UNIVERSIDAD TECNOLÓGICA INDOAMÉRICA

DIRECCIÓN DE POSGRADO

MAESTRÍA EN BIODIVERSIDAD Y CAMBIO CLIMÁTICO

TEMA:

POPULATION GENOMICS OF THE POISON FROG *EPIPEDOBATES ANTHONYI* ACROSS DIFFERENT ELEVATIONS OF SOUTHERN ANDES IN ECUADOR TO UNDERSTAND AND CONSERVE CURRENT GENETIC DIVERGENCE PROCESSES

Trabajo de investigación previo a la obtención del título de Magister en Biodiversidad y Cambio Climático

Autor(a)

Gabriela Alexandra Maldonado Castro

Tutor(a)

PhD. Mónica Isabel Páez Vacas

QUITO – ECUADOR

2022

AUTORIZACIÓN POR PARTE DEL AUTOR PARA LA CONSULTA, REPRODUCCIÓN PARCIAL O TOTAL, Y PUBLICACIÓN ELECTRÓNICA DEL TRABAJO DE TÍTULACIÓN

Yo, Gabriela Alexandra Maldonado Castro, declaro ser autor del Trabajo de Investigación con el nombre "Population genomics of the poison frog *Epipedobates anthonyi* across different elevations of southern Andes in Ecuador to understand and conserve current genetic divergence processes", como requisito para optar al grado de Magister en Biodiversidad y Cambio Climático y autorizo al Sistema de Bibliotecas de la Universidad Tecnológica Indoamérica, para que con fines netamente académicos divulgue esta obra a través del Repositorio Digital Institucional (RDI-UTI).

Los usuarios del RDI-UTI podrán consultar el contenido de este trabajo en las redes de información del país y del exterior, con las cuales la Universidad tenga convenios. La Universidad Tecnológica Indoamérica no se hace responsable por el plagio o copia del contenido parcial o total de este trabajo.

Del mismo modo, acepto que los Derechos de Autor, Morales y Patrimoniales, sobre esta obra, serán compartidos entre mi persona y la Universidad Tecnológica Indoamérica, y que no tramitaré la publicación de esta obra en ningún otro medio, sin autorización expresa de la misma. En caso de que exista el potencial de generación de beneficios económicos o patentes, producto de este trabajo, acepto que se deberán firmar convenios específicos adicionales, donde se acuerden los términos de adjudicación de dichos beneficios.

Para constancia de esta autorización, en la ciudad de Quito, a los 28 días del mes de octubre de 2022, firmo conforme:

Autor: Gabriela Alexandra Maldonado Castro

Firmado electrónicamente por:
GABRIELA ALEXANDRA
MALDONADO CASTRO

Firma: ……

Número de Cédula: 1717774291

Dirección: Pichincha, Quito, Chillogallo.

Correo Electrónico: gabylumaldonado@gmail.com

Teléfono: (02) 2634930 / 0995678852

APROBACIÓN DEL TUTOR

En mi calidad de Tutor del Trabajo de Titulación "POPULATION GENOMICS OF THE POISON FROG *EPIPEDOBATES ANTHONYI* ACROSS DIFFERENT ELEVATIONS OF SOUTHERN ANDES IN ECUADOR TO UNDERSTAND AND CONSERVE CURRENT GENETIC DIVERGENCE PROCESSES" presentado por Gabriela Alexandra Maldonado Castro, para optar por el Título Magister en Biodiversidad y Cambio Climático,

CERTIFICO

Que dicho trabajo de investigación ha sido revisado en todas sus partes y considero que reúne los requisitos y méritos suficientes para ser sometido a la presentación pública y evaluación por parte del Tribunal Examinador que se designe.

Quito, 24 de noviembre del 2022

MONICA ISABEL PAEZ VACAS

Digitally signed by MONICA ISABEL PAEZ VACAS
Date: 2022.11.24 22:15:10 -05'00'

.....

PhD. Mónica Isabel Páez Vacas

DECLARACIÓN DE AUTENTICIDAD

Quien suscribe, declaro que los contenidos y los resultados obtenidos en el presente trabajo de investigación, como requerimiento previo para la obtención del Título de Magister en Biodiversidad y Cambio Climático, son absolutamente originales, auténticos y personales y de exclusiva responsabilidad legal y académica del autor.

Quito, 24 de noviembre 2022

Firmado electrónicamente por:
GABRIELA ALEXANDRA
MALDONADO CASTRO

……

Gabriela Alexandra Maldonado Castro
171777429-1

APROBACIÓN TRIBUNAL

El trabajo de Titulación, ha sido revisado, aprobado y autorizada su impresión y empastado, sobre el Tema: POPULATION GENOMICS OF THE POISON FROG *EPIPEDOBATES ANTHONYI* ACROSS DIFFERENT ELEVATIONS OF SOUTHERN ANDES IN ECUADOR TO UNDERSTAND AND CONSERVE CURRENT GENETIC DIVERGENCE PROCESSES, previo a la obtención del Título de Magister en Biodiversidad y Cambio Climático, reúne los requisitos de fondo y forma para que el estudiante pueda presentarse a la sustentación del trabajo de titulación.

Quito, 24 de noviembre de 2022

NORA HELENA OLEAS GALLO
Firmado digitalmente por NORA HELENA OLEAS GALLO
Fecha: 2022.11.25 10:27:46 -05'00'

……..

PhD. Nora Helena Oleas Gallo

PRESIDENTE DEL TRIBUNAL

Firmado electrónicamente por:
CHRISTIAN DAVID SALAZAR VALENZUELA

……

PhD. Christian David Salazar Valenzuela

VOCAL

MONICA ISABEL PAEZ VACAS
Digitally signed by MONICA ISABEL PAEZ VACAS
Date: 2022.11.24 22:15:40 -05'00'

……

PhD. Mónica Isabel Páez Vacas

DIRECTORA

DEDICATORIA

¡A todas las mujeres que dedican su vida a descubrir, entender, y maravillarse por los procesos que ocurren en nuestro planeta Tierra a través de la ciencia!

"*En la vida no existe nada que temer, solo cosas que comprender*"
Marie Curie

AGRADECIMIENTO

Mi eterna gratitud a Mónica Páez-Vacas, quien no solo fue un pilar en el desarrollo de esta investigación, sino que fue quien nunca dejó de creer en mí, incluso cuando yo lo dejé de hacer. Mil gracias por convertirme en una mejor científica, gracias por tu paciencia y regaños, gracias por permitirme creer nuevamente que la ciencia es algo que se puede disfrutar sin el ego y con la complicidad y apoyo del otro.

Gracias infinitas a Patricia Salerno por el apoyo en el desarrollo de metodología, gracias por enseñarme sin egoísmo y con paciencia, gracias por compartirme las herramientas para mi futuro profesional.

Mil gracias a todas las mujeres que han cuidado de mi salud mental, que han sido cómplices de este camino llamado vida y quienes siempre me han regalado palabras reconfortantes Naty, Vero, Jhaeli, Pao, Mica, Karlita, Enmy, Amalia, Becha.

¡Gracias infinitas!

INDEX

CHAPTER I

CHAPTER II

CHAPTER III

CHAPTER IV

CHAPTER V

TABLE INDEX

FIGURE INDEX

UNIVERSIDAD TECNOLÓGICA INDOAMÉRICA
DIRECCIÓN DE POSGRADO
MAESTRÍA DE BIODIVERSIDAD Y CAMBIO CLIMÁTICO

TEMA: POPULATION GENOMICS OF THE POISON FROG *EPIPEDOBATES ANTHONYI* ACROSS DIFFERENT ELEVATIONS OF SOUTHERN ANDES IN ECUADOR TO UNDERSTAND AND CONSERVE CURRENT GENETIC DIVERGENCE PROCESSES.

Autora: Gabriela Alexandra Maldonado Castro
Tutora: Mónica Isabel Páez Vacas, PhD.

RESUMEN EJECUTIVO

La genómica proporciona nuevas herramientas para responder preguntas sobre especiación, conservación, evolución fenotípica, entre otras. La secuenciación de nueva generación (NGS) ha cambiado el estudio de la variación genética y la ha convertido en una herramienta accesible para el estudio en poblaciones naturales. El presente estudio tiene como objetivo investigar la variación genética a nivel del genoma de doce poblaciones de *Epipedobates anthonyi* (166 individuos) distribuidas a lo largo de cuatro transectos elevacionales (0-1800 msnm) y que presentan alta variación fenotípica tanto en color y tamaño. Luego de una serie de filtros a los datos iniciales obtuvimos dos matrices, una con 2706 Polimorfismos de Nucleótidos Simples (SNPs (matriz m3M5n5) y otra con 1706 SNPs (matriz m4M5n5), también identificamos 69 y 48 outlier-loci (putativamente bajo selección), respectivamente. Análisis de Componentes Principales (PCA) y Análisis Discriminante de Componentes Principales (DAPC) se aplicaron en las matrices y la estructura poblacional de *E. anthonyi* mostró que las poblaciones de alitudes medias y altas se agrupan, mientras que las poblaciones de altitudes bajas están más dispersas en el espacio multivariante mostrando también mayor variación genética. Encontramos poca divergencia en todo

el genoma en los loci neutrales y putativamente bajo selección (outlier) (F_{ST}). Esto podría indicar que estas poblaciones se encuentran en las primeras etapas de divergencia poblacional, y que los mecanismos evolutivos que operan a nivel del genoma (flujo genético, deriva genética) podrían estar explicando los patrones de variación genética poblacional, o que, existen solo algunos loci detrás de la variación fenotípica observada y que no pudieron ser detectados con nuestros análisis. Finalmente, nuestros resultados indican que la altitud puede ser un factor que modifique la estructura genética entre las poblaciones de *E. anthonyi*. Estos resultados resaltan importancia de comprender la trayectoria evolutiva en las especies de los Andes tropicales e implica que las estrategias de conservación deben presentar un enfoque holístico que incluya no sólo la protección del medio ambiente sino también la protección de la información genética en estas especies.

Palabras clave: anfibios, DAPC, genómica de poblaciones, IBD, loci, outlier loci, PCA, SNPs.

UNIVERSIDAD TECNOLÓGICA INDOAMÉRICA

DIRECCIÓN DE POSGRADO

MAESTRÍA DE BIODIVERSIDAD Y CAMBIO CLIMÁTICO

THEME: POPULATION GENOMICS OF THE POISON FROG *EPIPEDOBATES ANTHONYI* ACROSS DIFFERENT ELEVATIONS OF SOUTHERN ANDES IN ECUADOR TO UNDERSTAND AND CONSERVE CURRENT GENETIC DIVERGENCE PROCESSES

AUTHOR: Gabriela Alexandra Maldonado Castro

TUTOR: Mónica Isabel Páez Vacas, PhD.

ABSTRACT

Genomics provides new tools to answer questions about speciation, conservation, and phenotypic evolution, among others. Next-generation sequencing (NGS) has changed the study of genetic variation and made it an accessible tool for studying natural populations. The objective of this study is to investigate the genetic variation at the genome level of twelve populations of *Epipedobates anthonyi* (166 individuals) distributed along four elevational transects (0 – 1800 m asl) that present phenotypic variation in both color and size, using restriction site-associated DNA sequencing. After a series of filters to the initial genomic data, we obtained two matrices, one with 2706 Single Nucleotide Polymorphisms (SNPs) (m3M5n5 matrix) and another with 1706 SNPs (m4M5n5 matrix), we also identified 69 and 48 outlier-loci of (putatively under selection) respectively. Principal Component Analysis (PCA) and Principal Component Discriminant Analysis (DAPC) applied on the matrices show that populations of mid and high altitudes cluster, while populations of low altitudes are more dispersed in the multivariate space showing also greater genetic variation. We found little genome-wide divergence at neutral and putatively outlier loci (F_{ST}). This could indicate that these populations are in the early stages of divergence, and that

evolutionary mechanisms operating at the genome level (gene flow, genetic drift) could be explained the patterns of population genetic variation, or that there are just a few loci behind the observed phenotypic variation and that could not be detected with our analyses. Finally, our results suggest that altitude is a factor that shapes genetic structure among populations of *E. anthony*i. These findings highlight the importance of understanding the evolutionary trajectory of species from the tropical Andes and imply that conservation strategies must include not only environmental protection but also the protection of genetic information found in these species.

Keywords: amphibians, DAPC, genomics population, IBD, loci, outlier loci, PCA, SNPs.

CHAPTER I

THEME: Population genomics of the poison frog *Epipedobates anthonyi* across different elevations of southern Andes in Ecuador to understand and conserve current genetic divergence processes.

The tropical Andes region is one of the most biodiverse in the world, due to a high level of endemism and species richness in both plants and vertebrates. Sadly, also presents a progressive high rate of habitat loss (Myers et al., 2000, Smith et al., 2014). These characteristics define the region as a global biodiversity hotspot (Brummitt & Lughadha, 2003; Chazot et al., 2016; Hutter et al., 2017; Myers et al., 2000). However, the evolutionary mechanisms responsible for the region richness remain poorly understood (Cadena et al., 2012; Chazot et al., 2016; Hutter et al., 2017; Myers et al., 2000; Orme et al., 2005; Salazar-Valenzuela et al., 2019). In this region, Ecuador is one of the tropical countries with the highest biodiversity per unit area reflected in the species number for several groups such as plants, mammals, birds, reptiles, and amphibians (Ron et al., 2019; Iturralde-Pólit et al., 2017). Ecuador has 662 amphibian species, positioning it worldwide as the third country in amphibian richness (Ron et al., 2019).

Amphibian richness in the tropical Andes has been explained as a combination of high diversification rates and the timing of clade colonization (Hutter et al., 2017), as well as the higher speciation rates at higher altitude zones because of site topography and climatic variation (Guarnizo & Canatella, 2013). Dispersal and vicariance events that led to the differentiation of populations have remained opaque, but also Andean orogeny can promote diversification by providing novel and heterogeneous environments at different elevations, which in turn could promote local adaptation and ecological divergence (Ghalambor et al., 2006; Mayr, 1970 Schluter, 2000). However, this hypothesis can underestimate alternative mechanisms of speciation such as the ecological adaptation across complex environments (Sheu et al., 2020; Parini et al.,

2020). Recent research has concluded that besides geographic isolation, ecological traits in amphibian populations across different elevations could be also drivers of diversification (Cadena et al., 2011, Funk, 2005; Guarnizo & Canatella, 2013; Van Buskirk & Jansen van Rensburg, 2020). Ecological differences between populations can cause adaptive divergence in populations, which in turn could lead to ecological speciation (Schluter, 2001). Understanding how evolutionary forces can drive divergence in populations with phenotypic differences (clinal populations) across elevational gradients could explain the origin of the high biodiversity present in the region (Bolnick & Fitzpatrick, 2007; Richards et al., 2019). Understanding these processes is one of the pillars in the field of genetics, evolutionary ecology, and conservation genetics (Allendorf & Luikart, 2009; Catchen et al., 2013a).

Population genetics allows us to explore the mechanisms and evolutionary forces by studying the different levels of genetic variation within and between populations. One of the most popular molecular techniques is the use of microsatellites (Allendorf & Luikart, 2009). However, the emergence of genomics in the last decade, including the acquisition of more genetic data and the development of bioinformatics tools, has impacted the study of populations (Allendorf & Luikart, 2009; Schlötterer, 2004; Shaffer et al., 2015). Next-generation sequencing provides with high-quality data of millions of small DNA fragments at low cost, and bioinformatic tools to analyze them have changed the concept of population genetics (Rivera-Colón & Catchen, 2016).
The population genomics term has originated due to all the pointed approaches and mainly for processing bigger amounts of data to the genome-wide level. Also, it provides marker sets between 100s to 10000s regions with no requirement of prior genomic data at low cost through next-generation sequencing methods (Peterson et al., 2012). The main objective of population genomics is to analyze the patterns of change in genetic sequences, emphasizing differences within natural populations (Allendorf, 2016; Hartl, 2006; Shaffer et al., 2015;). Genomic tools now allow genome-wide studies, and theoretical advances can help design research strategies that use not just genomics but also field experiments and allows the evaluation of the processes and the

genetics of local adaptation. These advances are also helpful to investigate non-model species, whose adaptation patterns may differ from those of traditional model species (Savolainen et al., 2013).

Population genomics refers to the use of genome-wide data at numerous of variable loci to infer some characteristics of populations as demographics, evolutionary processes, describing population differences, inferring patterns of gene flow, identifying neutral genetic diversity, obtaining thousands of single nucleotide polymorphisms (SNPs) at both neutral and coding sites (Allendorf et al., 2022). It also provides information on local adaptation, effective population size, inbreeding depression (Allendorf et al., 2010; Black et al., 2001; Eguiarte et al., 2013). Over the past years, genomics has become a tool to understand evolutionary processes at different levels, even finding the responsible genes for local adaptation. These methods can still be challenging because the genome is large, and most signatures of selection are subtle (Whitlock & Lottherhos, 2015). Additionaly, amphibian speciation research has incorporated other dimensions of fields such as genetic mechanisms and evolutionary forces (Wollenberg et al., 2019).

Single nucleotide polymorphisms (SNPs) are the most abundant type of polymorphism in the genome (Baird et al., 2008; Brumfield et al., 2003). It is estimated that there is one in every 200 – 500 bp (base pairs) in many wild animal populations (Brumfied et al., 2003). SNPs at coding regions of the genome could identify phenotypic effects and thus detect traits related to fitness (Allendorf & Luikart, 2009). This means that SNPs could be powerful markers for separating the effects of demographics and history from natural selection (Luikart et al., 2003). The study of adaptive traits in natural populations of amphibians has been scarce (Guo et al., 2016). However, genomic studies at the population level would allow us to identify current processes of genetic divergence in Ecuadorian amphibians.

One of the Ecuadorian amphibian species suitable for studying these processes is *Epipedobates anthonyi*. This species is distributed along the western slopes of the Andes in southern Ecuador and northern Peru, and it occurs over a wide altitudinal range from sea level to 1800 (masl). Morphological and acoustic variation had been described in this poison species frog (Coloma et al., 2018; Funk et al., 2009; Santos et al., 2014). Páez-Vacas et al. (2022) found relatively low genetic divergence at nine neutral microsatellite loci in thirty-five populations of *E. anthonyi* along four altitudinal gradients. Phenotypic and genetic divergence were both explained by landscape resistance between sites. Moreover, all these variations occur along different altitudes, thus, these characteristics make this species an excellent candidate to study microevolutionary processes in natural populations.

OBJECTIVES

General objective

- Estimate the relationship between the population genetic structure of *Epipedobates anthonyi* across elevations in southern Ecuador using genomic data (SNPs).

Specific objectives

- To obtain a filtered and processed SNPs matrix of *Epipedobates anthonyi* populations.
- To characterize genomic clusters of *E. anthonyi* populations.
- To examine signals of possible selection based on detected putatively SNPs.
- To determine association between the elevational gradients and genetic structure.

CHAPTER II

METHODS

Study species, study area, and sampling

Epipedobates anthonyi is a poison frog that is distributed along an altitudinal range between sea level and 1800 (m asl) on the western Andean slopes of southern Ecuador and northwestern Peru (Ron et al., 2019; Frost, 2021). This species is abundant, exhibiting morphological and acoustic variation; and low genetic divergence in neutral microsatellite loci (Santos et al., 2014; Páez-Vacas et al. 2022). We collected 166 samples of *E. anthonyi* from 2012 to 2014, all of them were distributed in the southern Andes of Ecuador (Table 1, Figure 1). We studied 12 sites located along four transects: three elevational transects encompassing the altitudinal range of the species (900 – 1712 m asl) and one transect over a coastal lowland region (13 – 65 m asl).

