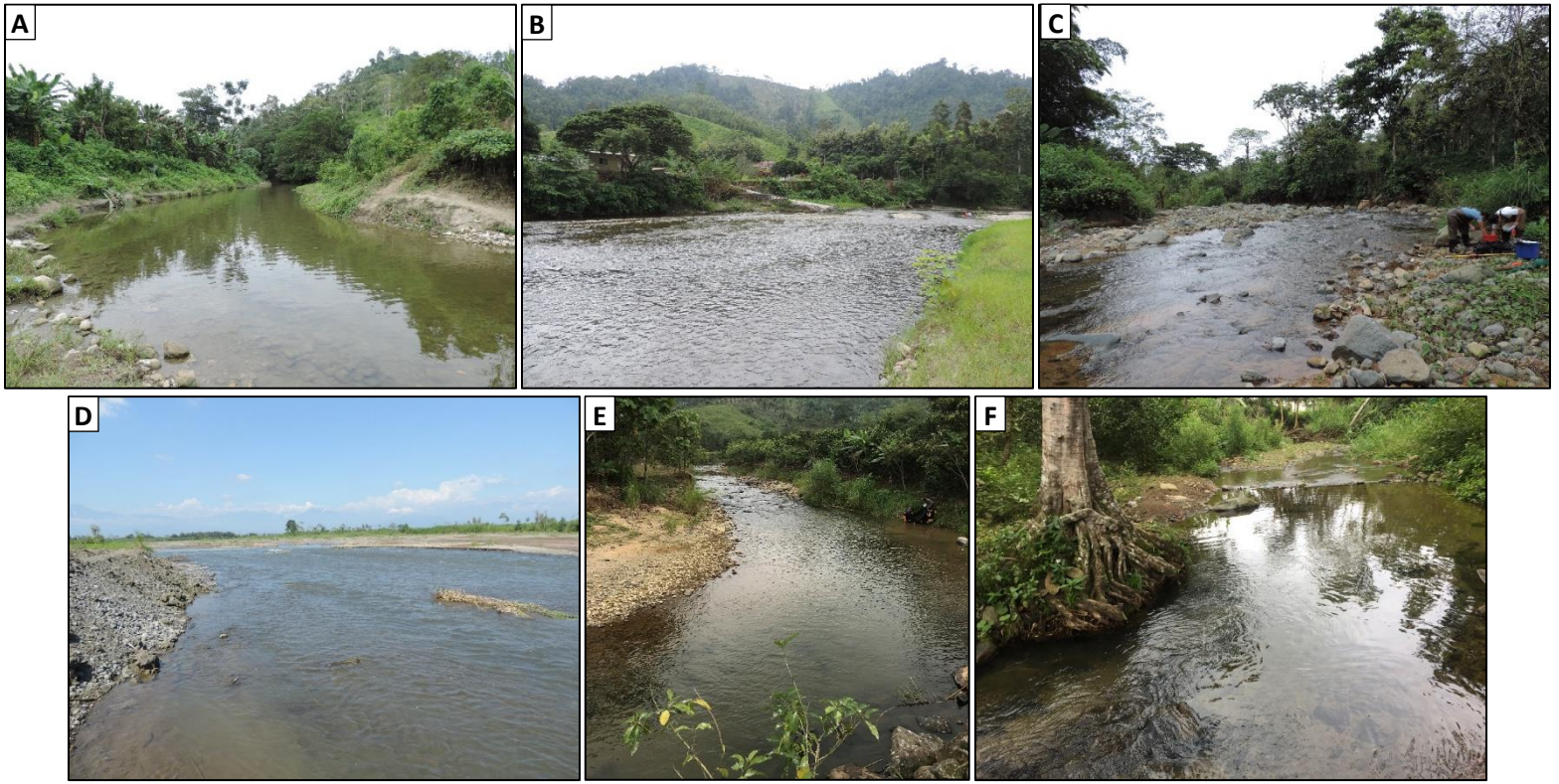
Appendix S2. Close up of the drainages San Pablo, Jubones, and Santa Rosa.



Appendix S3. Catalog information of specimens obtained from museums. *ROM (Royal Ontario Museum), STRI (Smithsonian Tropical Research Institute Panama), N (Sample size), ID (specimen Identification in this study).*

Source	Country	Drainage	Site	Code	N	ID	Species	Catalog
ROM	Ecuador	Esmeraldas	Silanchi	E9	8	T1 -T8	*Rhoadsia minor*	T13569 - T13576
	Ecuador	Guayas	Otongo	G7	4	T11- T14	*Rhoadsia altipinna*	T13677 - T13680
	Ecuador	Guayas	Chiguilpe	G6	2	T9 - T10	*Rhoadsia altipinna*	T13638 -T13639
	Ecuador	Taura	Cutahay	Ta	2	T15 - T16	*Rhoadsia altipinna*	T13874, T13868
	Colombia	Calima	Colombia	Par	4	T17 - T20	*Parastremma sadina*	T25079 - T25082
STRI	Panama	N/A	Guna Yala	Car	1	S1	*Carlana eigenmanni*	STRI-1676
	Nicaragua	N/A	Rio Mena	Car	1	S2	*Carlana eigenmanni*	STRI-14863
	Costa Rica	N/A	Alajuela	Car	1	S3	*Carlana eigenmanni*	STRI-2160
	Panama	N/A	Guna Yala	Car	1	S4	*Carlana eigenmanni*	STRI-2618
	Panama	N/A	Guna Yala	Car	1	S5	*Carlana eigenmanni*	STRI-4889
	Nicaragua	N/A	Chonatales	Car	1	S6	*Carlana eigenmanni*	STRI-14434
	Panama	N/A	Guna Yala	Car	1	S7	*Carlana eigenmanni*	STRI-1675

Appendix S4. Picture of the sites. A, Atacames river (At, 29 masl); B, Teaone river (E2B, 64 masl); C, Chague Grande river (G3, 401 masl); D, Chimbo river (G4, 62 masl); E, Caluma river (G5, 178 masl); F, Estero Mina (SP, 50 masl).



Appendix S5. Electrofishing Technique at the Caluma river (Top, 178 masl). *Rhoadsia* fish just sampled from the Chague Grande river (Bottom, 401 masl).

Appendix S6. Primers used for amplification of microsatellite data modified from Loh et al., 2013. *The first portion of the forward primer (F) represent the binding site for the universal M13 tag implemented in Schuelke's (2000) method. Bp (base pair), T_A (C°) (annealing temperature used for amplification (Celsius degree)). *Annealing temperature modified for optimization of PCR product.*

Locus		Primer Sequence (5'-3')	bp	T_A (C°)	Size range
RA-06	F	**TGTAAAACGACGGCCAGT**TCACGTATCCTCGCTTCCTC	38	54*	155-175
	R	AGTTGGAGTTGGATCGAACC	20		
RA-08	F	**TGTAAAACGACGGCCAGT**TCAGCATGATGACATATTGG	38	52*	130-140
	R	CGTGTGTAATGGTGATGTTG	20		
RA-09	F	**TGTAAAACGACGGCCAGT**CTGAGGCTGGCACATCTGTT	38	55	191-213
	R	TCCTCATAGAGGTCATAGAG	20		
RA-10	F	**TGTAAAACGACGGCCAGT**AACACTAACACAGGAGGAGG	38	55	255-318
	R	AGAAGATGTGGAGAAGCACC	20		
RA-13	F	**TGTAAAACGACGGCCAGT**AAGAGGACATGAGGAACGCT	38	60	238-335
	R	ACACACTCACACTGGCGGTG	20		
RA-16	F	**TGTAAAACGACGGCCAGT**ATCCTCTTCTTCCTGGCCAT	38	55	175-208
	R	ACATGGTAGGTGAATTGGCC	20		
RA-20	F	**TGTAAAACGACGGCCAGT**CTCTACAGTCACTCCTTCCA	38	50	160-285
	R	TGTGCATGCCTGTGTGTAAG	20		
RA-21	F	**TGTAAAACGACGGCCAGT**TTCTGCTGACATAATCCTGC	38	55	204-237
	R	TCATGGCATTACACACAGCT	20		
RA-22	F	**TGTAAAACGACGGCCAGT**CAACACAATTGGCTAACCTG	38	52	154-164
	R	GCAACCGATACACAGATATG	20		
RA-26	F	**TGTAAAACGACGGCCAGT**CCACTGCAATGAAGGAACTC	38	55	158-199
	R	TAGGTCAGGTGCATCAGTCC	20		
RA-27	F	**TGTAAAACGACGGCCAGT**GCAGGTACCATATACACAGT	38	52	136-203
	R	CACTTCTGTACTTCAGTAGC	20		
RA-30	F	**TGTAAAACGACGGCCAGT**GCGCTAAGTGCTATGCTGTC	38	55	192-210
	R	GTCGGCCTGATTGAGATCAT	20		