Table 1.

Epipedobates anthonyi localities. Twelve localities of *E. anthonyi* used in this study, each population has the ID code, sample size (n), altitude (m asl), transect, and province.

Population (ID_Code)	(n)	Transect	Locality	(masl)	Province
EA22	14	1	Jubones	230	Azuay
EA8	13	1	San Sebastian	776	Azuay
EA2	19	1	San Rafael	1546	Azuay
EA16	10	2	Quera	229	El Oro
EA6	14	2	Chilla	979	El Oro
EA15	11	2	Chilla	1666	El Oro
EA14	14	3	Saracay	276	El Oro
EA10	13	3	Moromoro	946	El Oro
EA11	27	3	Paccha	1712	El Oro
EA31	10	lowlands	El Progreso	51	El Oro
EA32	10	lowlands	El Retiro	13	El Oro
EA35	11	lowlands	San Vicente	65	El Oro

Total	**166**

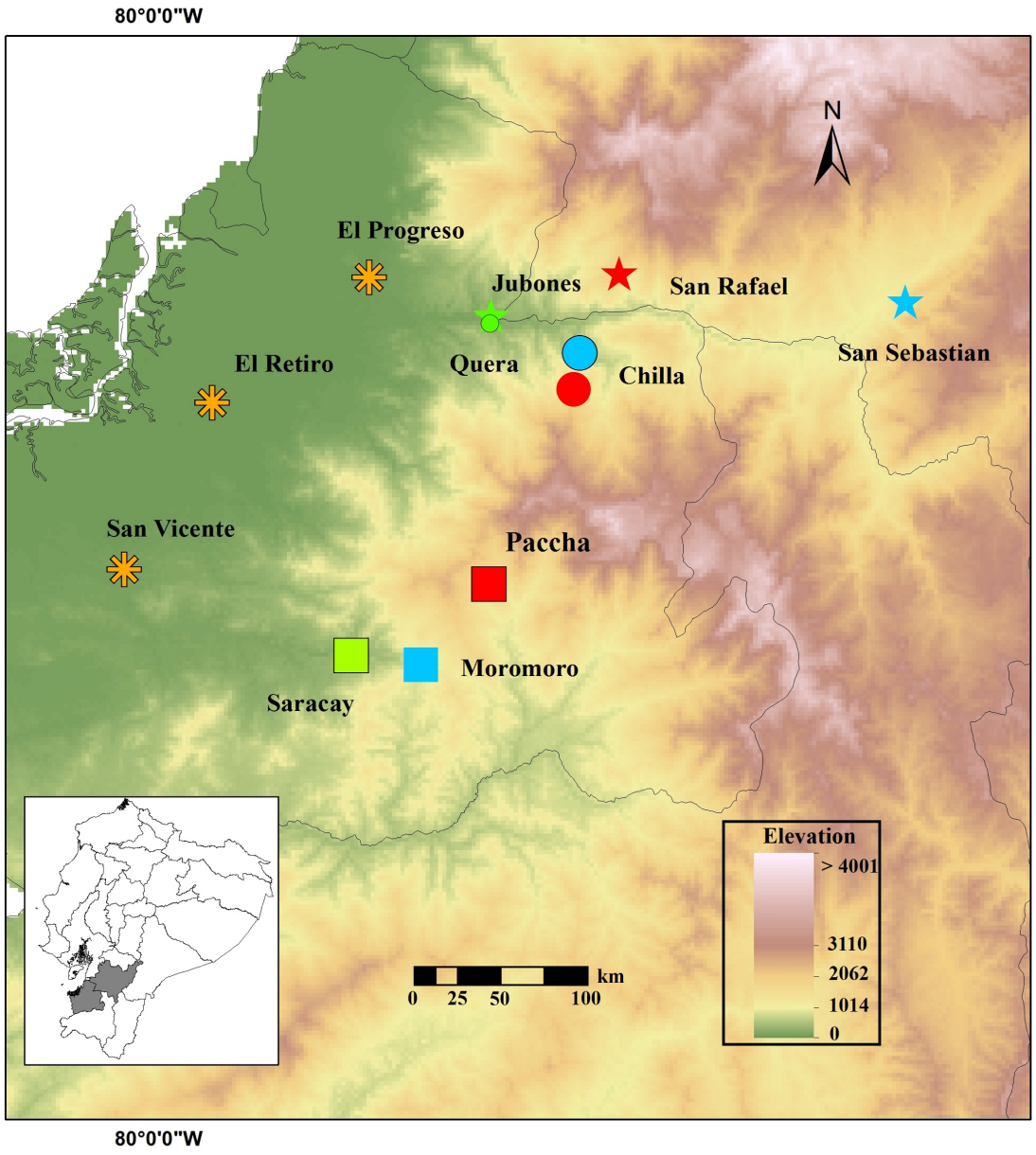


Figure 1. Locations of twelve populations of the species *Epipedobates anthonyi* studied. Populations from the same transect are represented by the same symbol (stars belong to transect one, circles to transect two, and squares to transect three). The symbol colors correspond to different altitudes, red represents populations from high altitudes (1546 – 1712 m asl), blue from medium altitudes (776 – 979 m asl), green from low altitudes (230, 276 m asl) and orange from the lowlands (13 – 65 m asl).

Tissues for genomic analyses (liver and muscle) were stored in 96% ethanol. Every specimen was fixed and preserved according to the standard museum collection protocols and deposited in the scientific collection of "Centro Jambatu de Investigación y Conservación de Anfibios". All collections were approved by the Ministerio de Agua y Ambiente del Ecuador (Permits N° 003-11 IC-FAU-DNB/MA and N° 001-13 IC-FAU-DNB/MA).

ddRad Genomic Data Generation

To characterize SNPs genotypes for genomic analysis, we used ddRAD sequencing (Peterson et al. 2012). Genomic DNA was extracted from the muscle of each specimen using a Qiagen DNA Blood and Tissue Kit and quantified with a Qubit 2.0 fluorometer. We constructed genomic libraries, using the restriction enzymes *EcoRI* and *NlaIII*. Forward unique barcodes (5 bp) for identifying each individual/sample and Illumina adapters were ligated to the digested DNA fragments normalized on a library basis to at least 250 ng. Finally, three libraries containing 166 samples with two pools each were sequenced in an Illumina HiSeq machine at the Oregon Genomics Core Facility to generate 97 bp single end reads.

Processing Illumina data

Raw data sequences obtained by Illumina Oregon Genomics Core Facility were processed in the **Stacks** v2.59 pipeline (Catchen et al., 2013 b). Stacks is designed to curate and assemble large numbers of sequences from multiple samples. All 166 samples were processed together with the following specifications: first, to recover the individual samples from the Illumina libraries *process_radtags* process was performed. Sequences were demultiplexed, and restriction sites and barcodes were trimmed. This process extracts all the samples (individuals) from the libraries and evaluates the read quality of the samples based on Phred scores. As we did not have a reference genome, the Stacks pipeline used was *DENOVO_MAP.PL* analysis, which ensues five major stages (ustacks, cstacks, sstacks, tsv2bam, gstacks, and populations).

Bioinformatics ddRAD genotyping

Once all data was demultiplexed, we ran 19 different *DENOVO_MAP.PL* assemblies. These trials were variations of the following Stacks parameters: (-m) a minimum number of identical reads to create a stack, (-M) number of mismatches allowed between loci in a single individual, and (-n) the number of mismatches allowed between loci when building catalog (Table 2). These parameters affect the results of allele numbers, allele identification, and SNP error rates (Mastretta-Yanes et al., 2015; Paris et al., 2017).

Table 2.

Epipedobates anthonyi demultiplexed data trials. Trials for demultiplexed data of *Epipedobates anthonyi* showing the parameters' values used of Stacks (-m, -M, -n). All trials started with the same initial sample (ni). Some samples were eliminated from *DENOVO_MAP.PL* analysis due to low coverage (equal or less than 4x) and every trial had a different final sample (nf). a* is the trial eliminated, e** and k** are the best *DENOVO_MAP.PL* permutations chosen.

Permutation	(-m)	(-M)	(-n)	(ni)	(nf)	Removed Samples
a*	3	2	3	166	166	0
b	3	3	2	166	166	0
c	3	3	3	166	166	0
d	3	5	4	166	166	0
e**	3	5	5	166	166	0
f	3	7	6	166	166	0
g	3	7	7	166	166	0
h	4	3	2	166	163	EA8_003, EA14_014, EA14_015
i	4	3	3	166	163	EA8_003, EA14_014, EA14_015
j	4	5	4	166	163	EA8_003, EA14_014, EA14_015
k**	4	5	5	166	162	EA8_003, EA6_010, EA14_014, EA14_015
l	4	7	6	166	163	EA8_003, EA14_014, EA14_015
m	4	7	7	166	163	EA8_003, EA14_014, EA14_015
n	5	3	2	166	158	EA8_003, EA14_014, EA14_015 EA14_013, EA6_002, EA6_010 EA6_019, EA8_010

o	5	3	3	166	158	EA8_003, EA14_014, EA14_015 EA14_013, EA6_002, EA6_010 EA6_019, EA8_010
p	5	5	4	166	158	EA8_003, EA14_014, EA14_015 EA14_013, EA6_002, EA6_010 EA6_019, EA8_010
q	5	5	5	166	158	EA8_003, EA14_014, EA14_015 EA14_013, EA6_002, EA6_010 EA6_019, EA8_010
r	5	7	6	166	158	EA8_003, EA14_014, EA14_015 EA14_013, EA6_002, EA6_010 EA6_019, EA8_010
s	5	7	7	166	158	EA8_003, EA14_014, EA14_015 EA14_013, EA6_002, EA6_010 EA6_019, EA8_010

The main *DENOVO_MAP.PL* Stack pipeline comprises the following stages: building loci in each sample (**ustacks**), creating the catalog of loci with all the samples (**cstacks**), and matching every single locus of each sample with the catalog created (**sstacks**). The final stages (**tsv2bam**) are to transpose all data from being stored per sample to be stored per-locus, and (**gstacks**) which compares every locus of each sample and identifies SNPs.

Before running *DENOVO_MAP.PL* for m=5, first, we decided to plot the coverage of populations EA14, EA6, and EA8 because in previous analyses (*DENOVO_MAP.PL* when m=4) their samples reported errors in the pipeline due to coverage of less than 4x. So, we had to identify samples with coverage of less than 5x and eliminate them for avoiding new errors when running *DENOVO_MAP.PL* for m=5 (Figure 2, Table 2). Over 19 DENOVO_MAP.PL trials, the **m3M2n3** combination was eliminated because it had the lowest coverage overall (Figure 3).

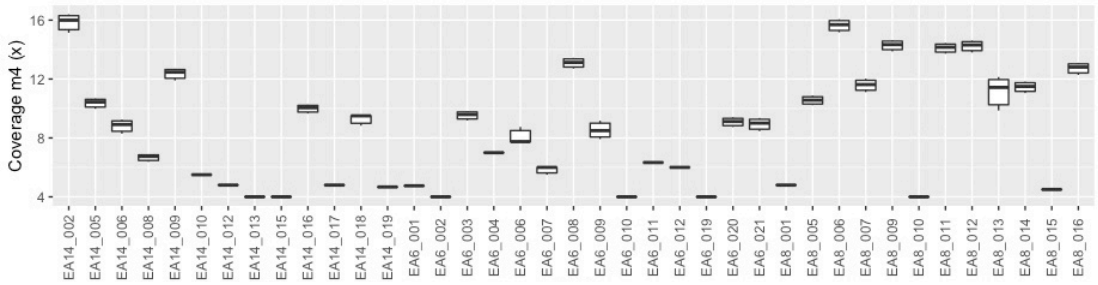


Figure 2. Coverage from *Epipedobates anthonyi* populations (EA14, EA6, EA8). Population samples of EA14, EA6, and EA8 with coverage less or equal to 4x were eliminated before running the DENOVO_MAP.PL when m=5.

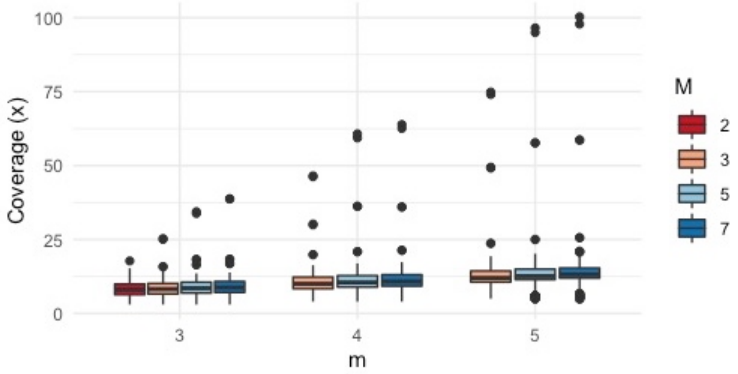


Figure 3. Mean coverage plot for the different combinations in *DENOVO_MAP.PL* with parameters -m and -M. Note the lowest coverage value when m3M2n2, and the highest coverage when m=5 accompanied by large variation.

To generate our first SNPs matrix, we used the **populations** Stacks program with **r = 0.8**, meaning that a locus must be present in 80% of individuals across populations, and **p = 1**, a locus must be present in a minimum one population (Paris et al., 2017). After running 18 trials, the best combinations were chosen to get a permutation of parameters that reduce the chances of over or under-merging loci (Trumbo et al., 2019). The analysis with the highest number of loci and variant sites were **m3M5n5** and **m4M5n5** (Figure 4). We exported the two best parameter permutations SNPs matrix in **populations** with the following requirements: **p = 1, r = 0.1,** and **--write-random-snp** (restricts data analyses to one random SNP per locus).

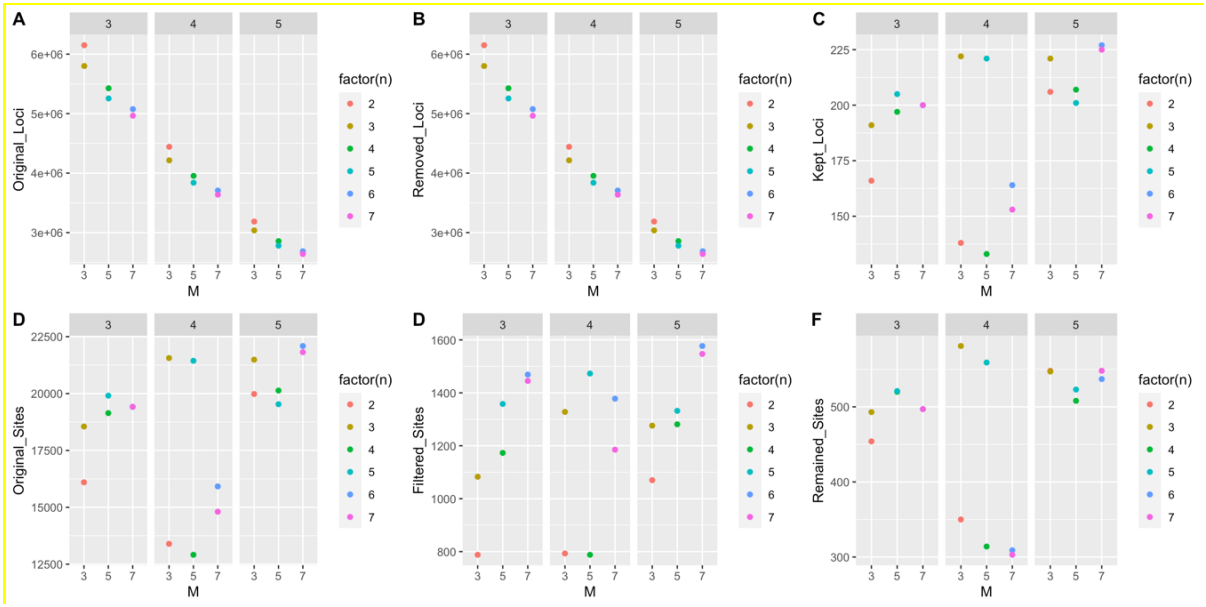


Figure 4. Original, removed, and kept loci and original, filtered, and remained sites according to the 18 Stacks parameters (-m, -M, -n) combinations. The analyses (when r=80) with the highest number of kept loci and remained variant sites were m3M5n5 and m4M5n5.

Data filtering

We used Variant Call Format **VCFTools** version 0.1.16 (Auton and Marcketta, 2009, Danecek et al., 2011) in the two best combinations (m3M5n5 and m4M5n5), the flags applied were: **max-missing (0.65, 0.70, 0.75, 0.80, 0.85)** which excludes sites based on the proportion of missing data, and **maf** (**0.01, 0.02, 0.05**), which only include sites with minor allele frequency equal or greater than the values chosen. All thirty combinations were exported in **Plink** version v1.07 (Purcell et all., 2009) and then, we applied the flag **--mind 0.5** to determine the samples with less than 50% of missing data. After evaluating the number of sites kept and the relevance of individuals removed within the populations, we kept two matrices: m3M5n5/max missing 0.70/maf 0.02 and m4M5n5/max missing 0.70/maf 0.02 (Table 3). For further analyses, we use the final loci list and retained individuals generated by Plink for each matrix as a **whitelist** file in the **populations** (p = 1, r = 0.1) program to obtain our final genotyping matrices (Catchen et al., 2013b).

Table 3.

Trials for VCFTools and Plink filters in m3M5n5 and m4M5n5 arrays. Possible and kept sites according to the flags used by VCFTools and removed and kept individuals after filtering with Plink. * Final matrices chosen for downstream analysis.

Arrays	(-max-missing)	Possible sites	Kept sites	(-maf)	Possible sites	Kept sites	(n)	Plink (rem_ind)	Plink (kept_ind)
m3M5n5	0.65			0.01	19438	9377	166	39	127
m3M5n5	0.65	463957	19438	0.02	19438	6839	166	39	127
m3M5n5	0.65			0.05	19438	4211	166	39	127
m3M5n5	0.70			0.01	9080	4025	166	33	133
m3M5n5*	**0.70**	**463957**	**9080**	**0.02**	**9080**	**2775**	**166**	**33**	**133**
m3M5n5	0.70			0.05	9080	1551	166	33	133
m3M5n5	0.75			0.01	2310	902	166	32	134
m3M5n5	0.75	463957	2310	0.02	2310	504	166	32	134
m3M5n5	0.75			0.05	2310	267	166	32	134
m3M5n5	0.80			0.01	80	28	166	29	137
m3M5n5	0.80	463957	80	0.02	80	16	166	29	137
m3M5n5	0.80			0.05	80	7	166	30	136
m3M5n5	0.85			0.01	0	0	166	0	166
m3M5n5	0.85	0	0	0.02	0	0	166	0	166
m3M5n5	0.85			0.05	0	0	166	0	166
m4M5n5	0.65			0.01	12980	6150	162	40	122
m4M5n5	0.65	333149	12980	0.02	12980	4458	162	40	122
m4M5n5	0.65			0.05	12980	2742	162	40	122
m4M5n5	0.70			0.01	5789	2515	162	35	127
m4M5n5*	**0.70**	**333149**	**5789**	**0.02**	**5789**	**1758**	**162**	**35**	**127**
m4M5n5	0.70			0.05	5789	996	162	35	127
m4M5n5	0.75			0.01	1350	514	162	29	133
m4M5n5	0.75	333149	1350	0.02	1350	332	162	28	134
m4M5n5	0.75			0.05	1350	168	162	28	134
m4M5n5	0.80			0.01	95	28	162	27	135
m4M5n5	0.80	333149	95	0.02	95	16	162	26	136
m4M5n5	0.80			0.05	95	8	162	27	135
m4M5n5	0.85			0.01	1	0	162	0	162
m4M5n5	0.85	333149	1	0.02	1	0	162	0	162
m4M5n5	0.85			0.05	1	0	162	0	162

SNP outliers' identification

We used two packages in **R** version 4.1.1 to identify outlier loci potentially under selection: **PCAdapt** version 4.3.3 (Luu et al., 2017) and **OutFLANK** version 0.2 (Whitlock & Lotterhos, 2015). Following filtering, the **populations** program in Stacks was re-run using SNPs whitelists and SNPs blacklist, to generate information on the final matrices containing putatively neutral and under-selection SNPs.