Appendix S7. Group assignment for the AMOVA based on cytochrome b (Cyt-b), cytochrome oxidase I (COI), and microsatellites (Mstls).

	By drainage		By clade		By cluster (K=10)	
	Groups	Sites	Groups	Sites	Groups	Sites
Cyt-b	Sua	Su	Clade I	Su, At, E2B, E9, E3, E4, E6, E8, G6, G7, G1, G2	1	Su, At
	Atacames	At			2	E2B, E9, E4
	Esmeraldas	E2B, E9, E3, E4, E6, E8			3	E6
					4	E8
	Guayas	G6, G7,G1, G2, G5, G3, G4			5	G6, G7, G1
			Clade II	G5, G3, G4, Ta, SP, J1, J2, J3, J4, J5, SR	6	G2
	Taura	Ta			7	G5, G3, G4, Ta
	San Pablo	SP			8	SP, J4, J5
	Jubones	J1, J2, J3, J4, J5			9	J1,J2, J3
	Santa Rosa	SR			10	SR
COI	Sua	Su	Clade I	Su, At, E2B, E1, E9, E3, E4, E5, E6, E7, E8, G6, G7, G1, G2	1	Su, At
	Atacames	At			2	E2B, E1, E9, E3, E4
	Esmeraldas	E2B, E1, E9, E3, E4, E5, E6, E7, E8			3	E5, E6
					4	E7, E8
	Guayas	G6, G7,G1, G2, G5, G3, G4			5	G6, G7, G1
			Clade II	G5, G3, G4, Ta, SP, J1, J2, J3, J4, J5, SR	6	G2
	Taura	Ta			7	G5, G3, G4, Ta
	San Pablo	SP			8	SP, J4, J5
	Jubones	J1, J2, J3, J4, J5			9	J1,J2, J3
	Santa Rosa	SR			10	SR
Mstls	Atacames	At	Clade I	At, E2B, E4, E6, E8, G1, G2	1	At
	Esmeraldas	E2B, E4, E6, E8			2	E2B, E4
					3	E6
					4	E8
	Guayas	G1, G2, G3, G4			5	G1
			Clade II	G3, G4, SP, J1, J2, J3, J4, J5, SR	6	G2
	San Pablo	SP			7	G3, G4
	Jubones	J1, J2, J3, J4, J5			8	SP, J4, J5
					9	J1,J2, J3
	Santa Rosa	SR			10	SR

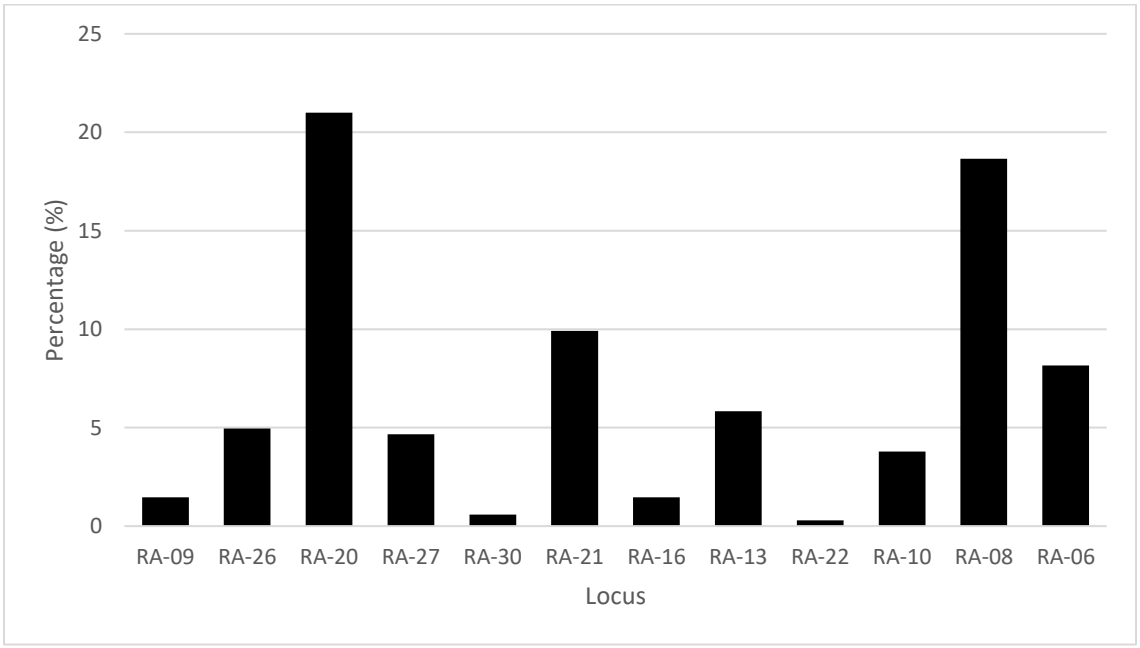
Appendix S8. Neutrality test of populations of *Rhoadsia* based on the mitochondrial gene cytochrome b. *N (sample size), S (number of polymorphic sites), Pi (mean number of pairwise differences), Ob.A (observed number of alleles), Exp.A (expected number of alleles).*

Drainage	Site	N	S	Pi	Ob.A	Exp.A	Tajima's D	P-value	Fu's Fs	P-value
Sua	**Su**	3	0	0.00	1	0.00	0.00	1.000	0.00	N.A
Atacames	**At**	23	1	0.09	2	1.31	-1.16	0.156	-0.99	0.072
Esmeraldas	**E2B**	18	1	0.11	2	1.36	-1.16	0.133	-0.79	0.091
	E9	8	2	0.79	3	2.45	0.07	0.621	-0.22	0.222
	E4	22	2	0.92	3	3.53	1.54	0.934	1.13	0.755
	E6	25	3	0.99	3	3.79	0.60	0.743	1.41	0.789
	E8	21	2	1.03	2	3.70	1.96	0.989	3.07	0.916
Guayas	**G6**	2	5	5.00	2	1.83	0.00	1.000	1.61	0.494
	G7	4	5	2.50	2	2.72	-0.80	0.164	2.60	0.866
	G1	24	12	2.03	5	5.68	-1.26	0.105	1.03	0.725
	G2	15	6	1.14	3	3.56	-1.33	0.069	1.22	0.722
	G5	2	1	1.00	2	1.50	0.00	1.000	0.00	0.257
	G3	11	6	2.25	4	4.44	0.40	0.678	1.05	0.736
	G4	18	7	2.54	5	5.77	0.84	0.809	1.18	0.743
Taura	**Ta**	2	1	1.00	2	1.50	0.00	1.000	0.00	0.266
San Pablo	**SP**	11	4	1.42	3	3.58	0.14	0.636	1.32	0.761
Jubones	**J1**	24	4	0.94	3	3.65	-0.33	0.394	1.26	0.744
	J2	24	4	0.99	3	3.74	-0.22	0.463	1.37	0.782
	J3	24	0	0.00	1	0.00	0.00	1.000	0.00	N.A
	J4	20	0	0.00	1	0.00	0.00	1.000	0.00	N.A
	J5	24	0	0.00	1	0.00	0.00	1.000	0.00	N.A
Santa Rosa	**SR**	21	3	1.03	3	3.70	0.61	0.754	1.34	0.779