Population genomic structure

Basic population statistics such as heterozygosity, π (nucleotide diversity), and F_{ST} fixation index were obtained in the **populations** program in Stacks (Catchen et al., 2013b). The final matrices were used as imputed data sets for Principal Component Analyses (**PCA**) and Discriminant Analysis of Principal Components (**DAPC**) in the program **Adegenet** version (2.1.7) (Jombart, 2008; Jombart & Collins, 2022).

Isolation by distance

Linear regressions were performed in **R** version (4.1.1) to explore if there is any signal for a relationship between geographic distance and pairwise genetic distance in all populations.

CHAPTER III

RESULTS

Genotyping and filtering data

Demultiplexing raw data in Stacks resulted in 482 691 496 total reads across all 166 samples in the three libraries. Of the total reads, 0.249% (1 201 716) were low-quality reads and 91.172% (440 080 539) reads were maintained for the following analyses (Table 4). Demultiplexed data was mapped with an average coverage of 8.61x for m = 3, 10.95x for m = 4, and 13.72x for m = 5 (Figure 5).

Table 4.

Demultiplexed data on *process_radtags* program. Number and reads type after demultiplexing the raw data of three libraries with the *process_radtags* program in Stacks. Note that there are six rows in the column library name because each library has two files.

Library name	Retained reads	Low quality	Barcode not found	Rad cutsite not found	Total reads
LIB6POOL1	124950915	1061326	16112843	2109522	144234606
LIB6POOL2	7591	61	18046	13324	39022
LIB7POOL1	91170838	43547	1477855	1459726	94151966
LIB7POOL2	54477551	26593	778233	780782	56063159
LIB8POOL1	131734711	55683	2502417	3349743	137642554
LIB8POOL2	37738933	14506	11865956	940794	50560189
Total 3	**440080539**	**1201716**	**32755350**	**8653891**	**482691496**
%	**91.172**	**0.249**	**6.786**	**1.793**	**100**

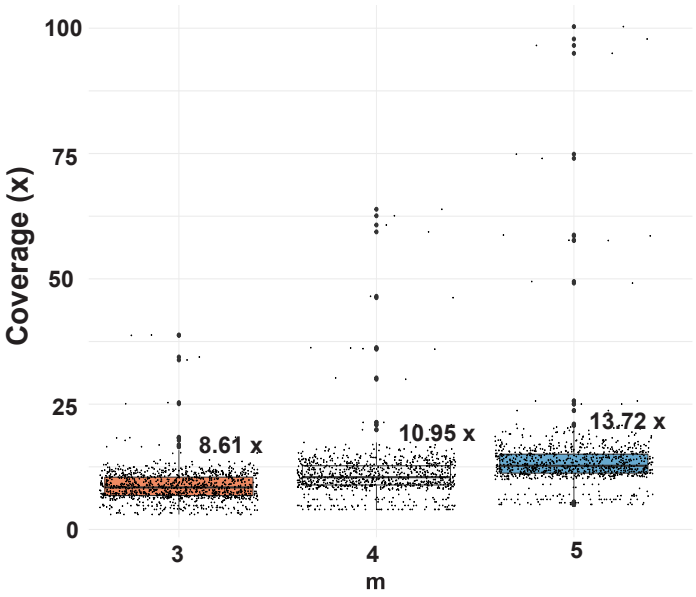


Figure 5. *Epipedobates anthonyi* mean coverage plot for the different combinations in *DENOVO_MAP.PL* with parameters -m; 8.61x for m=3, 10.95x for m=4, and 13.72x for m=5. Each point represents one individual/sample. Note that the distribution of all the points in the different boxplots is similar with the three different m values.

Using optimal parameter settings in DENOVO_MAP.PL, the best two choices, due to the highest number of loci and sites, were m3M5n5 with 19 908 sites and 5 255 146 loci, whereas the m4M5n5 trial got 21 441sites and 3 838 574 loci (Table 5, Figure 4).

Table 5.

Loci and sites according to the best trials in *DENOVO_MAP.PL* and populations program.

Trial	Populations parameter	Original Loci	Removed Loci	Kept Loci	Original Sites	Filtered Sites	Remained Sites	Mean Sites
m3M5n5	r = 0.8	5255146	5254941	205	19908	1358	521	97.11
m3M5n5	r (default)	5255146	0	5255146	505009466	13	3604804	96.10
m4M5n5	r = 0.8	3838574	3838353	221	21441	1473	559	97.02
m4M5n5	r (default)	3838574	0	3838574	368894790	17	2528694	96.10

Outlier SNPs identification

Two final matrices remained after filtering individuals and loci for missing data and loci for minor allele frequencies. The **m3M5n5** matrix with **2775** SNPs among **133** individuals and the **m4M5n5** matrix with **1758** SNPs among **127** samples (Table 3).

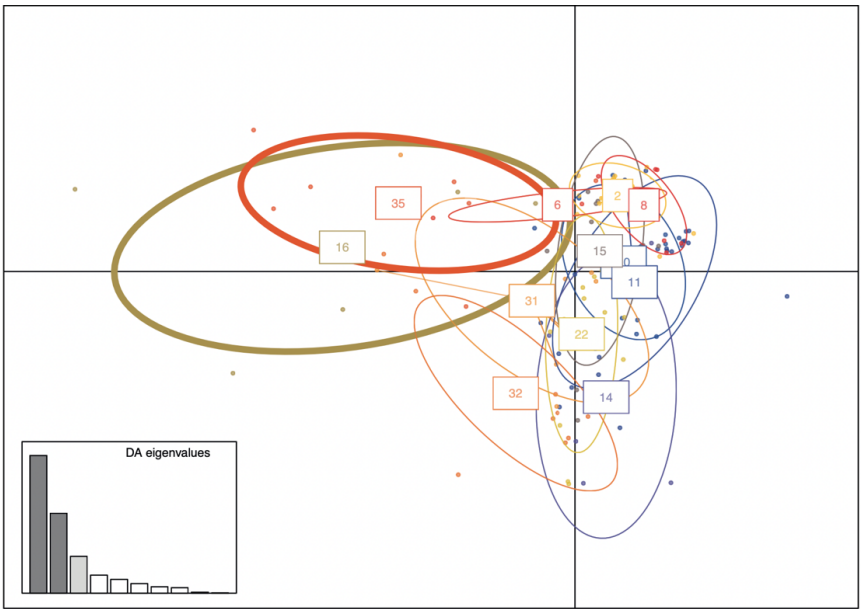
These two matrices were used for identifying outlier SNPs. **PCAdapt** detected 69 putatively adaptative outlier SNPs in the m3M5n5 matrix, and 48 putatively adaptative outlier SNPs in the m4M5n5 (Table 6, Table S3). We also run the analyses removing two low-elevation populations that had high genetic variation, which could increase background F_{ST}, and therefore, hamper outlier detection. However, those results did not differ much from the ones with all populations; therefore, we show results for all populations included. The outlier SNPs identified were the base to create the whitelist and blacklist file used for downstream neutral and possible selection populations analysis.

Table 6. Putatively outlier and neutral SNPs identified in PCAdapt.

Trial	Populations (n)	SNPs (n)	qvalue outliers	Benjamini Hochberg outliers	Bonferroni Correction outliers	Samples
m3M5n5	12	2775	69	69	20	133
m3M5n5*	10	2775	60	60	19	118
m4M5n5	12	1758	48	48	27	127
m4M5n5*	10	1758	53	53	24	115

* Trials that were run without the populations EA16-Quera, and EA35-San Vicente.

A



B

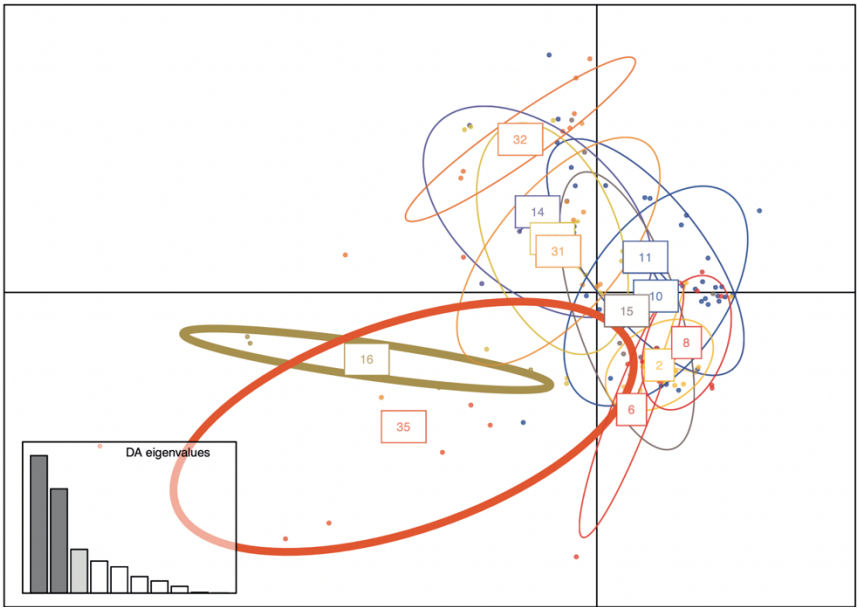


Figure 6. Population structure of *Epipedobates anthonyi*. (A) DAPC with 2775 SNPs in the m3M5n5 matrix, (B) general DAPC with all the 1758 SNPs in the m4M5n5 matrix. Note that the bigger circles represent the populations 16 and 35 that have higher variation in comparison with the rest of the demes. Each circle and color represent a different population, and each point is a sample. The x and y axes indicate the first and second principal components of DAPC, respectively. The left inset shows a barplot for DA-eigenvalues (measure of the amount of variability retained).

Population genomics

Twelve populations were geographically identified for this study, and population genetic statistics were estimated for each defined deme in **populations** by Stacks (Table 7). Standard population genomic parameters calculated in the m3M5n5 matrix with 2775 SNPs were: observed heterozygosity (range = 0.112 – 0.152), expected heterozygosity (0.113 – 0.163), inbreeding coefficient (F_{is} = 0.023 – 0.120), and nucleotide diversity (π = 0.00139 – 0.0018). For the m4M5n5 matrix with 1758 SNPs: observed heterozygosity (0.120 – 0.157), expected heterozygosity (0.116 – 0.164), inbreeding coefficient (0.022 – 0.119), and nucleotide diversity (π = 0.0014 – 0.0018) across all sites (variant and fixed). Population EA14 (Saracay) was the most diverse based on nucleotide diversity (π = 0.00183) in both matrices. The fixation index (F_{ST}) was obtained among all populations for the two matrices (m3M5n5 and m4M5n5). The

F_{ST} in the m3M5n5 matrix ranged between 0.0149223 and 0.127945; the F_{ST} in the m4M5n5 matrix ranged between 0.0149357 and 0.124082 (Table S1). Before running the population analysis, the DAPCs identified two populations (EA16-Quera, and EA35-San Vicente) with a higher variation within the group (Figure 6). We run the populations program eliminating these two populations, but there were no differences with the analyses made with twelve populations, so we show results with all the study sites.

Table 7.

Population statistics for *Epipedobates anthonyi.* (Pop_alt) identification and altitude populations, (Ho) observed heterozygosity, (He) expected heterozygosity, (F_{is}) inbreeding coefficient, and (π) nucleotide diversity.

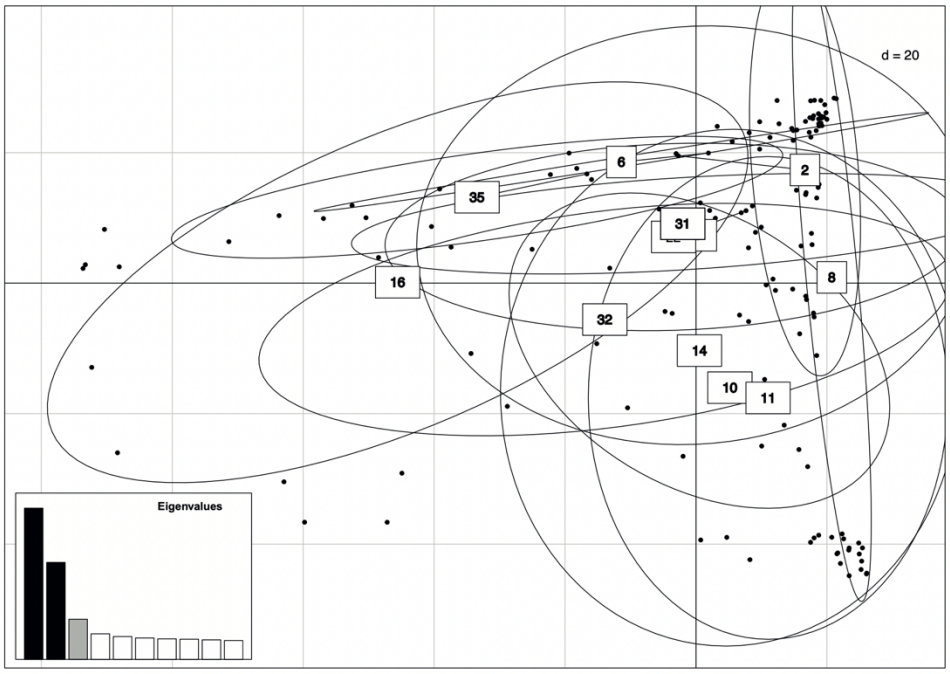
	m3M5n5				m4M5n5			
Pop_alt	**Ho**	**He**	**Fis**	**Pi**	**Ho**	**He**	**Fis**	**Pi**
EA2-high	0.123	0.147	0.111	0.00157	0.131	0.153	0.105	0.00163
EA15-high	0.137	0.159	0.092	0.00174	0.139	0.159	0.087	0.00174
EA11-high	0.143	**0.163**	0.066	0.00172	0.149	**0.164**	0.057	0.00174
EA8-mid	0.119	0.160	0.113	0.00177	0.126	0.164	0.110	0.00181
EA6-mid	0.121	**0.113**	**0.023**	**0.00139**	0.126	**0.116**	**0.022**	**0.00144**
EA10-mid	0.112	0.151	**0.120**	0.00165	**0.120**	0.156	**0.119**	0.00170
EA22-low	**0.152**	0.162	0.057	0.00175	**0.157**	0.160	0.036	0.00173
EA16-low	0.112	0.131	0.080	0.00155	0.129	0.121	0.035	0.00152
EA14-low	0.141	0.154	0.071	**0.00181**	0.146	0.156	0.066	**0.00183**
EA31-lowl	0.140	0.151	0.057	0.00167	0.146	0.153	0.054	0.00171
EA32-lowl	0.125	0.149	0.099	0.00164	0.138	0.154	0.076	0.00171
EA35-lowl	**0.112**	0.127	0.067	0.00142	0.124	0.133	0.054	0.00151

Population structure with neutral and outlier loci

We applied Principal Component Analyses (**PCAs**) and Discriminant Analysis on Principal Components (**DAPCs)** with the m3M5n5 and m4M5n5 matrices using **Adegenet** version (2.1.7) (Jombart, 2008; Jombart & Collins, 2022). The plots with both matrices look relatively similar (Figures 7-10). Main population clustering

reflected geographical distribution. All the populations from mid and high altitudes over the complete, neutral, and putative outlier SNPs matrices were grouped together (Figures 11-14). Geographical distribution by transects also was identified in the clustering, transects one and three are the most conspicuous (Figures 15 - 18).

A



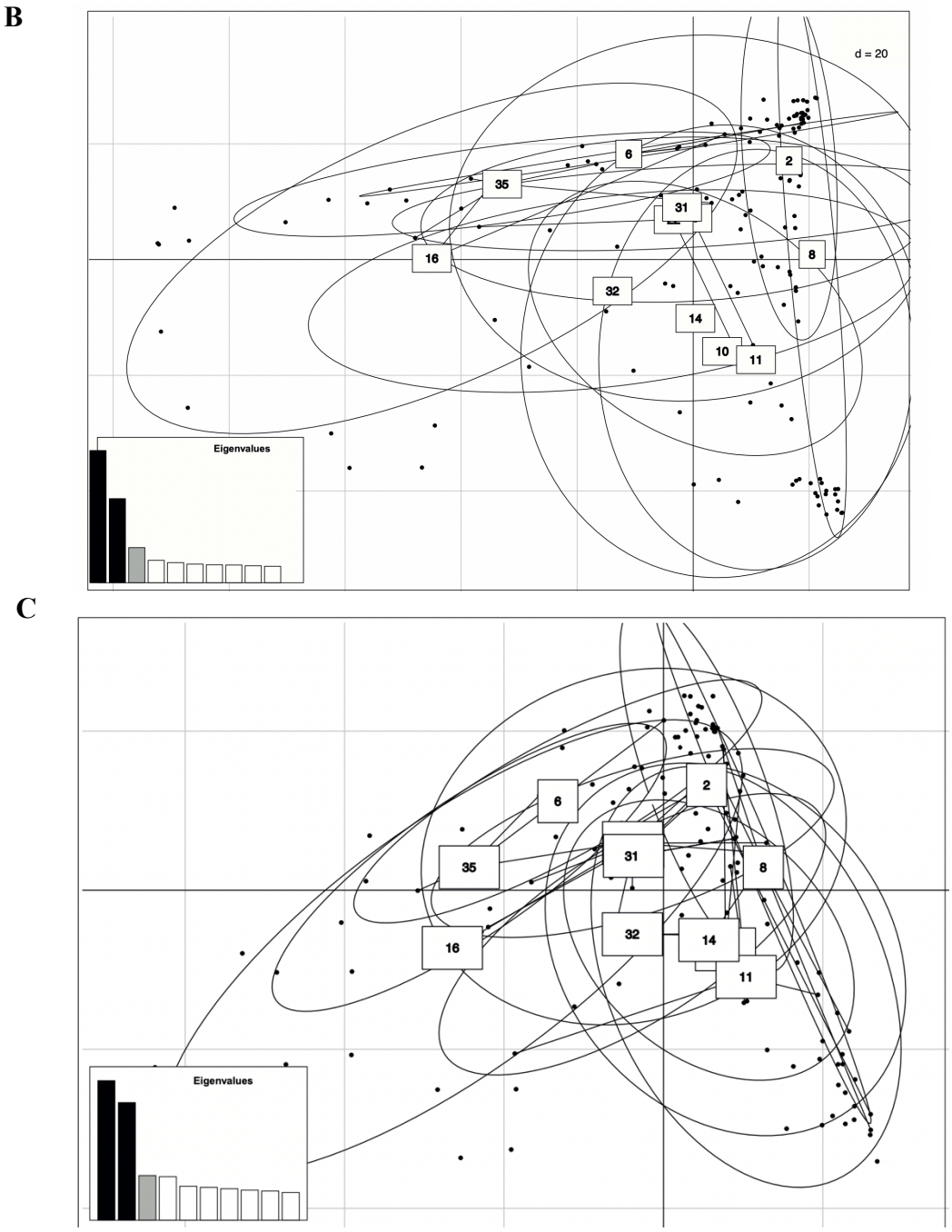
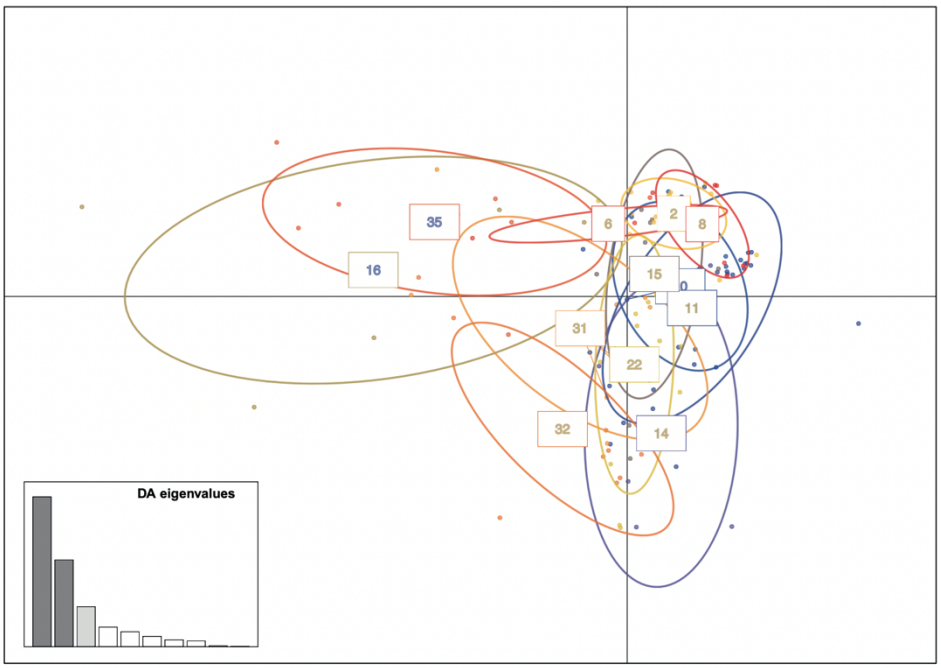
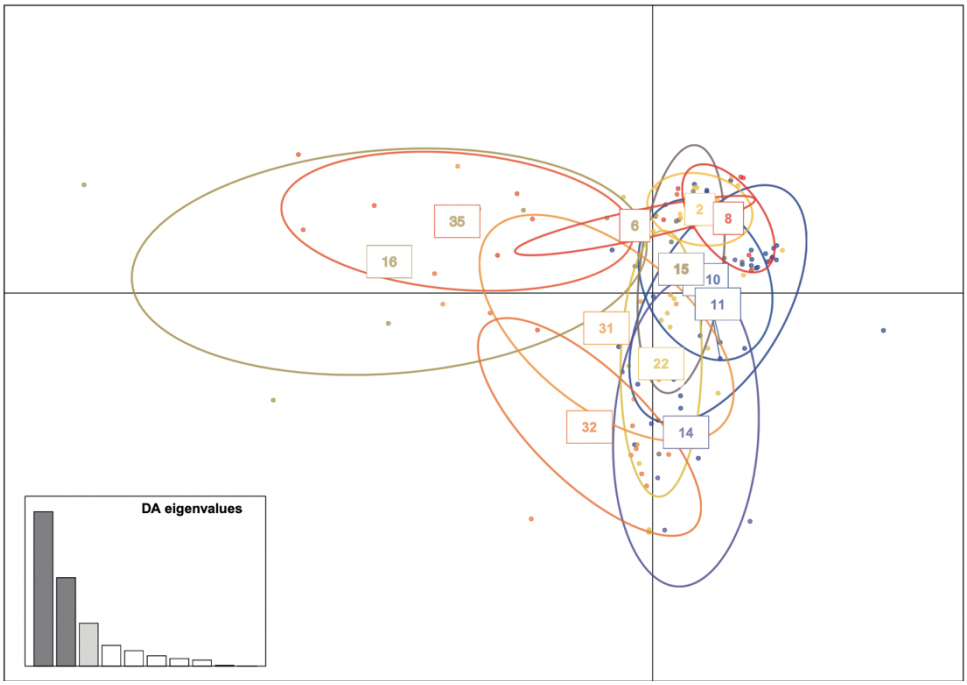


Figure 7. PCAs (Principal Component Analysis) for *Epipedobates anthonyi* population structure **m3M5n5** matrix. A) PCA from the 2775 SNPs matrix, B) PCA from the 2706 neutral SNPs matrix and C) PCA from the 69 putative outlier SNPs matrix. Each circle represents a different population, and each point is a sample. The x and y axes indicate the first and second principal components of PCA, respectively. The left inset shows a barplot for eigenvalues (measure of the amount of variability retained).

A


35
6
2
8
16
15
10
11
31
22
32
14
DA eigenvalues

B


35
6
2
8
16
15
10
11
31
22
32
14
DA eigenvalues

C

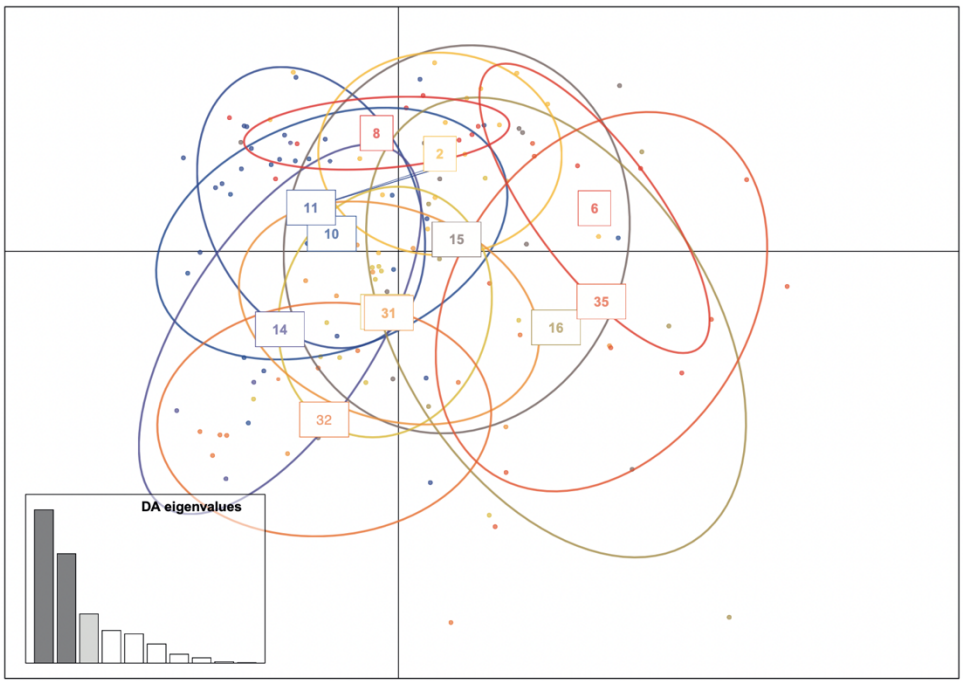
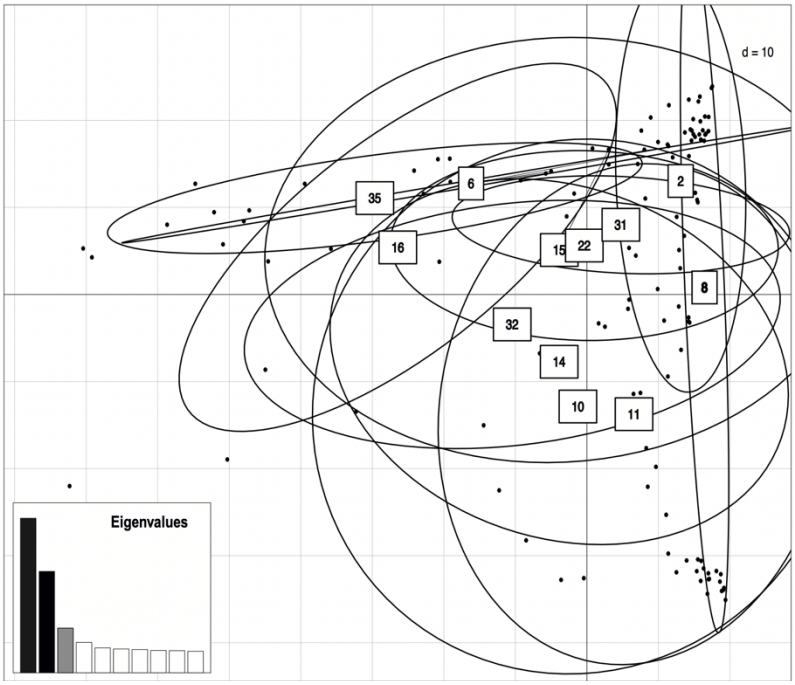
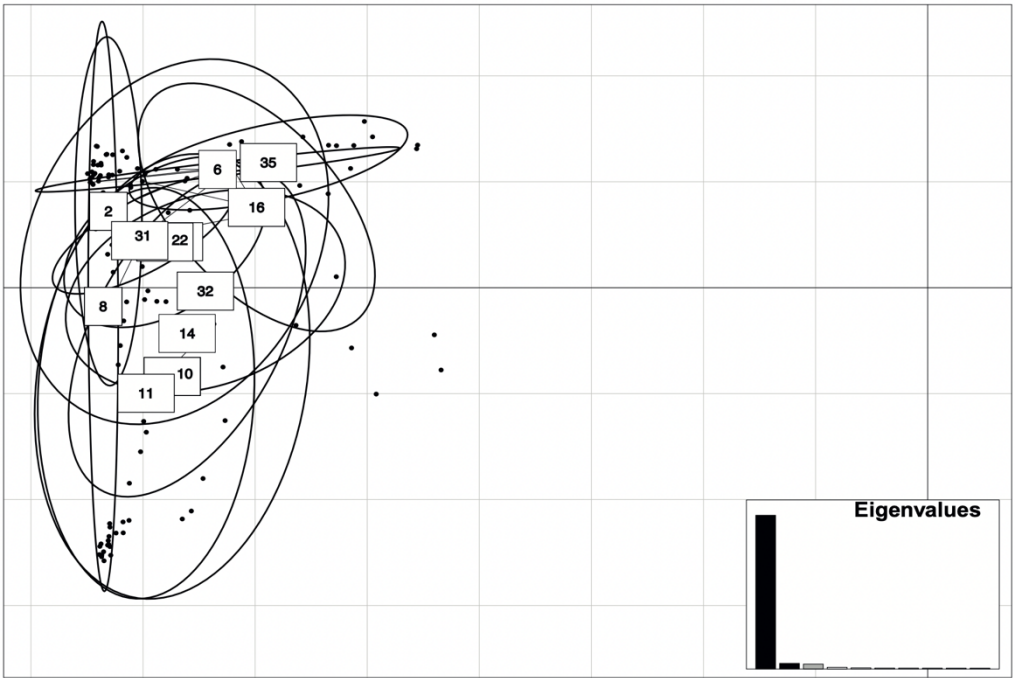


Figure 8. DAPCs (Discriminant Analysis of Principal Components) for *Epipedobates anthonyi* population structure **m3M5n5** matrix. A) DAPC from the 2775 SNPs matrix, B) DAPC from the 2706 neutral SNPs matrix and C) DAPC from the 69 putative outlier SNPs matrix. Each circle and color represent a different population, and each point is a sample. The x and y axes indicate the first and second principal components of DAPC, respectively. The left inset shows a barplot for DA-eigenvalues (measure of the amount of variability retained).

A


d = 10
35
6
2
31
16
15
22
8
32
14
10
11
Eigenvalues

B


6
35
2
16
31
22
32
8
14
10
11
Eigenvalues

C

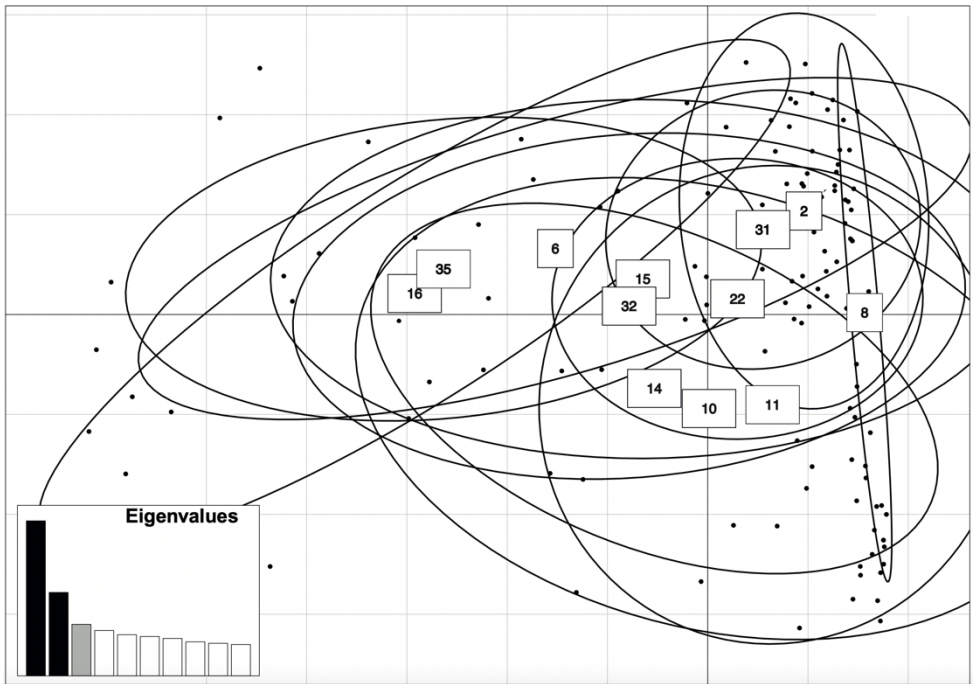
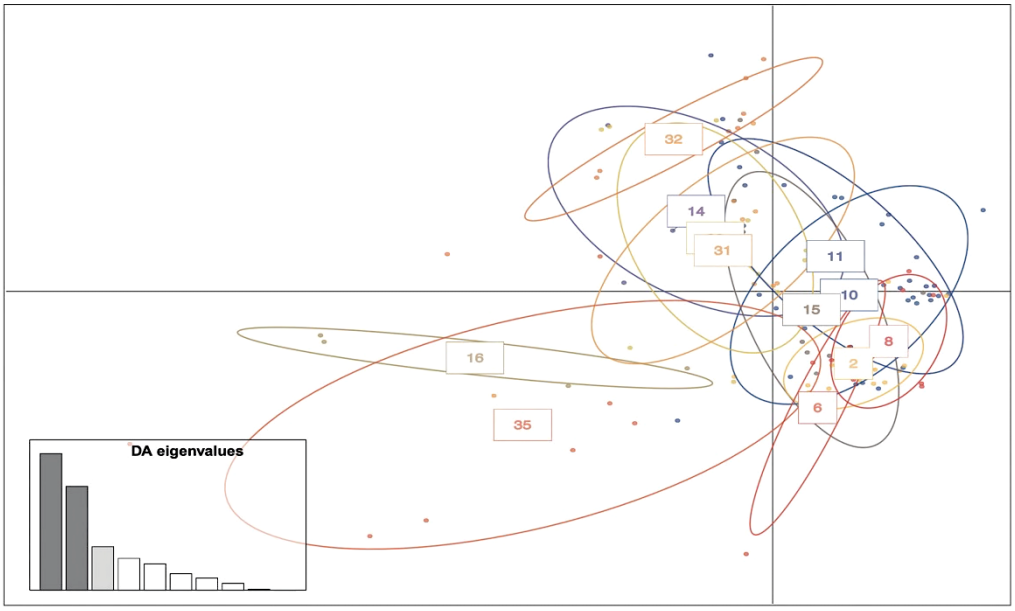
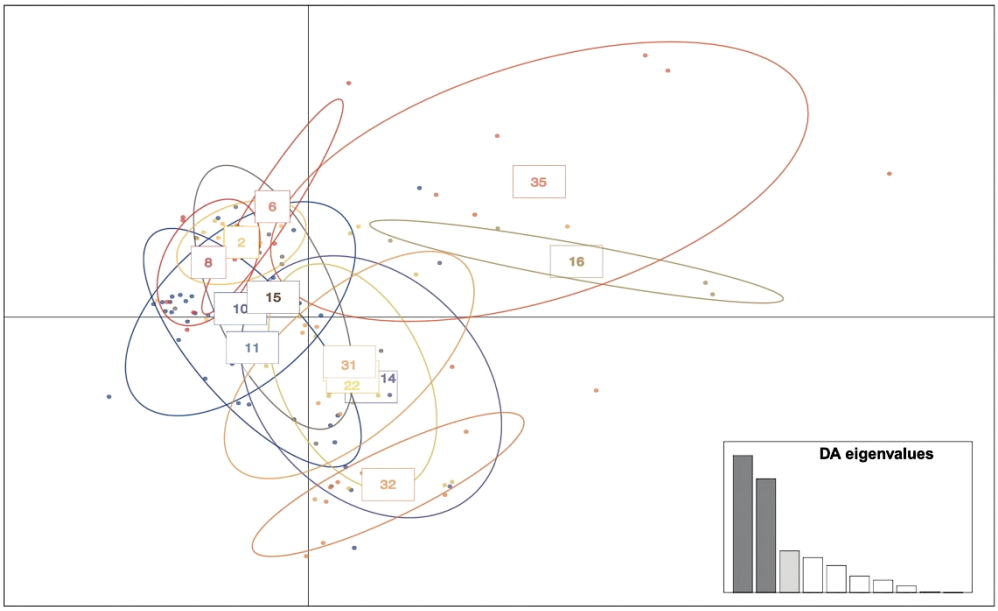


Figure 9. PCAs (Principal Component Analysis) for *Epipedobates anthonyi* population structure **m4M5n5** matrix. A) PCA from the 1758 SNPs matrix, B) PCA from the 1710 neutral SNPs matrix and C) PCA from the 48 putative outlier SNPs matrix. Each circle represents a different population, and each point is a sample. The x and y axes indicate the first and second principal components of PCA, respectively. The inset shows a barplot for eigenvalues (measure of the amount of variability retained).