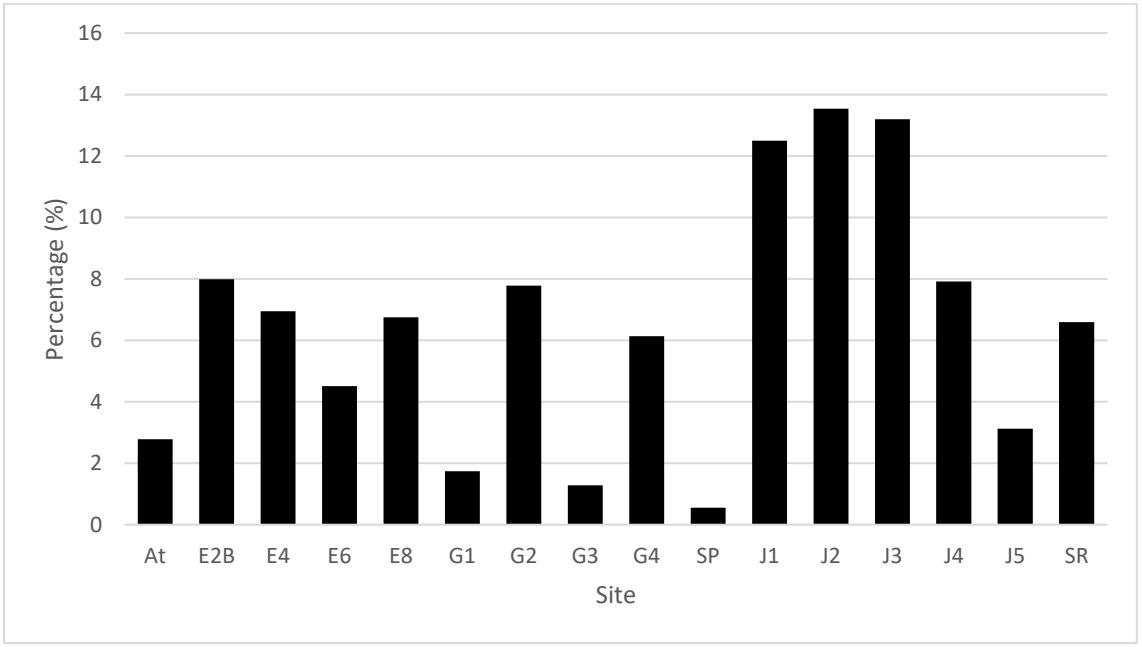
Appendix S9. Neutrality test of populations of *Rhoadsia* based on the mitochondrial gene cytochrome oxidase I. *N (sample size), S (number of polymorphic sites), Pi (mean number of pairwise differences), Ob.A (observed number of alleles), Exp.A (expected number of alleles).*

Drainage	Site	N	S	Pi	Ob.A	Exp.A	Tajima's D	P-value	Fu's Fs	P-value
Sua	**Su**	2	0	0.00	1	0.00	0.00	1.000	0.00	N.A.
Atacames	**At**	7	1	0.29	2	1.60	-1.01	0.232	-0.09	0.229
Esmeraldas	**E2B**	7	0	0.00	1	0.00	0.00	1.000	0.00	N.A.
	E1	6	0	0.00	1	0.00	0.00	1.000	0.00	N.A.
	E9	8	3	1.43	3	3.17	0.97	0.850	0.87	0.712
	E3	10	2	0.67	3	2.43	-0.18	0.378	-0.27	0.322
	E4	8	0	0.00	1	0.00	0.00	1.000	0.00	N.A.
	E5	8	0	0.00	1	0.00	0.00	1.000	0.00	N.A.
	E6	15	1	0.13	2	1.41	-1.16	0.175	-0.65	0.116
	E7	8	1	0.25	2	1.57	-1.05	0.192	-0.18	0.187
	E8	8	5	1.82	3	3.52	-0.25	0.416	1.37	0.781
Guayas	**G6**	2	2	2.00	2	1.67	0.00	1.000	0.69	0.383
	G7	4	2	1.00	2	2.08	-0.71	0.266	1.10	0.627
	G1	16	4	1.31	4	3.90	0.27	0.661	0.35	0.573
	G2	15	5	1.24	2	3.71	-0.66	0.261	3.20	0.936
	G5	2	0	0.00	1	0.00	0.00	1.000	0.00	N.A.
	G3	8	7	2.50	3	4.00	-0.35	0.408	2.08	0.865
	G4	7	6	3.43	2	4.17	2.03	0.991	4.83	0.977
Taura	**Ta**	2	6	6.00	2	1.86	0.00	1.000	1.79	0.507
San Pablo	**SP**	6	0	0.00	1	0.00	0.00	1.000	0.00	N.A.
Jubones	**J1**	9	3	1.44	4	3.34	1.18	0.897	-0.29	0.328
	J2	8	2	0.79	3	2.45	0.07	0.590	-0.22	0.250
	J3	6	4	1.87	4	3.11	0.36	0.675	-0.62	0.203
	J4	8	3	1.64	4	3.37	1.73	0.970	-0.24	0.331
	J5	10	1	0.56	2	2.23	1.46	0.970	1.10	0.658
Santa Rosa	**SR**	16	2	0.53	3	2.43	-0.33	0.371	-0.29	0.292

Appendix S10. Missing microsatellite data per locus.



Appendix S11. Missing microsatellite data per site.



Appendix S12. Haplotype frequency of population of *Rhoadsia* in western Ecuador based on mitochondrial gene cytochrome b.

Site	N	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12	H13	H14	H15	H16	H17	H18	H19	H20	H21	H22	H23	H24	H25	H26	H27	H28
Su	3	-	-	-	-	-	-	-	-	-	-	-	3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
At	23	-	-	-	-	-	-	-	-	-	-	-	22	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-
E2B	18	-	-	-	-	-	-	-	-	-	-	17	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-
E9	8	-	-	-	-	-	-	-	-	3	4	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
E4	22	-	-	-	-	-	-	-	-	6	3	13	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
E6	25	-	-	-	-	-	-	-	-	7	-	17	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-
E8	21	-	-	-	-	-	-	-	-	12	-	9	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
G6	2	-	-	-	-	-	-	-	-	1	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-
G7	4	-	-	-	-	-	-	-	-	3	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-
G1	24	-	-	-	-	-	-	3	-	18	-	-	-	-	-	1	-	-	-	-	1	-	-	-	-	-	1	-	-
G2	15	-	-	-	-	-	-	1	-	9	-	-	-	-	-	-	-	-	5	-	-	-	-	-	-	-	-	-	-
G5	2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	1
G3	11	-	2	-	6	-	-	-	-	-	-	-	-	1	2	-	-	-	-	-	-	-	-	-	-	-	-	-	-
G4	18	-	7	-	-	-	-	-	-	-	-	-	-	2	7	-	-	-	-	-	-	-	-	-	1	1	-	-	-
Ta	2	-	-	-	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
SP	11	-	-	2	-	-	-	-	-	-	-	-	-	-	-	-	3	6	-	-	-	-	-	-	-	-	-	-	-
J1	24	16	-	-	-	6	-	-	2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
J2	24	10	-	-	-	12	-	-	2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
J3	24	24	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
J4	20	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	20	-	-	-	-	-	-	-	-	-	-	-
J5	24	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	24	-	-	-	-	-	-	-	-	-	-	-
SR	21	14	-	-	-	-	6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-
TOTAL	346	64	9	2	7	19	6	4	4	59	7	57	25	3	9	3	3	50	5	1	1	1	1	1	1	1	1	1	1
Freq.		0.311	0.044	0.010	0.034	0.092	0.029	0.019	0.019	0.286	0.034	0.277	0.121	0.015	0.044	0.015	0.015	0.243	0.024	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005

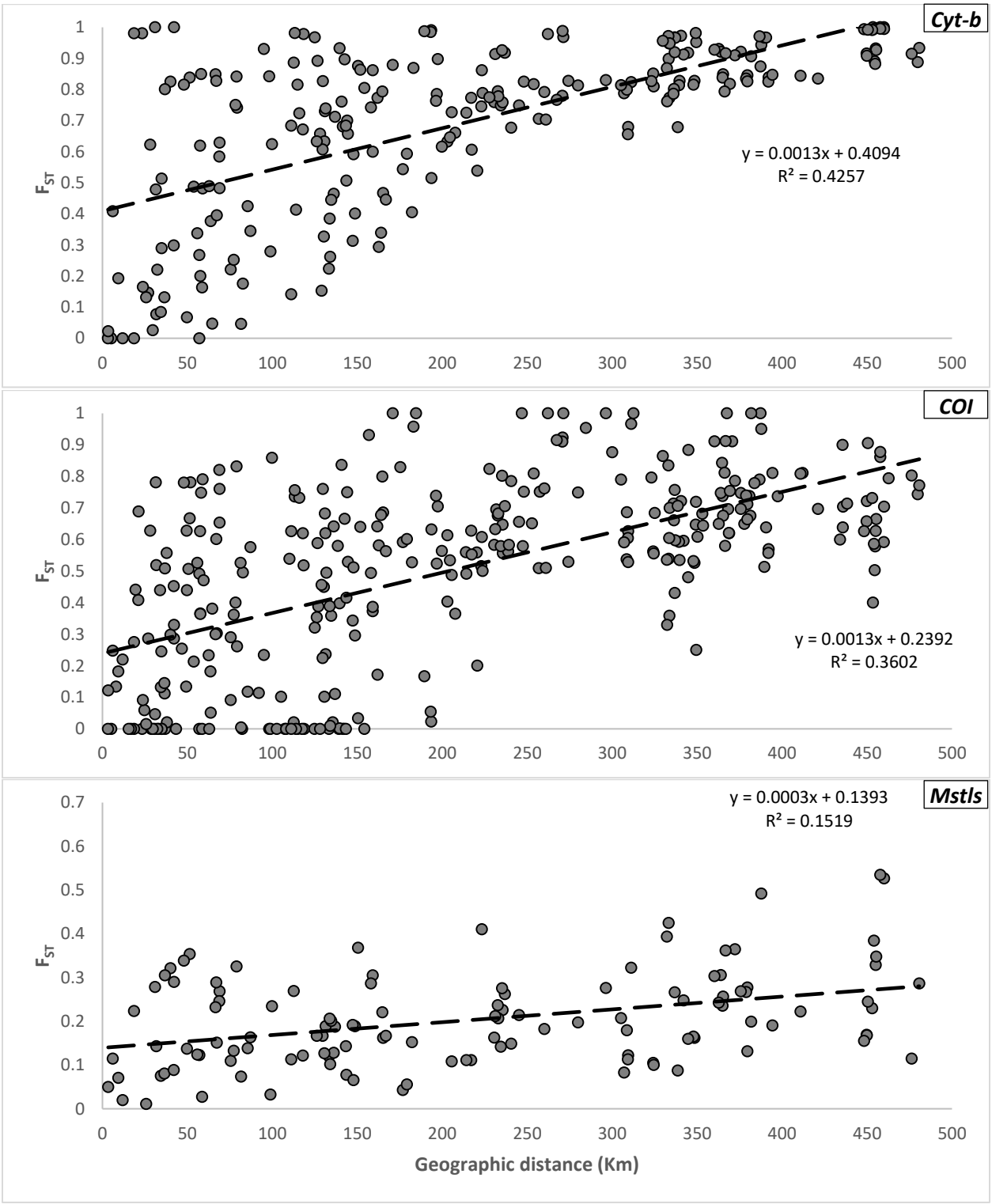
Appendix S13. Haplotype frequency of population of *Rhoadsia* in western Ecuador based on mitochondrial gene cytochrome oxidase I.

Site	N	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12	H13	H14	H15	H16	H17	H18
Su	2	-	-	2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
At	7	-	-	6	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-
E2B	7	-	-	7	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
E1	6	-	-	6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
E9	8	-	-	2	-	-	-	-	-	-	-	-	4	-	2	-	-	-	-
E3	10	-	-	7	-	-	-	-	-	-	-	-	2	1	-	-	-	-	-
E4	8	-	-	8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
E5	8	-	-	8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
E6	15	-	-	14	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-
E7	8	-	-	7	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-
E8	8	5	-	2	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-
G6	2	1	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
G7	4	3	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
G1	16	11	1	1	3	-	-	-	-	-	-	-	-	-	-	-	-	-	-
G2	15	13	-	-	-	-	-	2	-	-	-	-	-	-	-	-	-	-	-
G5	2	-	-	-	-	-	-	2	-	-	-	-	-	-	-	-	-	-	-
G3	8	-	-	-	-	-	-	5	-	-	-	-	-	-	-	2	1	-	-
G4	7	-	-	-	-	-	-	4	-	-	-	-	-	-	-	-	3	-	-
Ta	2	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	1
SP	6	-	-	-	-	-	-	6	-	-	-	-	-	-	-	-	-	-	-
J1	9	-	-	-	1	2	4	-	-	2	-	-	-	-	-	-	-	-	-
J2	8	-	-	-	4	3	-	-	-	1	-	-	-	-	-	-	-	-	-
J3	6	-	-	-	2	1	-	2	-	-	1	-	-	-	-	-	-	-	-
J4	8	-	-	-	-	-	3	1	-	-	3	1	-	-	-	-	-	-	-
J5	10	-	-	-	-	-	-	5	-	-	5	-	-	-	-	-	-	-	-
SR	16	-	-	-	4	11	1	-	-	-	-	-	-	-	-	-	-	-	-
TOTAL	206	33	1	72	16	17	9	27	1	3	9	1	6	1	2	2	4	1	1
Freq.		0.160	0.005	0.350	0.078	0.083	0.044	0.131	0.005	0.015	0.044	0.005	0.029	0.005	0.010	0.010	0.019	0.005	0.005