A


32
14
31
11
10
15
8
2
6
16
35
DA eigenvalues

B


35
6
2
8
16
15
10
11
31
14
22
32
DA eigenvalues

C

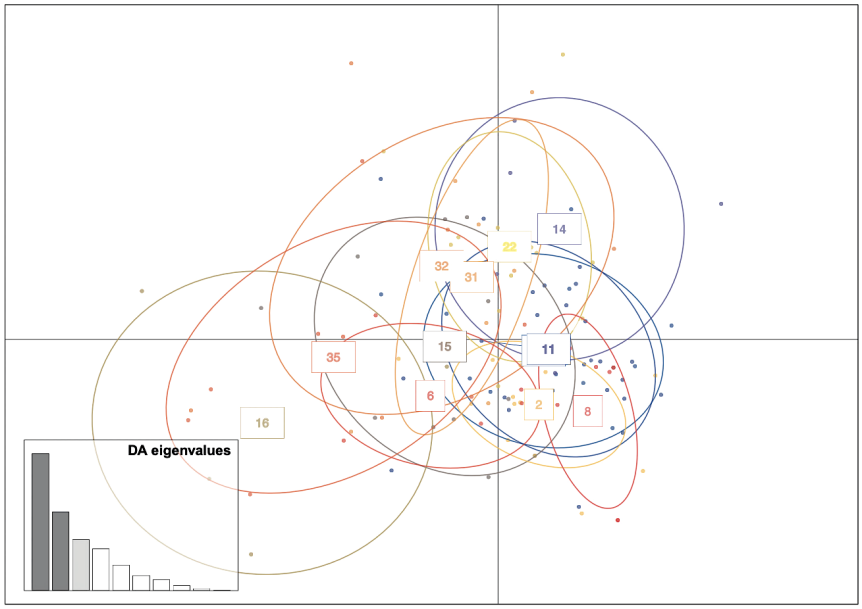
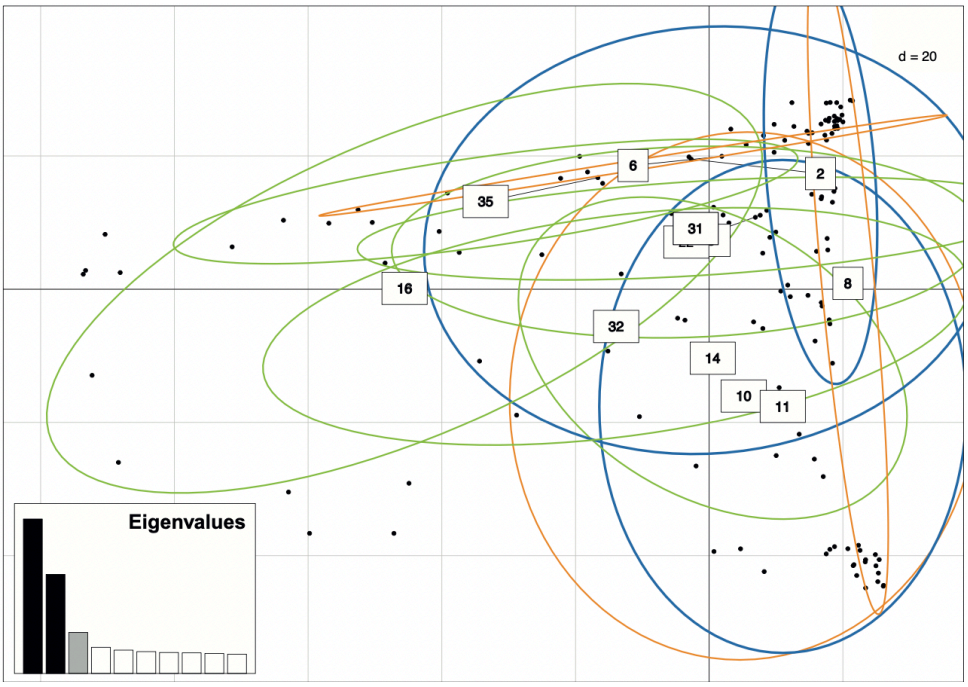
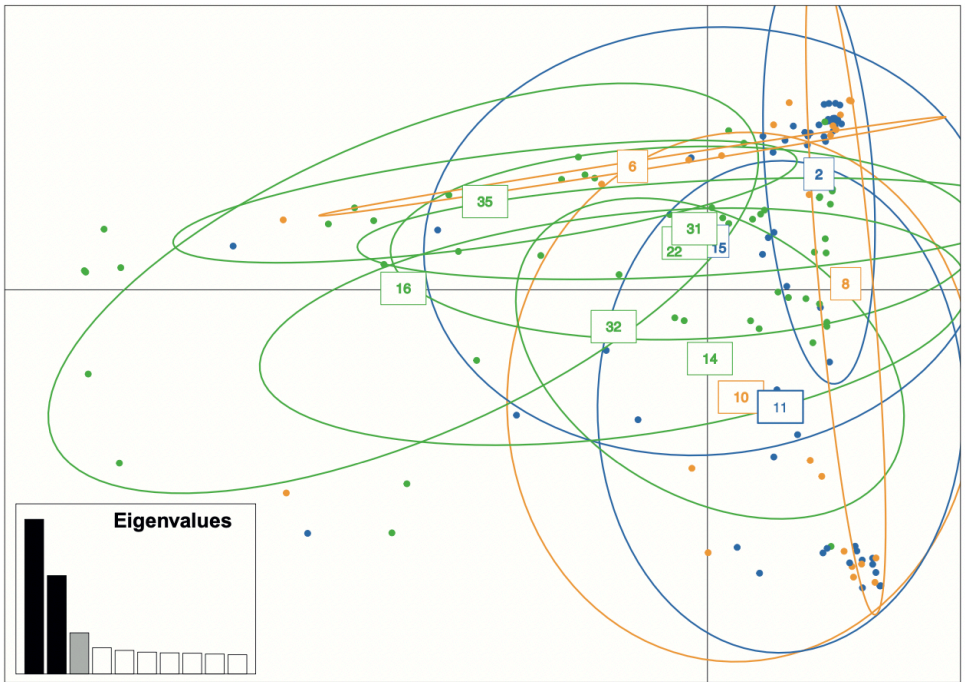


Figure 10. DAPCs (Discriminant Analysis of Principal Components) for *Epipedobates anthonyi* population structure **m4M5n5** matrix. A) DAPC from the 1758 SNPs matrix, B) DAPC from the 1710 neutral SNPs matrix and C) DAPC from the 48 putative outlier SNPs matrix. Each circle and color represent a different population, and each point is a sample. The x and y axes indicate the first and second principal components of DAPC, respectively. The left inset shows a barplot for DA-eigenvalues (measure of the amount of variability retained).

A


d = 20
6
2
35
31
16
8
32
14
10
11
Eigenvalues

B


6
2
35
31
22
15
16
8
32
14
10
11
Eigenvalues

C

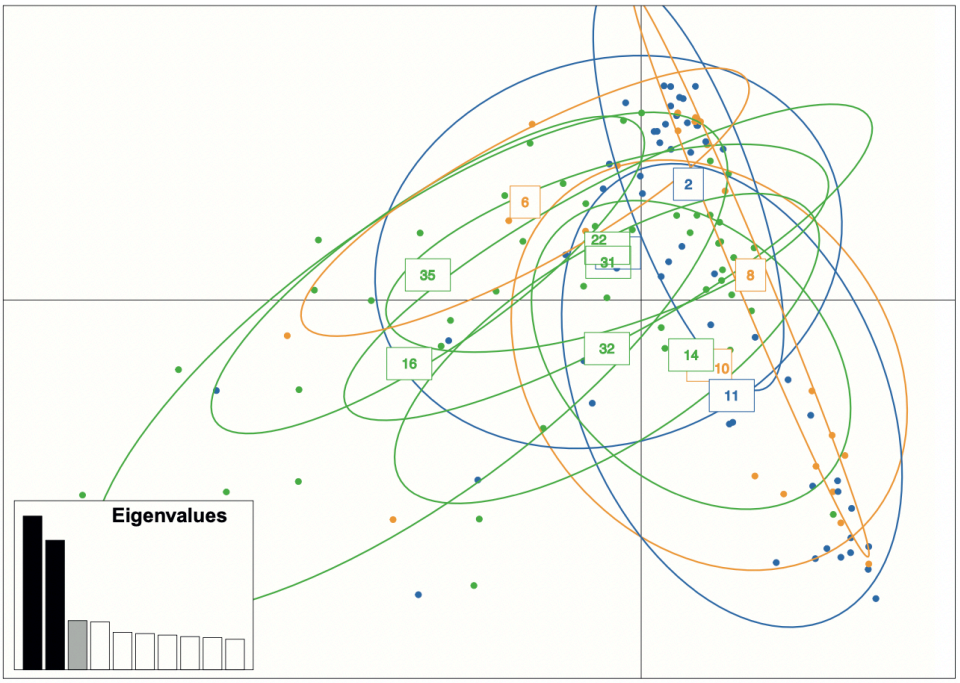
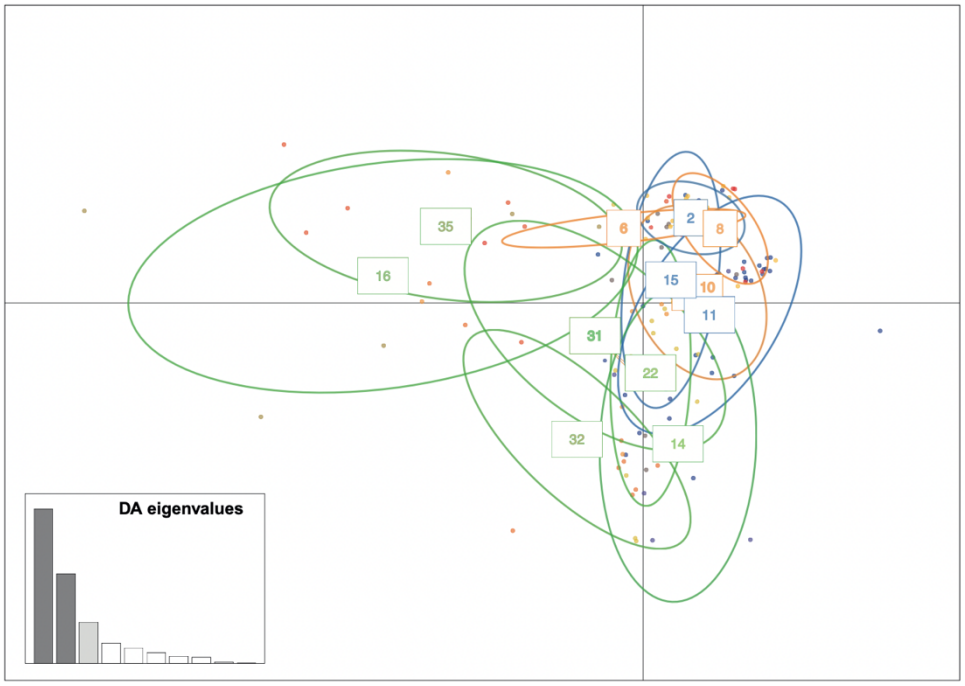
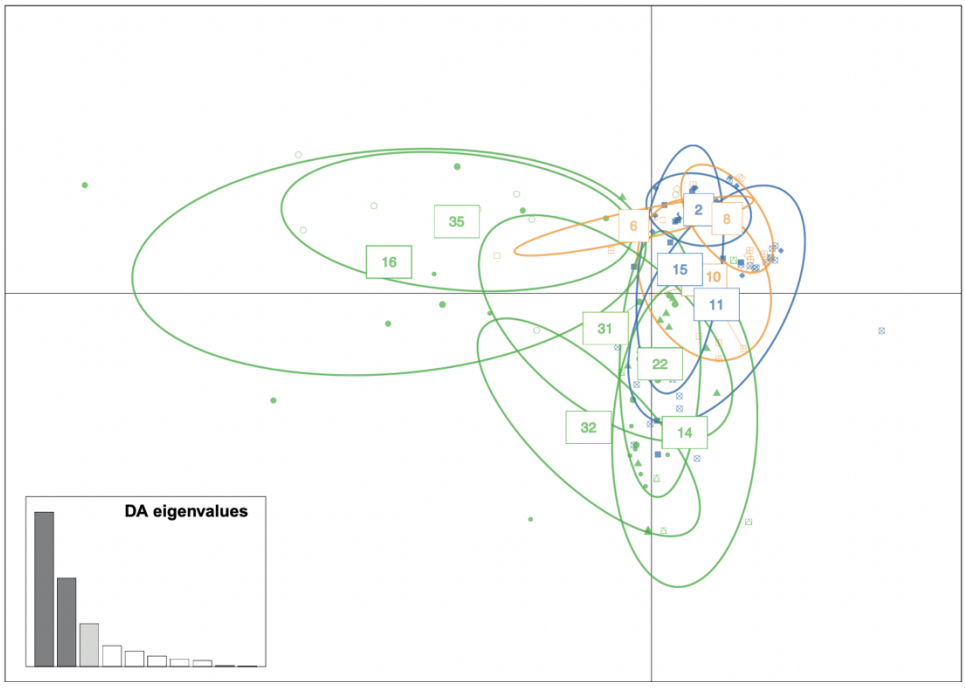


Figure 11. PCAs (Principal Component Analysis) for *Epipedobates anthonyi* population structure **m3M5n5** matrix according to the distribution of the altitudes. A) PCA from the 2775 SNPs matrix, B) PCA from the 2706 neutral SNPs matrix and C) PCA from the 69 putative outlier SNPs matrix. Each circle represents a different population, and each point is a sample. Altitudes are shown as green color for low altitudes, blue for mid altitudes, and orange for the highest altitudes. The x and y axes indicate the first and second principal components of PCA, respectively. The left inset shows a barplot for eigenvalues (measure of the amount of variability retained).

A


35
16
6
2
8
15
10
11
31
22
32
14
DA eigenvalues

B


35
16
6
2
8
15
10
11
31
22
32
14
DA eigenvalues

C

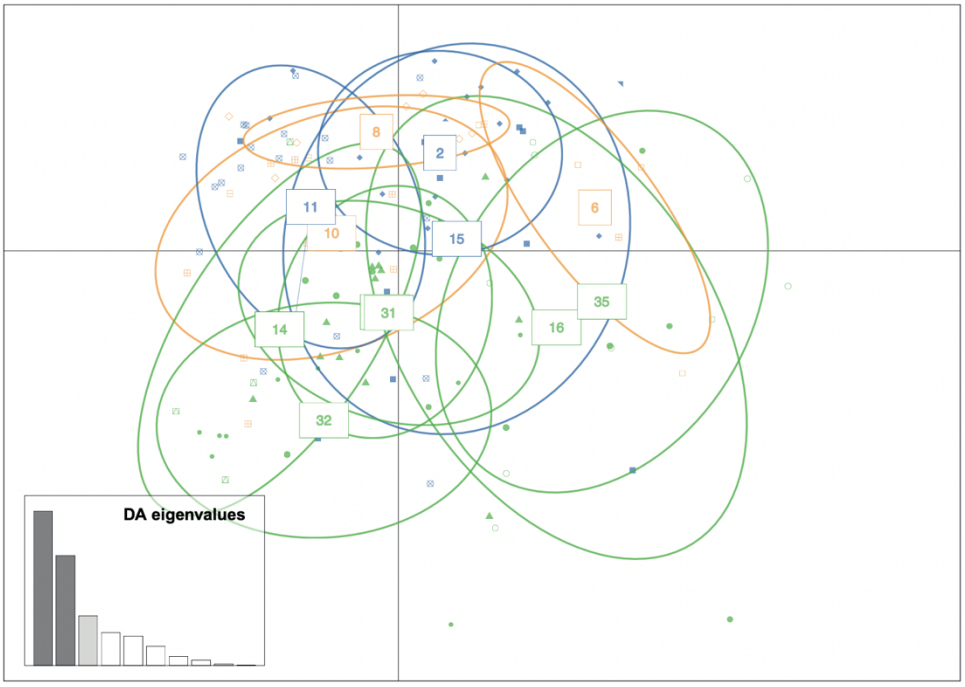
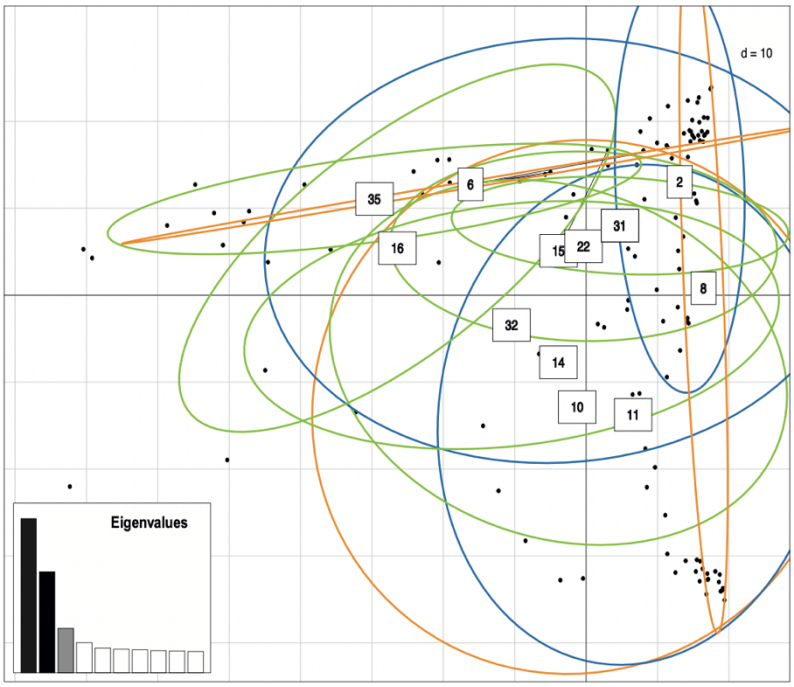
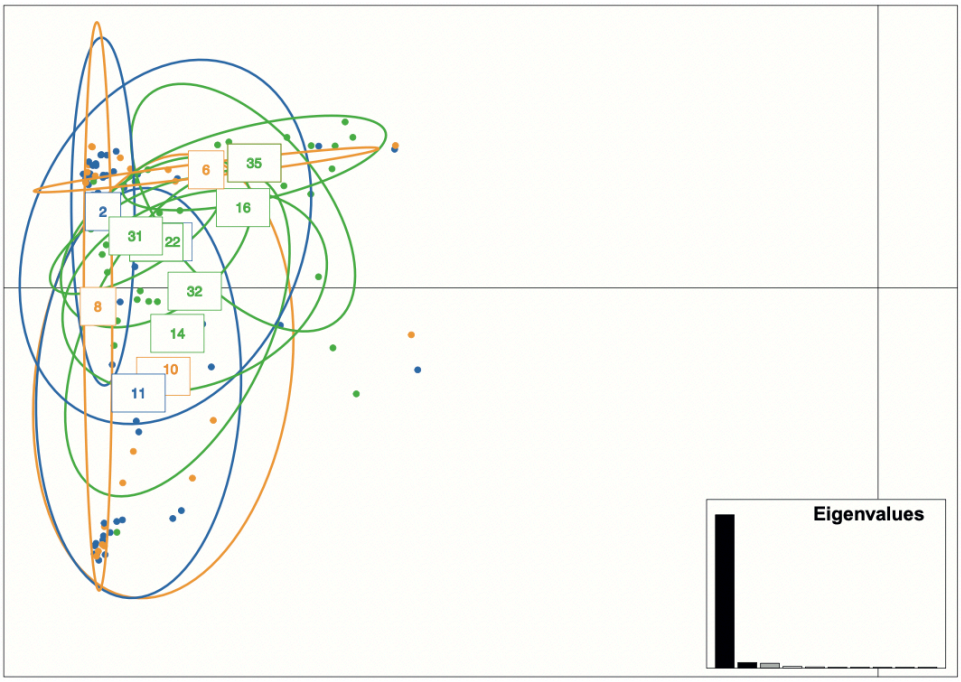


Figure 12. DAPCs (Discriminant Analysis of Principal Components) for *Epipedobates anthonyi* population structure **m3M5n5** matrix according to the distribution of the altitudes.A) DAPC from the 2775 SNPs matrix, B) DAPC from the 2706 neutral SNPs matrix and C) DAPC from the 69 putative outlier SNPs matrix. Each circle represents a different population, and each point is a sample. Altitudes are shown as green color for low altitudes, blue for mid altitudes, and orange for the highest altitudes. The x and y axes indicate the first and second principal components of DAPCP, respectively. The left inset shows a barplot for DA-eigenvalues (measure of the amount of variability retained).