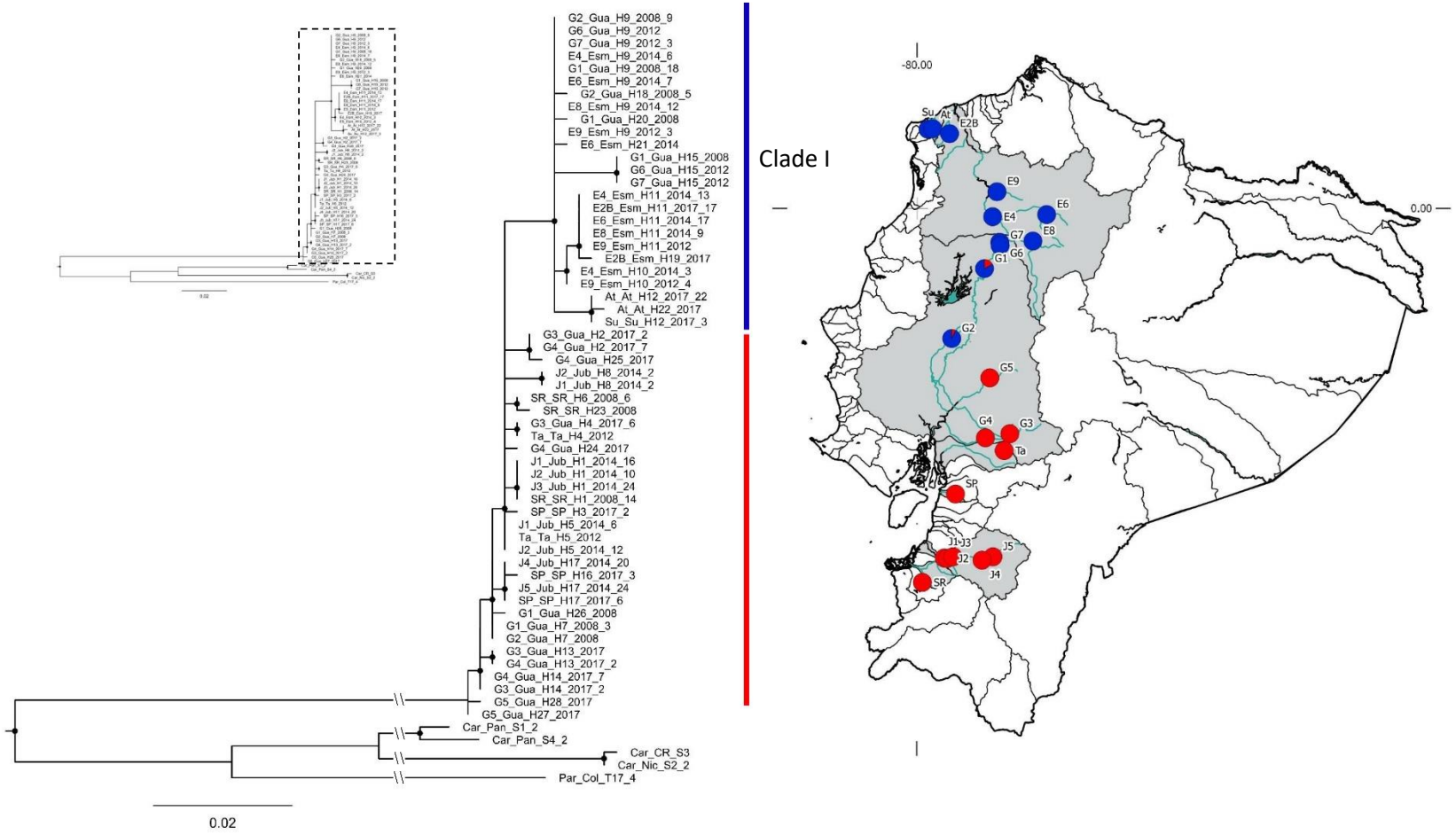
Appendix S14. Number of allele per site per locus. *AVE (average of allele per site) SE (Standard error).*

Drainage	Atacames	Esmeraldas				Guayas				San Pablo	Jubones					Santa Rosa
Locus	At	E2B	E4	E6	E8	G1	G2	G3	G4	SP	J1	J2	J3	J4	J5	SR
RA-09	2	7	3	3	4	4	4	5	6	4	4	4	4	3	3	7
RA-26	2	7	7	5	4	11	6	3	5	1	5	5	6	2	1	10
RA-20	4	12	9	6	5	14	6	9	11	7	4	6	4	4	5	8
RA-27	3	13	12	7	8	15	6	8	14	8	15	15	8	6	7	12
RA-30	2	2	2	3	3	4	3	5	6	5	4	5	4	3	3	5
RA-21	4	3	8	4	4	10	6	6	8	5	5	7	7	4	4	6
RA-16	2	5	6	3	5	8	5	7	7	4	4	7	7	3	3	4
RA-13	9	15	25	6	9	22	13	10	19	7	6	5	6	5	6	12
RA-22	1	5	3	3	4	5	3	3	5	5	4	4	3	5	3	4
RA-10	2	14	15	7	10	15	10	7	9	9	7	7	8	5	5	13
RA-08	2	3	4	3	3	4	3	5	5	2	5	1	4	3	2	4
RA-06	3	5	4	3	4	6	6	7	8	4	5	4	3	3	2	7
Total	36	91	98	53	63	118	71	75	103	61	68	70	64	46	44	92
AVE	3	7.6	8.2	4.4	5.3	9.8	5.9	6.3	8.6	5.1	5.7	5.8	5.3	3.8	3.7	7.7
SE	0.6	1.3	1.9	0.5	0.7	1.7	0.9	0.6	1.2	0.7	0.9	1	0.5	0.3	0.5	1

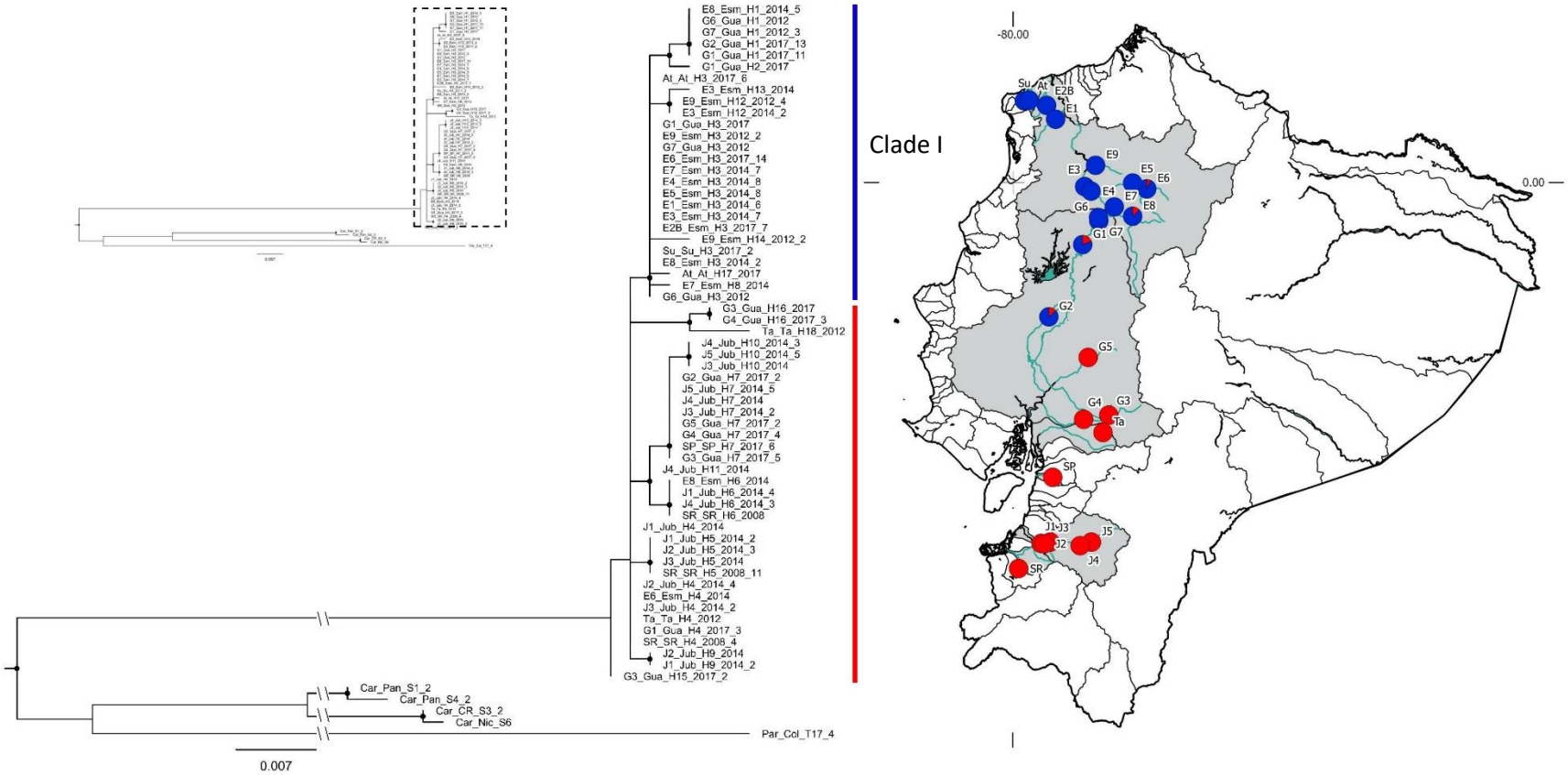
Appendix S15. Relationship between pairwise FST and geographic distance of populations of *Rhoadsia* in Ecuador according to the cytochrome b (Top), cytochrome oxidase I (Center), and microsatellites (Bottom) data. *Mantel test p value was <0.0001 for all the genetic markers.*



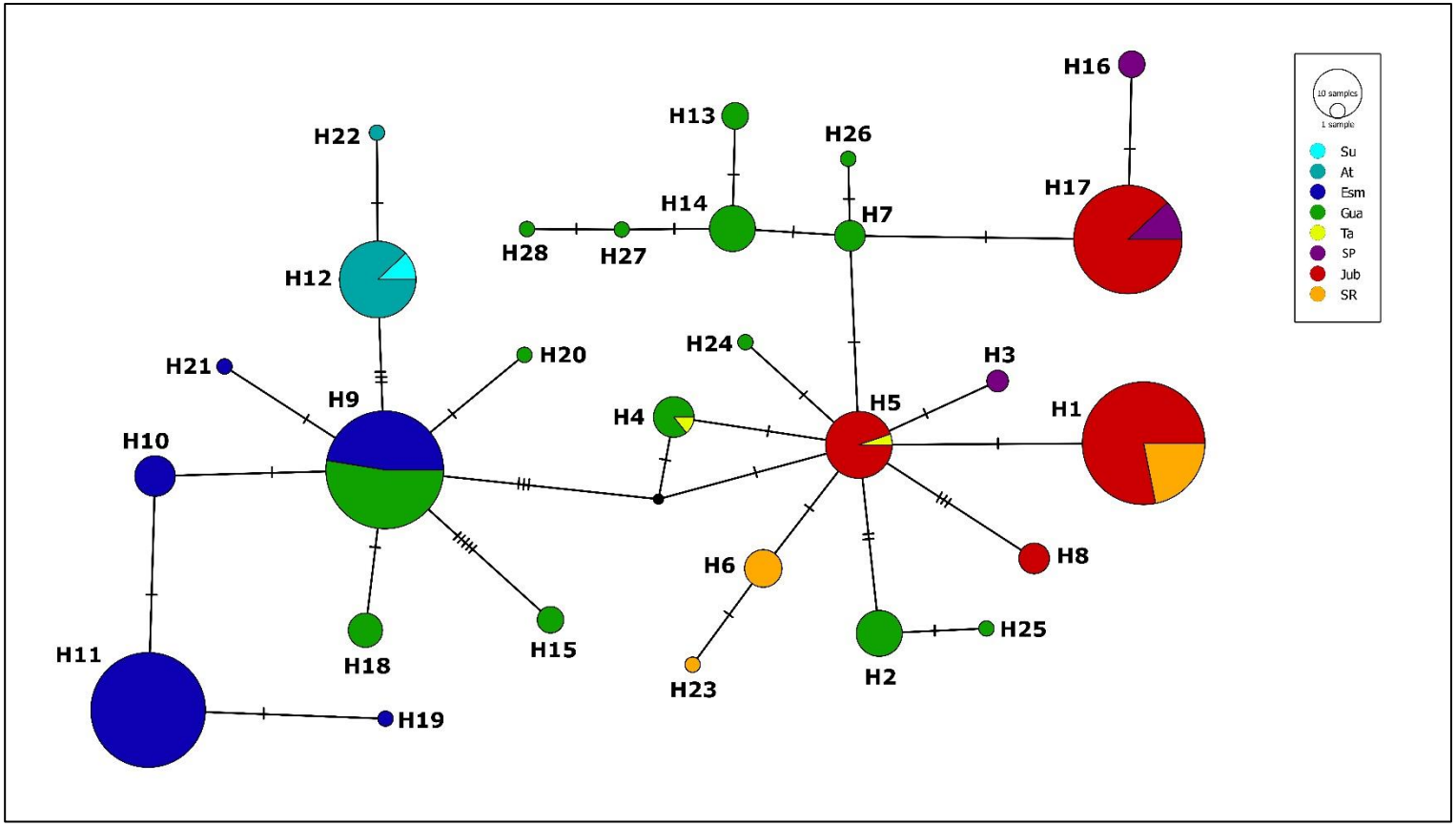
Appendix S16. Maximum likelihood phylogenetic tree based on the mitochondrial gene Cytochrome b. Branch *support values are based on the approximate likelihood ratio test, Shimodaira-Hasegawa-Like (aLRT SH-Like). Values greater or equal than 0.7 are represented with black dots at their node. Clade I (blue) comprised populations from the northern region while the rest (red) is represented by populations from the southern region.*



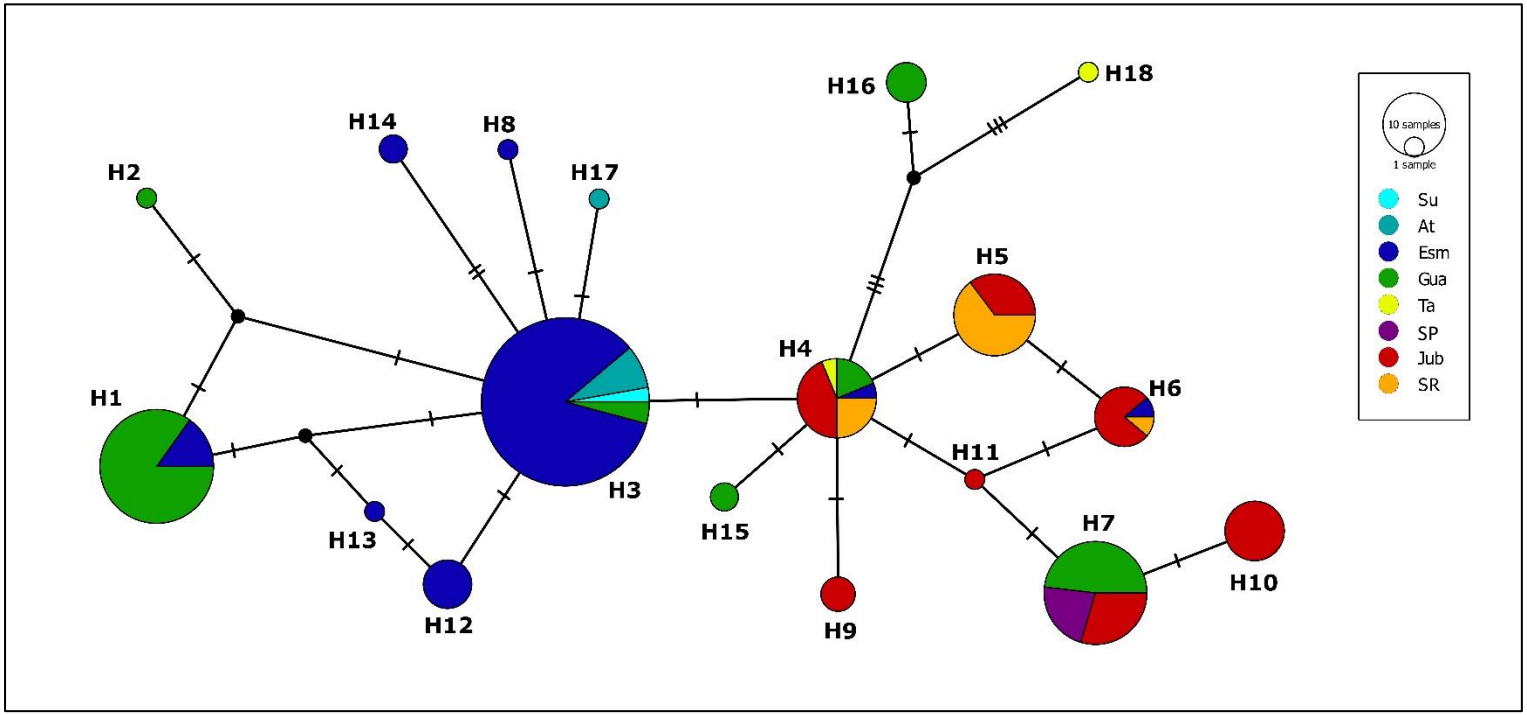
Appendix S17. Maximum likelihood phylogenetic tree based on the mitochondrial gene cytochrome oxidase I. Branch *support values are based on the approximate likelihood ratio test, Shimodaira-Hasegawa-Like (aLRT SH-Like). Values greater or equal than 0.7 are represented with black dots at their node. Clade I (blue) comprised populations from the northern region while the rest (red) is represented by populations from the southern region.*