A


d = 10
35
6
2
16
31
15
22
8
32
14
10
11
Eigenvalues

B


6
35
2
16
31
22
32
8
14
10
11
Eigenvalues

C

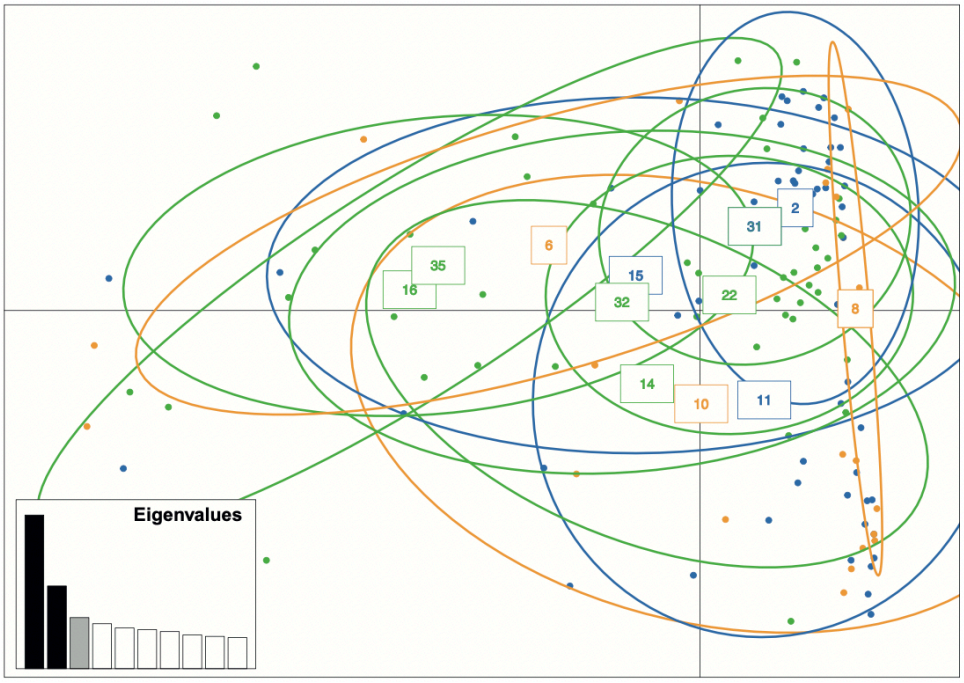
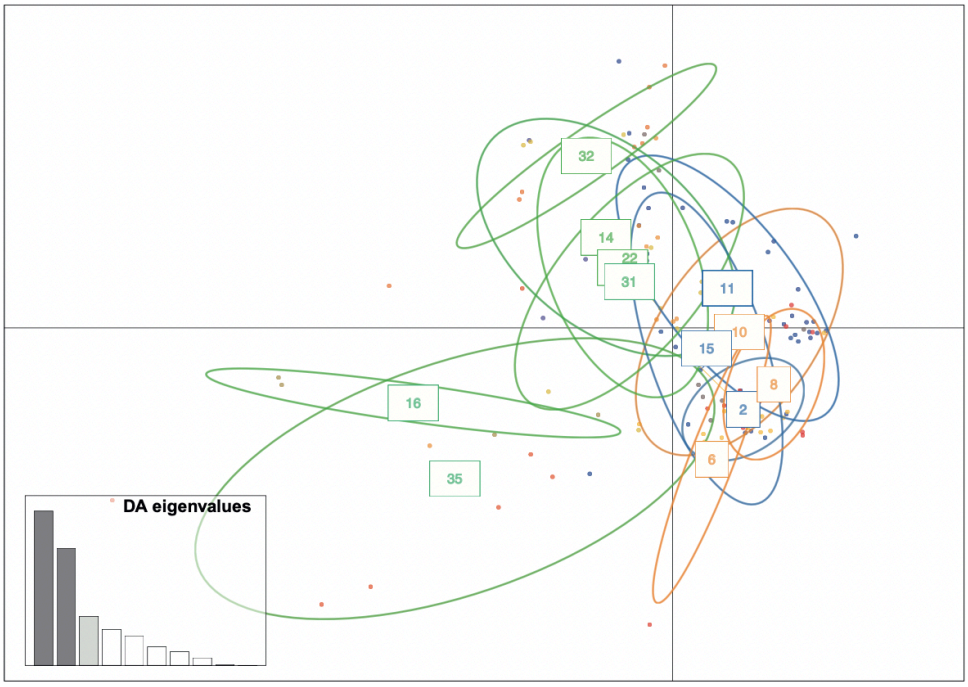
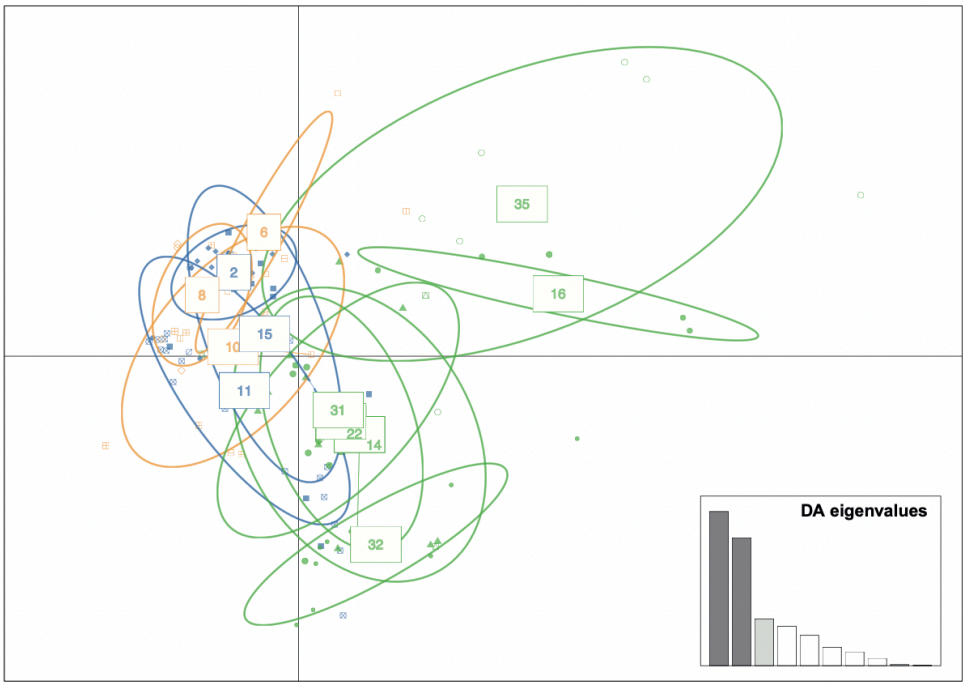


Figure 13. PCAs (Principal Component Analysis) for *Epipedobates anthonyi* population structure **m4M5n5** matrix according to the distribution of the altitudes. A) PCA from the 1758 SNPs matrix, B) PCA from the 1710 neutral SNPs matrix and C) PCA from the 48 putative outlier SNPs matrix. Each circle represents a different population, and each point is a sample. Altitudes are shown as green color for low altitudes, blue for mid altitudes, and orange for the highest altitudes. The x and y axes indicate the first and second principal components of PCA, respectively. The left inset shows a barplot for eigenvalues (measure of the amount of variability retained).

A


32
14
22
31
11
10
15
8
2
6
16
35
DA eigenvalues

B


35
6
2
8
16
15
10
11
31
22
14
32
DA eigenvalues

C

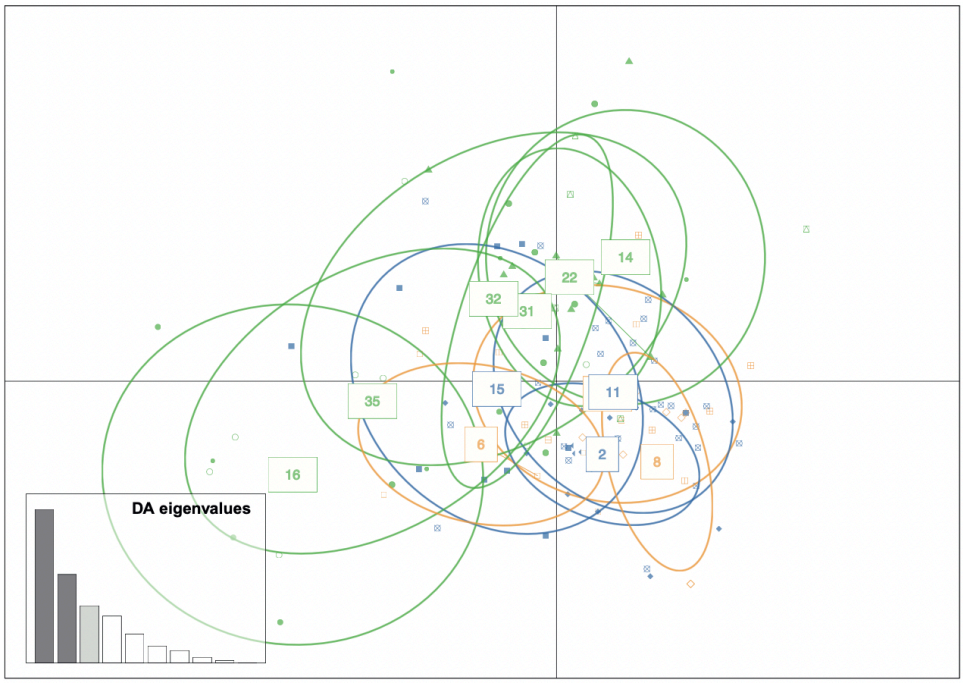
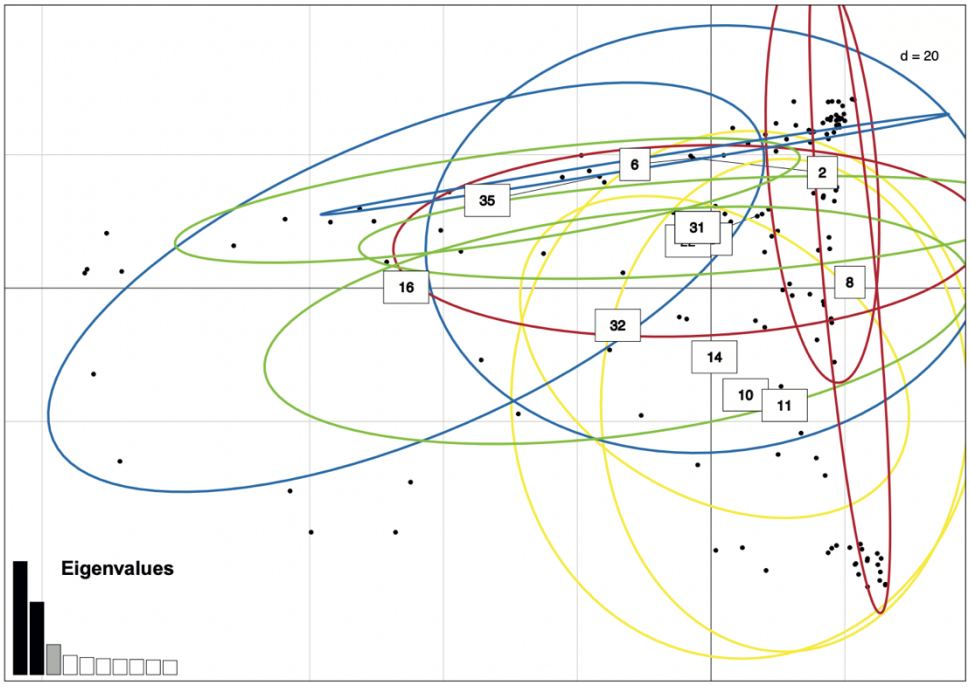
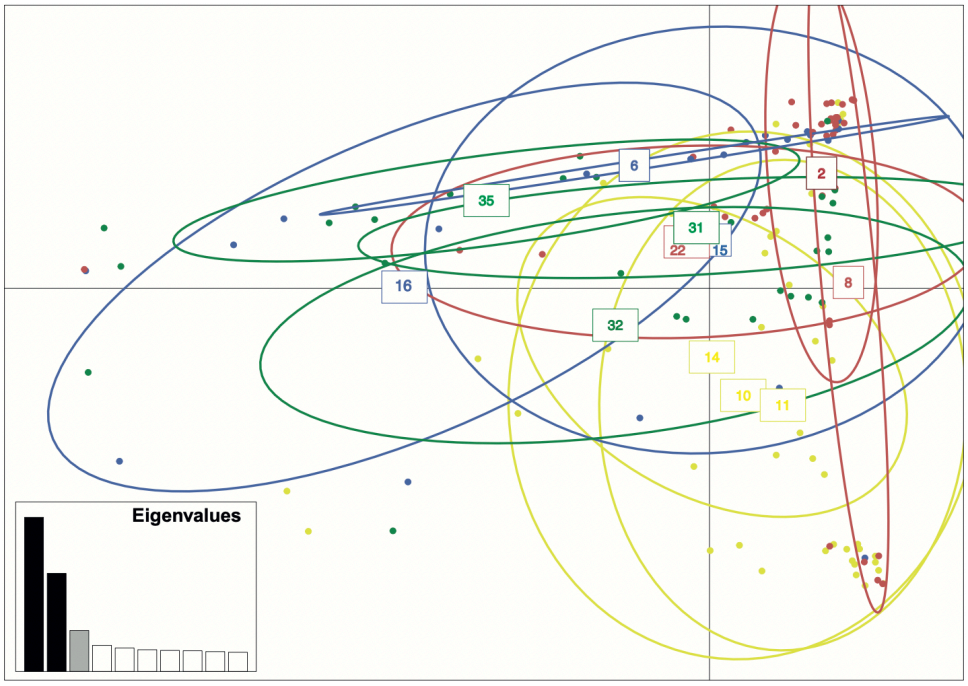


Figure 14. DAPCs (Discriminant Analysis of Principal Components) for *Epipedobates anthonyi* population structure **m4M5n5** matrix according to the distribution of the altitudes. A) DAPC from the 1758 SNPs matrix, B) DAPC from the 1710 neutral SNPs matrix and C) DAPC from the 48 putative outlier SNPs matrix. Each circle represents a different population, and each point is a sample. Altitudes are shown as green color for low altitudes, blue for mid altitudes, and orange for the highest altitudes. The x and y axes indicate the first and second principal components of DAPCP, respectively. The left inset shows a barplot for DA-eigenvalues (measure of the amount of variability retained).

A


d = 20
6
2
35
31
16
8
32
14
10
11
Eigenvalues

B


6
2
35
31
22
15
16
8
32
14
10
11
Eigenvalues

C

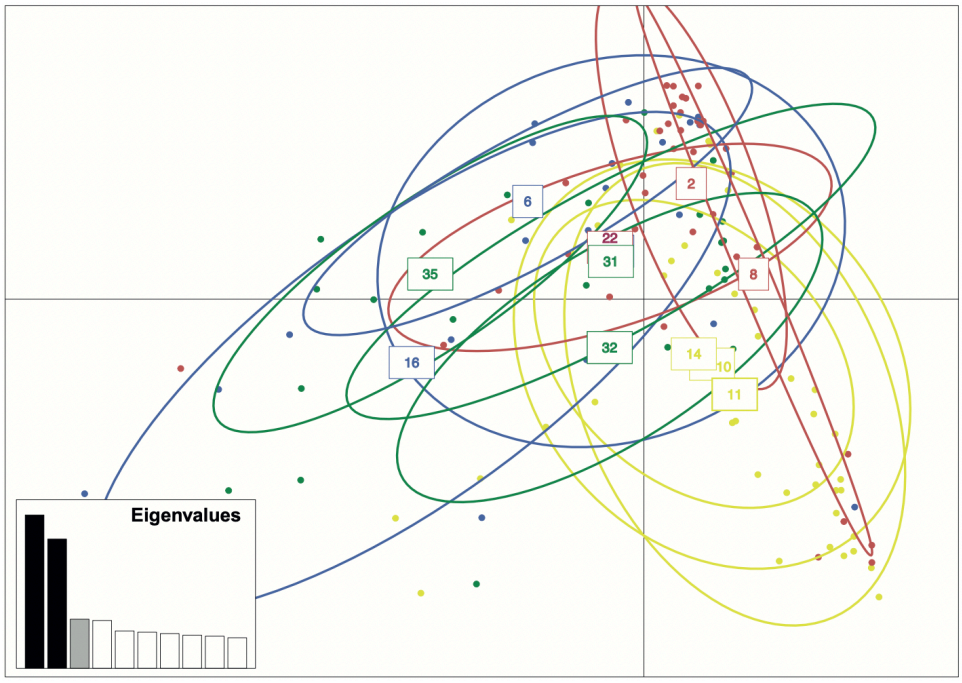
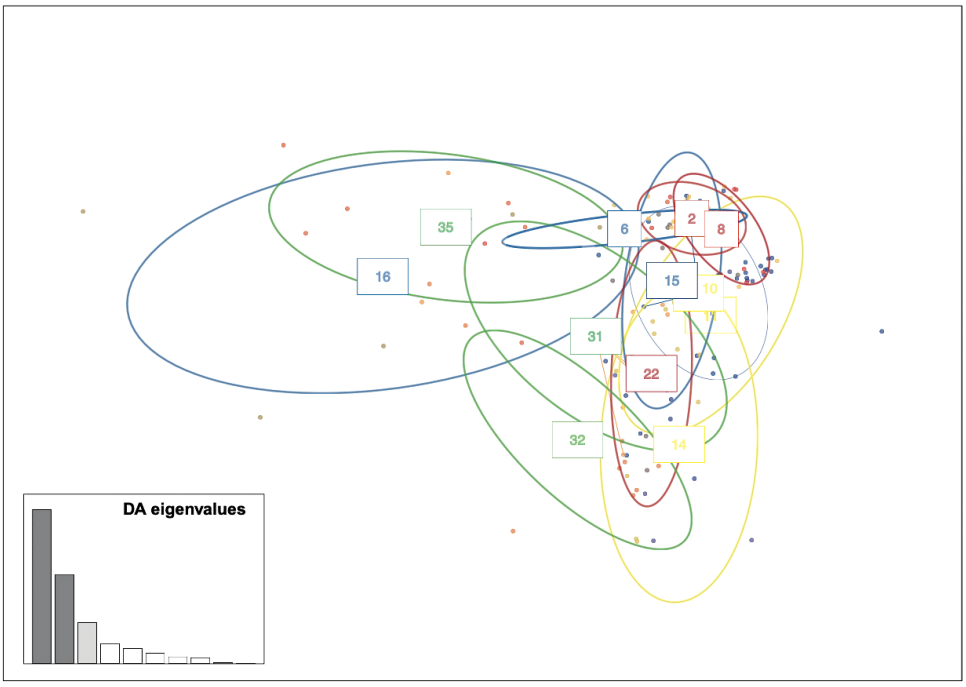
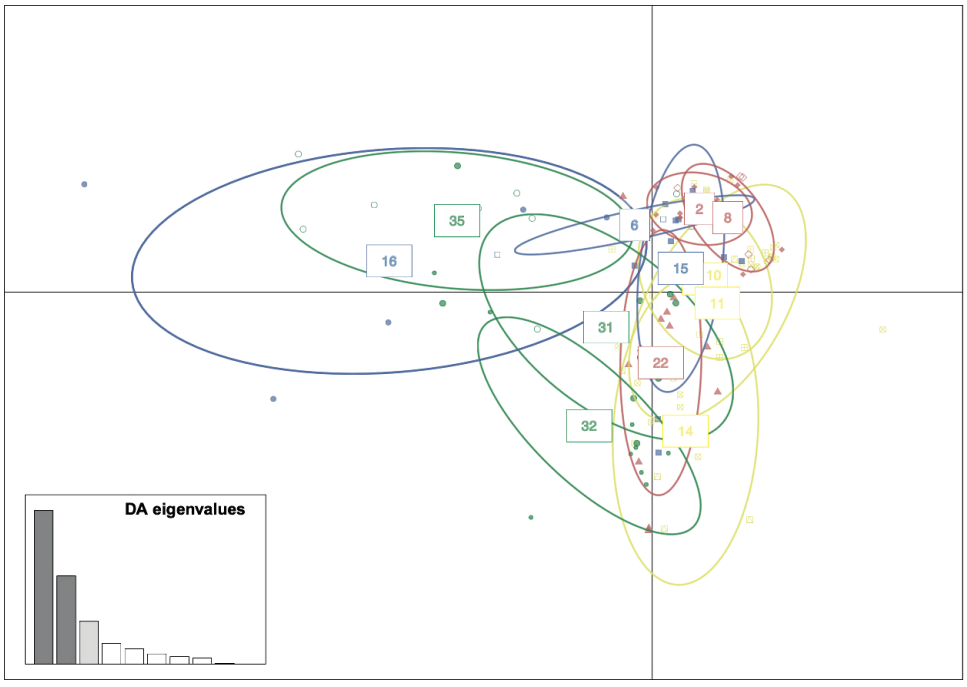