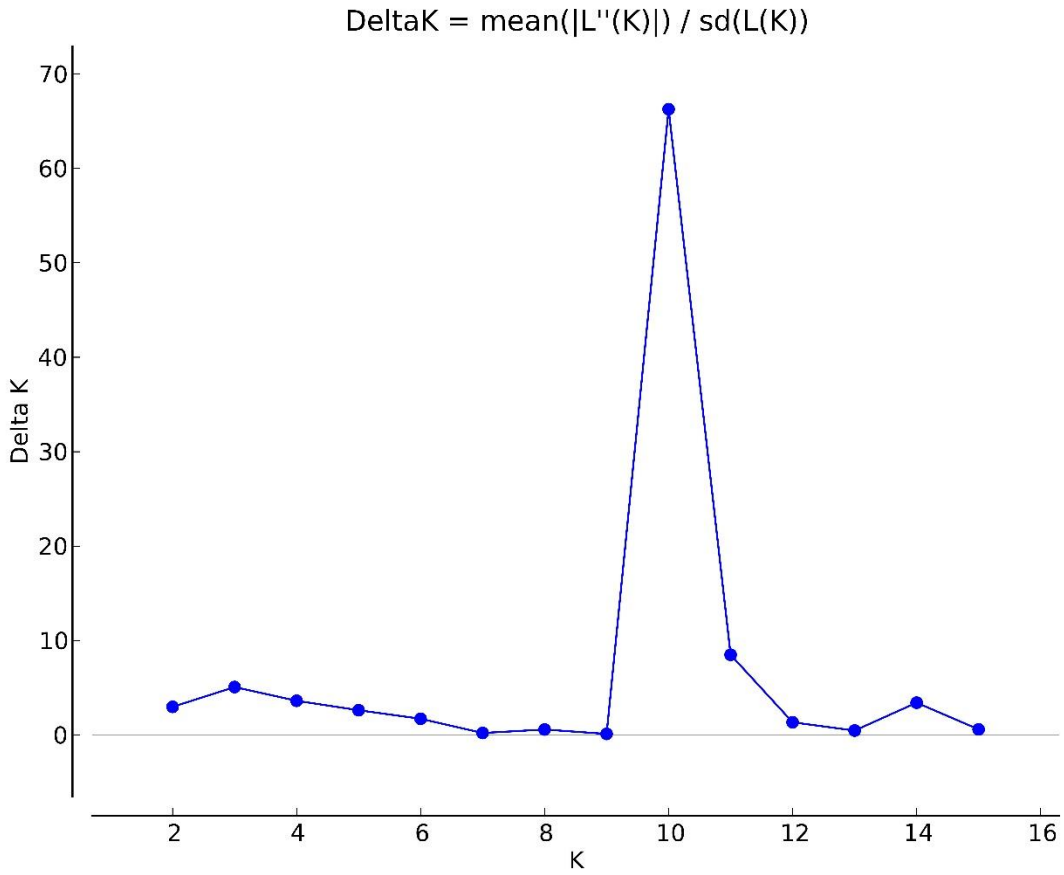
Appendix S18. Haplotype network based the mitochondrial gene cytochrome b per drainage. *Size of circle represent the haplotype frequency. The number of mutations between the haplotypes are represented by hatch marks. Drainages were represented by different colors (See legend).*



Appendix S19. Haplotype network based the mitochondrial gene cytochrome oxidase I per drainage. *Size of circle represent the haplotype frequency. The number of mutations between the haplotypes are represented by hatch marks. Drainages were represented by different colors (See legend).*



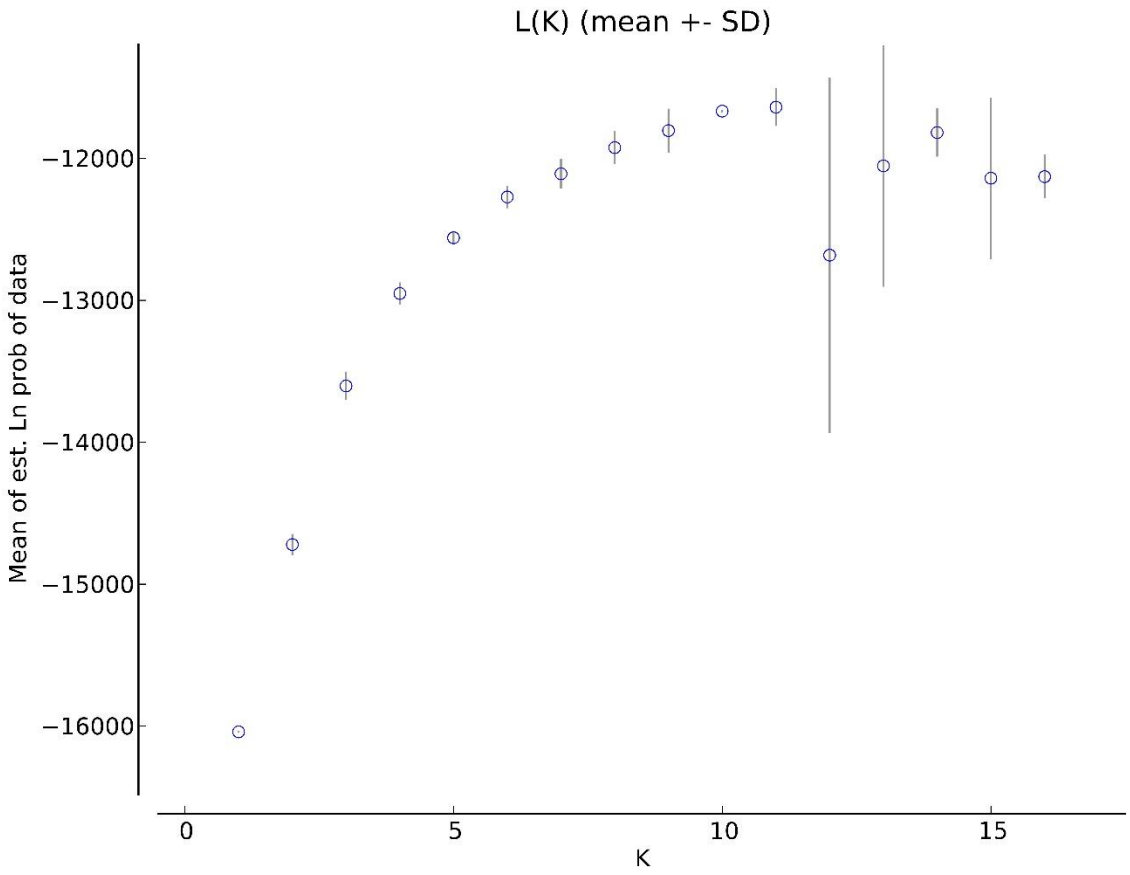
Appendix S20. Method used by Evanno et al. (2005) to determine the most likely number of cluster (K) inferred in STRUCTURE based on the microsatellite data. *This method follows the equation Delta K = (mean(|Ln"(K)|)/sd(Ln(K) where |Ln"(K)| is the absolute value of the second order rate of change of the mean likelihood distribution and sd(Ln(K) is the standard deviation of the mean natural log probability of the data. The highest value for Delta K represents the "true" K (K=10). See appendix S21.*