Figure 15. PCAs (Principal Component Analysis) for *Epipedobates anthonyi* population structure **m3M5n5** matrix according to the distribution of the transects. A) PCA from the 2775 SNPs matrix, B) PCA from the 2706 neutral SNPs matrix and C) PCA from the 69 putative outlier SNPs matrix. Each circle represents a different population, and each point is a sample. Transects are shown as red for transect one, blue for transect two, yellow for transect three, and green for lowlands. The x and y axes indicate the first and second principal components of PCA, respectively. The left inset shows a barplot for eigenvalues (measure of the amount of variability retained).

A


35
6
2
8
16
15
10
31
22
32
14
DA eigenvalues

B


35
6
2
8
16
15
10
11
31
22
32
14
DA eigenvalues

C

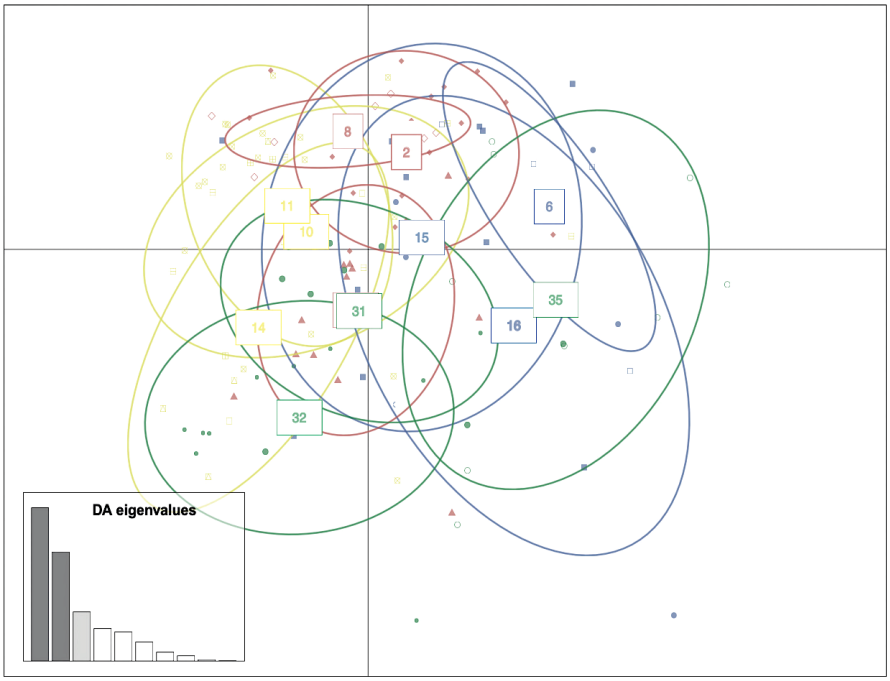
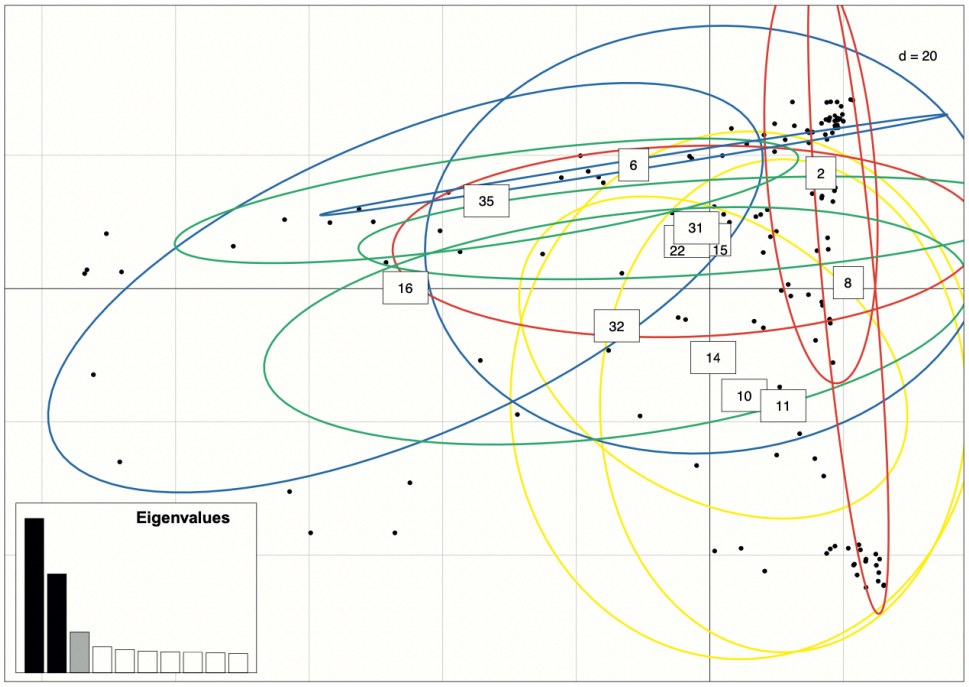
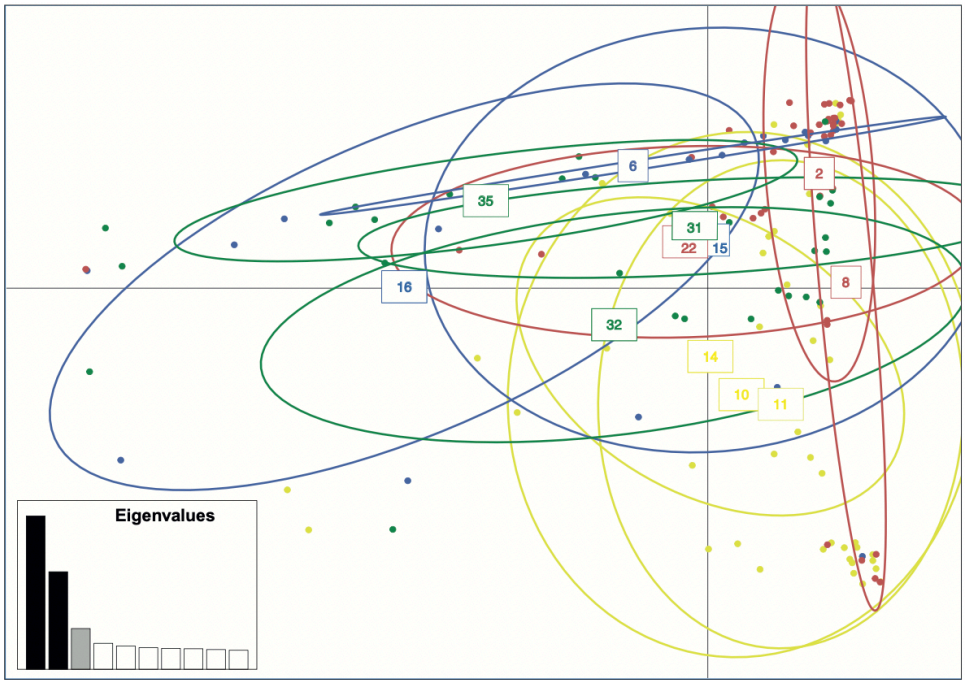


Figure 16. DAPCs (Discriminant Analysis of Principal Components) for *Epipedobates anthonyi* population structure **m3M5n5** matrix according to the distribution of the altitudes. A) DAPC from the 2775 SNPs matrix, B) DAPC from the 2706 neutral SNPs matrix and C) DAPC from the 69 putative outlier SNPs matrix. Each circle represents a different population, and each point is a sample. Transects are shown as red for transect one, blue for transect two, yellow for transect three, and green for lowlands. The x and y axes indicate the first and second principal components of DAPCP, respectively. The left inset shows a barplot for DA-eigenvalues (measure of the amount of variability retained).

A


d = 20
6
2
35
31
22
15
16
8
32
14
10
11
Eigenvalues

B


6
2
35
31
22
15
16
8
32
14
10
11
Eigenvalues

C

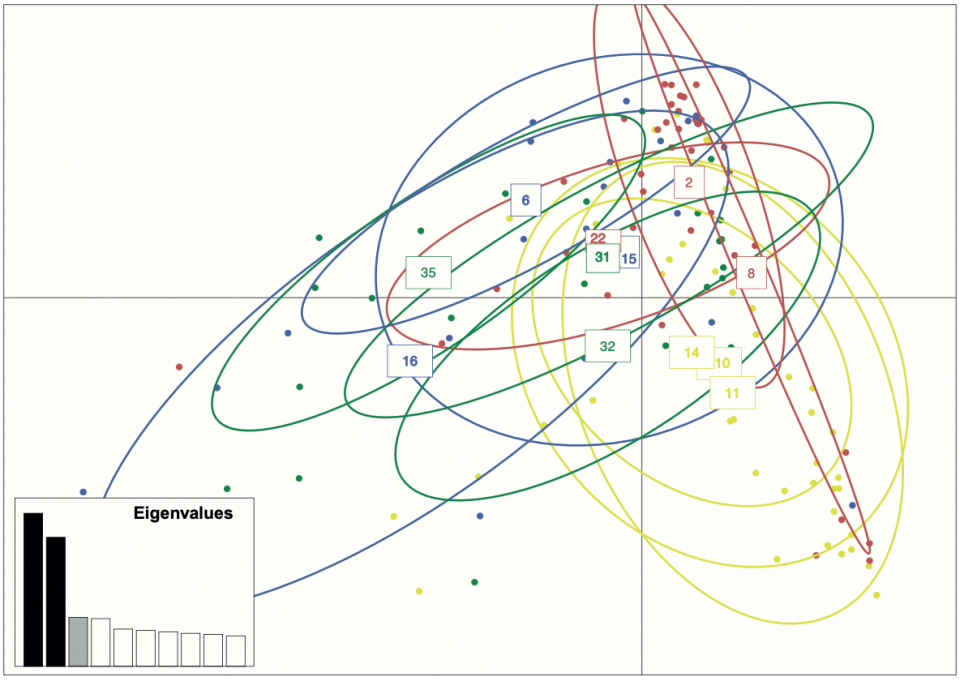
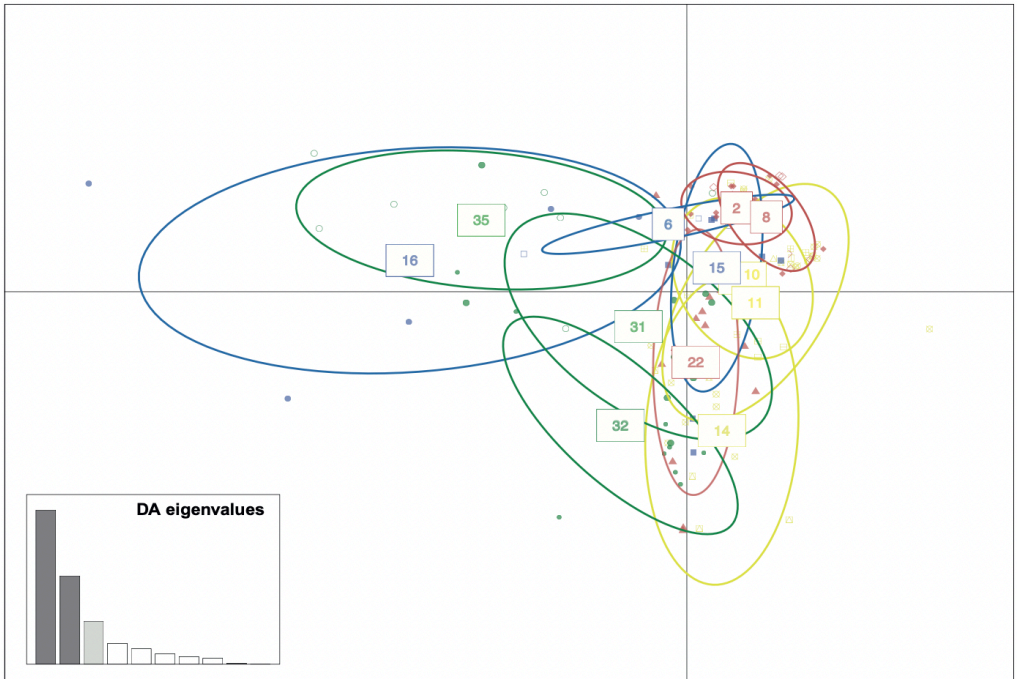


Figure 17. PCAs (Principal Component Analysis) for *Epipedobates anthonyi* population structure **m4M5n5** matrix according to the distribution of the transects. A) PCA from the 1758 SNPs matrix, B) PCA from the 1710 neutral SNPs matrix and C) PCA from the 48 putative outlier SNPs matrix. Each circle represents a different population, and each point is a sample. Transects are shown as red for transect one, blue for transect two, yellow for transect three, and green for lowlands. The x and y axes indicate the first and second principal components of PCA, respectively. The left inset shows a barplot for eigenvalues (measure of the amount of variability retained).

A


35
6
2
8
16
15
10
11
31
22
32
14
DA eigenvalues

B


35
6
2
8
16
15
10
11
31
22
32
14
DA eigenvalues

C

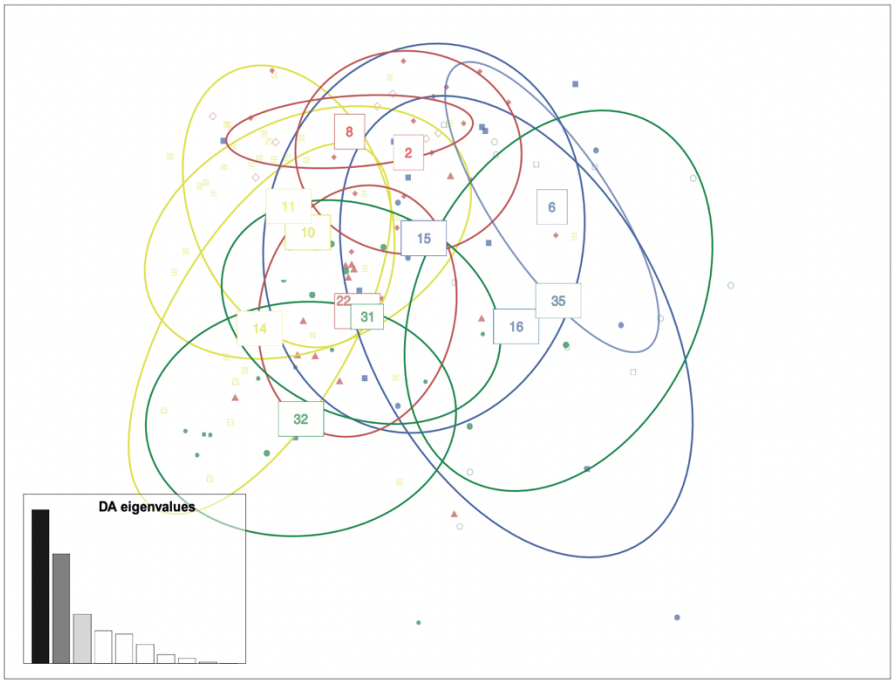
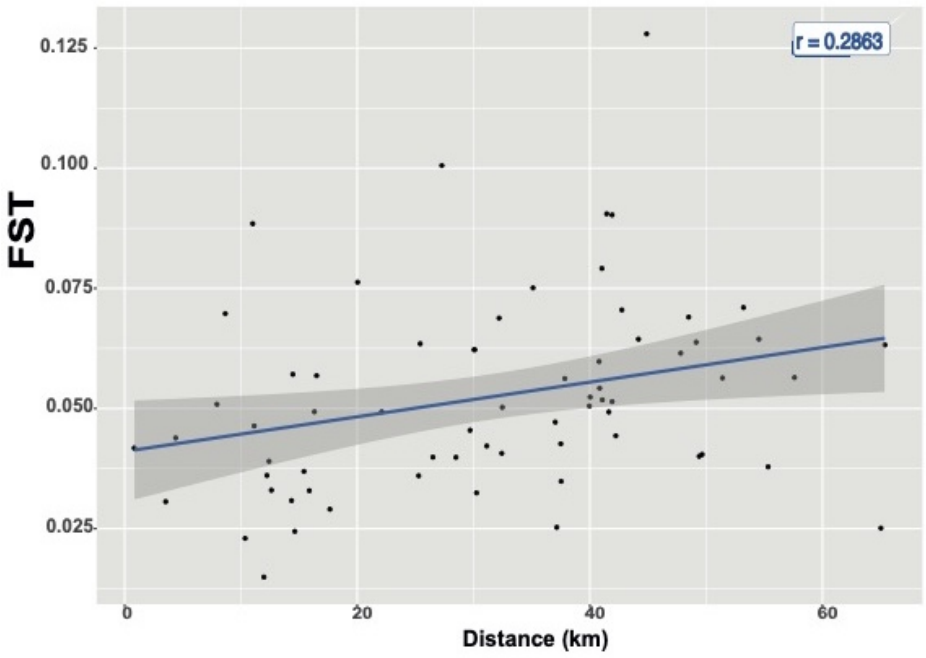


Figure 18. DAPCs (Discriminant Analysis of Principal Components) for *Epipedobates anthonyi* population structure **m4M5n5** matrix according to the distribution of the altitudes. A) DAPC from the 1758 SNPs matrix, B) DAPC from the 1710 neutral SNPs matrix and C) DAPC from the 48 putative outlier SNPs matrix. Each circle represents a different population, and each point is a sample. Transects are shown as red for transect one, blue for transect two, yellow for transect three, and green for lowlands. The x and y axes indicate the first and second principal components of DAPCP, respectively. The left inset shows a barplot for DA-eigenvalues (measure of the amount of variability retained).