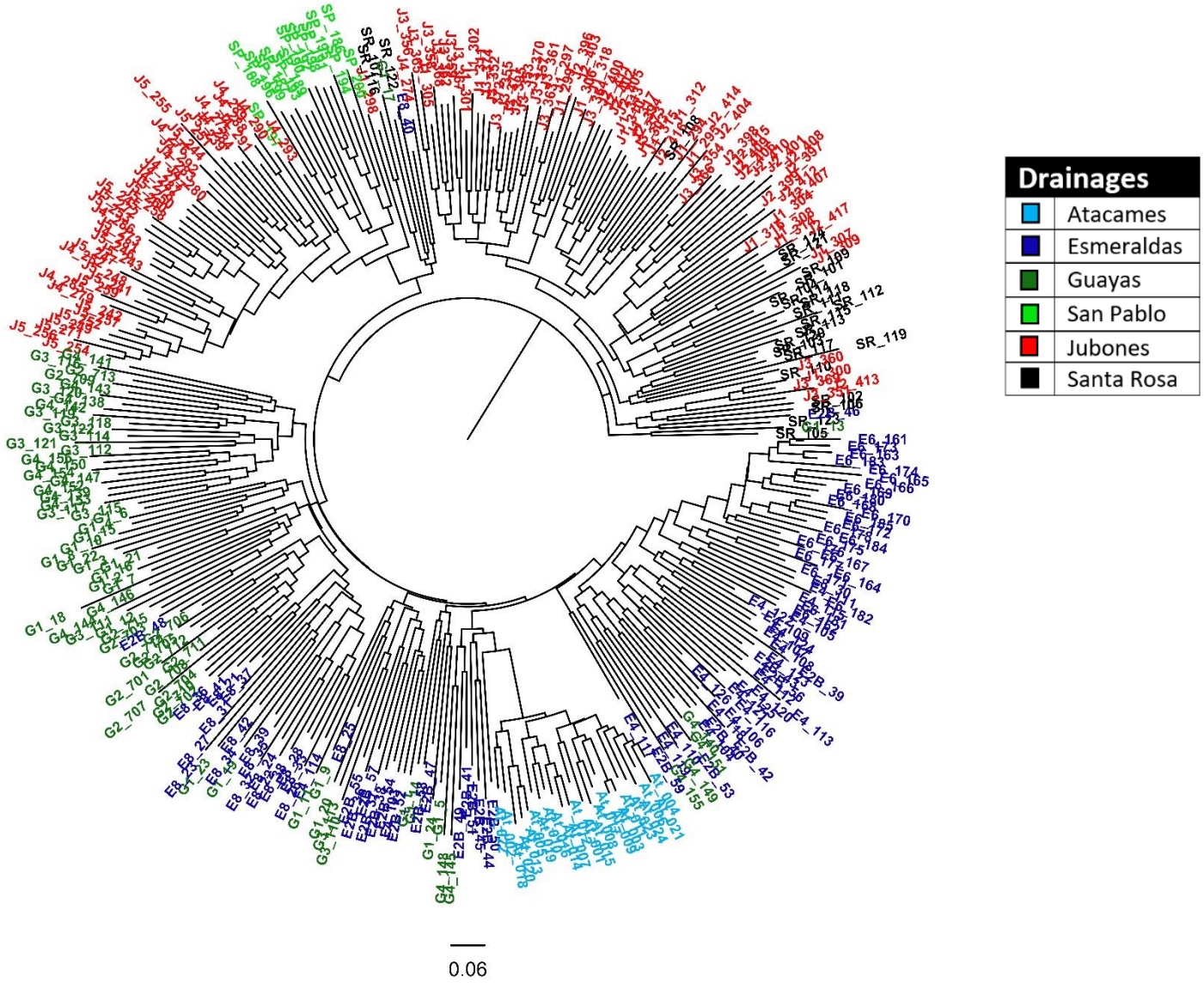
Appendix S21. Calculation of Delta K according to Evanno et al. (2005) method. *Value of the selected K is highlighted (K=10).*

K	Reps	Mean LnP(K)	Stdev LnP(K)	Ln'(K)	|Ln''(K)|	Delta K
1	20	-16040.995	0.333	NA	NA	NA
2	20	-14720.945	67.954	1320.050	202.520	2.980
3	20	-13603.415	91.826	1117.530	465.980	5.075
4	20	-12951.865	71.437	651.550	258.755	3.622
5	20	-12559.070	40.149	392.795	105.280	2.622
6	20	-12271.555	72.212	287.515	123.975	1.717
7	20	-12108.015	96.998	163.540	20.655	0.213
8	20	-11923.820	110.750	184.195	63.995	0.578
9	20	-11803.620	147.562	120.200	16.905	0.115
10	20	-11666.515	1.649	137.105	109.285	66.273
11	20	-11638.695	126.061	27.820	1070.970	8.496
12	20	-12681.845	1244.342	-1043.150	1672.510	1.344
13	20	-12052.485	844.982	629.360	394.355	0.467
14	20	-11817.480	162.871	235.005	556.915	3.419
15	20	-12139.390	562.906	-321.910	333	0.592
16	20	-12128.300	147.084	11.090	NA	NA

Appendix S22. Method used by Pritchard et al. (2000) to determine the most likely number of cluster (K) inferred in STRUCTURE based on the microsatellite data. *The "true" value for K is determined according to the highest mean Ln(K) which is followed by a plateau (K=10).*



Appendix S23. Neighbor joining of individuals based on microsatellites data using the Nei's DA genetic distance method (Nei et al., 1983) in the program Population 1.2.32.



Appendix S24. Proportion of membership of each cluster estimated in STRUCTURE (K=10).

	1	2	3	4	5	6	7	8	9	10
At	0.967	0.004	0.006	0.004	0.004	0.003	0.003	0.002	0.003	0.004
E2B	0.025	0.851	0.008	0.020	0.032	0.012	0.028	0.005	0.007	0.012
E4	0.025	0.739	0.036	0.022	0.122	0.010	0.013	0.009	0.015	0.008
E6	0.007	0.021	0.860	0.078	0.010	0.006	0.006	0.003	0.003	0.005
E8	0.013	0.013	0.007	0.914	0.012	0.016	0.006	0.006	0.005	0.008
G1	0.010	0.041	0.004	0.087	0.663	0.092	0.045	0.009	0.014	0.037
G2	0.005	0.009	0.020	0.016	0.049	0.841	0.027	0.009	0.009	0.015
G3	0.002	0.007	0.020	0.007	0.011	0.023	0.905	0.006	0.008	0.010
G4	0.005	0.053	0.006	0.010	0.040	0.064	0.780	0.022	0.010	0.011
SP	0.003	0.005	0.004	0.006	0.013	0.010	0.045	0.890	0.015	0.010
J1	0.004	0.005	0.005	0.007	0.007	0.005	0.006	0.004	0.938	0.018
J2	0.003	0.004	0.003	0.006	0.006	0.008	0.009	0.009	0.939	0.012
J3	0.003	0.005	0.003	0.005	0.009	0.007	0.007	0.010	0.935	0.015
J4	0.004	0.005	0.007	0.005	0.005	0.006	0.003	0.956	0.004	0.005
J5	0.002	0.003	0.004	0.003	0.003	0.003	0.003	0.975	0.002	0.002
SR	0.004	0.008	0.004	0.005	0.019	0.008	0.010	0.008	0.058	0.875