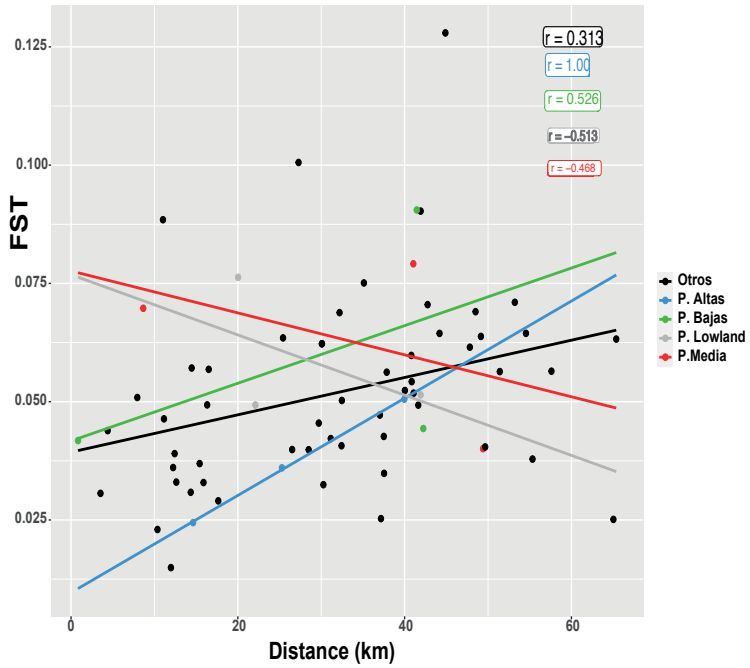
Isolation by distance

The results of our exploratory analysis based on linear regressions exhibit a genome-wide pattern of isolation by distance. *Epipedobates anthonyi* localities had overall genomic distance positively correlated with geographic distances (r = 0.286; P-value = 0.01881). However, when data were categorized by groups of altitudes only populations in high (r = 1; P-value =0.01686; Table 8) and low altitudes (r = 0.526; P-value = 0.6477) presented a positive correlation. Additionally, data were categorized by transects and transect 1 (r = 0.798; P-value = 0.4119) and transect 2 (r = 0.238; P-value = 0.8468) showed a positive correlation between geographic and genetic distances (Figure 19, Table 8).

A



B



C

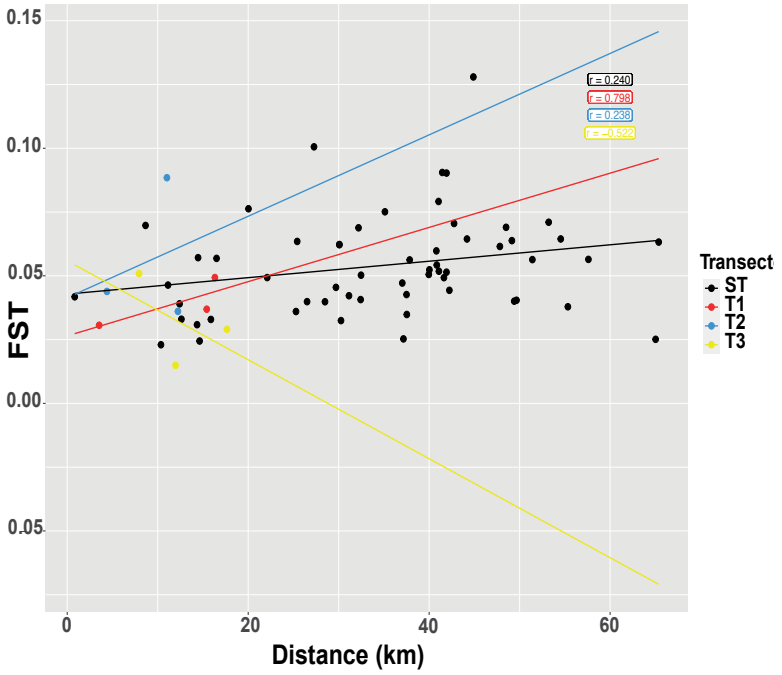


Figure 19. Regression plots of genetic differentiation (F_{ST}) and distance (km). A) General tendency of increasing F_{ST} per km, no data categorization, all populations included. B) Linear regressions according to altitude (blue = high altitudes; green = low altitudes; red = mid-altitude; gray = lowland). C) Linear regressions according to transects (red = transect 1; blue = transect 2; yellow = transect 3).

Table 8

Linear regression between genetic distance (F_{ST}) and geographic distance in populations of *Epipedobates anthonyi*. Categories refers to the comparisons made between populations.

Categories		n	R^2	R^2 adj.	r	p-value
	General (all populations)	67	0.082	0.067	0.286	0.018
	Same altitude only	12	0.032	-0.064	0.179	0.576
	Different altitude only	55	0.097	0.080	0.313	0.019
Altitudes	High-High only	3	0.999	0.998	1	0.016
	Low-Low only	3	0.276	-0.447	0.526	0.647
	Lowlands-Lowlands	3	0.263	-0.474	-0.513	0.657
	Mid-Mid only	3	0.218	-0.562	-0.468	0.690
	Within Transect 1	3	0.636	0.273	0.798	0.411
Transects	Within Transect 2	3	0.056	-0.886	0.238	0.846
	Within Transect 3	3	0.272	-0.454	-0.522	0.650

CHAPTER IV

DISCUSSION

The use of genomics provides new tools to address questions about speciation, conservation, and phenotypic evolution among others. One of these is the possibility of identifying putative loci under selection or detecting loci responsible for local-clinal adaptation (Allendorf, 2017; Savolainen et al. 2013). Genomic approaches have been made possible by the development of DNA sequencing technologies, making low-cost DNA sequencing possible, and increasing computational power. Next-generation sequencing (NGS) has changed the study of genetic variation and made it an accessible tool for the study of natural populations (Allendorf et al., 2022). However, the processing of the amounts of genetic information represents a challenge, that us researchers must overcome. Herein, we aimed to study genetic variation at the genome level of *Epipedobates anthonyi*, a species that exhibits phenotypic variation in color and size (Páez-Vacas et al., 2022). After following a set of filters, we obtained 2706 SNPs (matrix m3M5n5) and 1706 SNPs (matrix m4M5n5). Population structure in *Epipedobates anthonyi* showed that populations from mid and high putative populations cluster together, whereas populations from low elevation are more spread in the multivariate space. When we searched for outlier loci (putatively under selection) we did not find differences with neutral loci in population structure. This might indicate that this species is in the early stages of divergence, that evolutionary mechanisms that operate at the genome level (gene flow, genetic drift) could be explaining patterns of population genetic variation, or that, there is just a few loci behind the observed phenotypic variation that could not be detected with our analyses.

One of the first and most used analysis pipelines for genomic information is Stacks, which focuses on RAD data (Catchen et al., 2011). Stacks is designed to assemble loci based on short-read sequences obtained from restriction enzyme-based protocol (Catchen et al., 2011; Catchen, 2013 a; Rivera-Colón & Catchen, 2016). In this study, our raw data was almost 500 million total reads, this vast quantity of information should

be processed in a way to reduce the loss of a significant number of reads (Rivera-Colón & Catchen, 2016). By using *process_radtags* in Stacks, only 0.2% were low-quality reads, and 6.7% did not have a barcode, which leads to 91.17% of reads retained. Normally, it is expected to retain between 80-90 % of the total reads (Ali et al., 2016; Rivera-Colón & Catchen, 2016). When the amount of low-quality or no barcode found reads represent more than 20%, it is necessary to review the wet protocol used, whether the combination of restriction enzymes is the correct one, or if there is any problem with the sequence lane (Paris et al., 2017). Nonetheless, we did not encounter this problem.

De novo assembly is the pipeline used to analyze RADseq data without a reference genome, and optimizing its parameters (m, M, and n) to get loci is crucial. One of the metrics to choose a good combination of parameters after running *de novo* is the coverage of the samples. Ideally, coverage should be between 20 - 30x coverage (Rivera-Colón & Catchen, 2016). The processed data mapped 8.61x for m = 3 and 10.95x for m = 4. Considering that the study species is diploid, these values are 17.22x and 21.9x, meaning that each individual locus was observed in 17.22 and 21.9 independent reads. The -M value is related to the level of biological polymorphism present in a species (Rivera-Colón & Catchen, 2016), which is why we should avoid using a higher value when processing short reads as in the case of *E. anthonyi*. The simulation on many iterations of parameters will lend to different results of the number of original loci, removed loci, original sites, and remained sites (Table 5). Additionally, to optimize the highest number of polymorphic loci the r80 rule was applied to all iterations, this occurred when n = M and n = M-1 (Paris et al., 2017). Considering all these points, our best parameter iteration was m3M5n5 and m4M5n5 (Table 5, Figure 4), although the differences in the number of SNPs retained, near to one thousand, in the matrices, when putative under selection SNPs (outlier) were detected, there was not a significant difference in the number found (69 vs. 48).

Tools for estimating population structure from genetic data are now used in a wide variety of applications in population genetics. However, inferring population structure in large modern data sets imposes severe computational challenges (Raj et al., 2014). The amount of information obtained from many sampled individuals, genetic markers, and set of populations have become it difficult to describe population genetic pattern. As a hypothesis-free method, Principal Component Analysis (PCA) visualizes the data's inherent structure. PCA and Discriminant Analysis of Principal Components (DAPC) had been primarily used to study population structure (Dellinger et al., 2022), DAPC describes a model with genetic differences among a set of populations (Thia, 2022). In this study, twelve populations were defined a priori according to their geographical distribution. The population structure of *Epipedobates anthonyi* showed by PCAs and DAPCs (Figures 7 - 18) denotes how mid and high altitudes populations cluster together (Figures 11 - 14); furthermore, when data is categorized in transects (Figures 15 - 18), DAPCs maximize the differences among defined groups (Abegaz et al., 2019; Thia, 2022), and we identified the clustering of transects one and three.

Natural populations are shaped by their evolutionary history. It is crucial to interpret population structure in an evolutionary context (Kitada et al., 2021). One of the methods most used to differentiate between populations is F_{ST}, based on the proportion of the total genetic differentiation among groups of individuals (Wright, 1965). We found low levels of genetic differentiation between twelve populations using SNPs (F_{ST} mean = 0.05) which is not significative different from the value (F_{ST} mean = 0.04) found by Páez-Vacas et al. (2022) in the analysis based on microsatellites. Even so, *E. anthonyi* presents high phenotypic variation and the low genetic divergence shown is not necessarily inconsistent with extreme color divergence, because the color pattern is controlled by few genes (Koch & Behnecke, 2000). Showing that even at a genome-wide scale *E. anthonyi* has little divergence, even when SNPs can find possible signals of selection this study was not the case, our results were like neutral and putative loci under selection.

The genome-wide population-specific FST estimator can identify the source population and trace a possible evolutionary history of the derived populations based on genetic diversity inferring that populations closest to the ancestral population have the highest heterozygosity (Kitada et al., 2021). In our results, the low-altitude populations present the highest heterozygosity values (0.152) and the highest nucleotide diversity value (0.00181) those populations could be the first groups that eventually widespread and next to the dynamic pattern of multiple dispersals could have distributed until what we know as populations from high altitudes such as the endemic and diverse clade of poison frogs (Dendrobatidae) (Santos et al., 2009). Most species-level diversity originated after episodes of major Andean uplift in the Neogene period, suggest that dispersal and differentiation on a matrix previously shaped by large-scale landscape events (Smith et al., 2014). The aposematism in poison frog diversification has a central role including diet specialization, visual and acoustic communication, and reproductive biology all these aspects may facilitate the diversification of parental care in dendrobatids and might serve as an evolutionary platform while parental care can diversify as predation pressure is reduced (Carvajal-Castro et al., 2021).

Most of the studies have found high levels of genetic divergence among populations in adaptative divergence (Muir et al., 2014; Roland et al., 2017), however, we found little divergence genome-wide at neutral and putative loci under selection (F_{ST}). This could indicate these populations are in the early stages of population divergence. In addition, lineages with lower dispersal ability are expected to accrue genetic differentiation between populations at a relatively higher rate than more dispersive lineages, leading to a higher rate of speciation (Smith et al., 2014) as in the case of the high-altitude populations in *E. anthonyi*. Although putative loci under selection were identified, it was not possible to make inferences of adaptive selection along environmental gradients (elevation), which could depend on many other factors as well as how evolutive forces and evolutionary processes are acting on populations shaping diversity in natural systems (Clark et al., 2022; Savolainen et al, 2013; Shuichi et al.,

2021;). Consequently, F_{ST} outliers are only candidate loci of local adaptation and require further investigation to determine whether they are involved in this phenomenon.

Evolutionary theory predicts that divergent selection pressures across elevational gradients could cause adaptive divergence and reproductive isolation in the process of ecological speciation (Funk et al., 2015). The *Epipedobates* clade has shown differences in environmental characteristics of the habitat where they live, and altitude was identified as the most important environmental variable explaining the differences in the frequencies and note duration of the advertisement call of the clade (Baquero, 2012). Phenotypic divergence in *E. anthonyi* due to environmental differences at varying elevations (Páez-Vacas et al., 2022) indicates that greater isolation of high-elevation populations can reduce gene flow and increase the genetic divergence and reduce genetic variation (Funk et al., 2005) as our results showed in the exploratory analysis in linear regression evaluating the pattern of isolation by distance (IBD) (Figure 19) (Funk et al., 2005; Nali et al., 2020). Signals of adaptation in high-altitude amphibian populations have been reported (Yang et al., 2016), and few putative loci under selection could likely play a major role in the adaptation process in adults of *E. anthonyi*.

Our results underscore that altitude may be a factor shaping genetic structure among *Epipedobates anthonyi* populations. This has implications to understand the evolutionary trajectories of species in the tropical Andes, but also implications of their fate facing a changing world. We highlight that conservation strategies must focus on a holistic approach that includes not only the protection of the environment but also the protection of the genetic information found in species. High-elevation populations are genetically different from the low-elevation ones, showed less genetic variation, and therefore, might be more vulnerable to climate change. Even when *E. anthonyi* is a common poison frog species, this study shows that elevation influences its population structure. Other studies have pointed out that amphibians show sensitivity to

environmental changes in temperature, precipitation, and habitat transformation as future urbanization could result in fragmented populations (Pintanel et al., 2019; Trumbo et al., 2016). Thus, understanding how environmental variation along elevation affects evolutionary trajectories of natural populations can provide insights on how they will respond to future changes and how can we tackle their effects.

CHAPTER V

CONCLUSIONS

Epipedobates anthonyi populations exhibit a remarkable phenotypic variation in color and size, our results underscore that altitude may be a factor shaping its genetic structure accompanied with signals of isolation by distance that indicate that the greater isolation of high-elevation populations can increase genetic divergence and reduce genetic variation.

We obtained two matrices after applying pipelines and filters in Stacks, 2706 SNPs (matrix m3M5n5) and 1706 SNPs (matrix m4M5n5) based on the simulation on many iterations of the parameters -m, -M, and -n. These matrices produced almost 500 millions of original reads with coverage of 17.22x and 21.9x.

Population structure in *Epipedobates anthonyi* showed that populations from mid and high-altitude cluster together, whereas populations from low elevation are more spread in the multivariate space. Also, we were able to identify the clustering of transects one and three.

Putative under-selection SNPs (outlier) were detected: 69 for m3M5n5 matrix and 48 for the m4M5n5. We found little differences genome-wide at neutral and putative loci (F_{ST}); this could indicate these populations are in the early stages of population divergence, or that only a few loci that we were not able to detect with our analyses are responsible for phenotypic variation.

RECOMMENDATIONS

Working on genomic data brings challenging decisions when filtering massive amounts of information. Thus, we recommend that all the processes must be evaluated according to the metrics of each pipeline. If low-quality information represents more than 20% of the raw data it will be necessary to review the wet protocol used, whether if the combination of restriction enzymes is the correct one, or if there is any problem with the sequence lane. The analyses presented herein, are the initial steps in our datasets. We will perform Genotype-by-environment (GEA) and Genome-wide Association Analyses (GWAS) to further investigate the evolutionary processes in this species.

Understanding the processes that underline population divergence is important, yet there is scarce information about the roles of distance, patterns of phenotypic divergence, under-selection, and neutral genomic-wide information in amphibians. It is necessary to increase studies to a genome-wide level to obtain information about how amphibian populations are changing on that level.

Finally, the adaptability and resilience of wild populations, particularly those with restricted ranges and low genetic divergence, are threatened by rapid climate change. Hence, there is important to maintain conservation approaches that include not only the protection of the environment but also the protection of the genetic information found in this species.

REFERENCES

Abegaz, F., Chaichoompu, K., Génin, E., Fardo, D. W., König, I. R., Mahachie John, J. M. & Van Steen, K. (2019). Principals about principal components in statistical genetics. *Briefings in Bioinformatics*, *20*(6), 2200-2216. https://doi.org/10.1093/bib/bby081

Ali, O. A., O'Rourke, S. M., Amish, S. J., Meek, M. H., Luikart, G., Jeffres, C., & Miller, M. R. (2016). RAD capture (Rapture): flexible and efficient sequence-based genotyping. *Genetics*, *202*(2), 389-400. https://doi.org/10.1534/genetics.115.183665

Allendorf, F. W. (2017). Genetics and the conservation of natural populations: allozymes to genomes. *Molecular Ecology*, *26*(2), 420-430. https://doi.org/10.1111/mec.13948

Allendorf, F. W., & Luikart, G. (2009). Conservation and the genetics of populations. John Wiley & Sons.

Allendorf, F. W., Hohenlohe, P. A., and Luikart, G. (2010). Genomics and the future of conservation genetics. *Nature Reviews Genetics*, *11*(10), 697–709. https://doi.org/10.1038/nrg2844

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., Selker, E. U., Cresko, W. A. & Johnson, E. A. (2008). Rapid SNP Discovery and Genetic Mapping Using Sequenced RAD Markers. *PLoS ONE*, *3*(10), e3376. https://doi.org/10.1371/journal.pone.0003376

Baquero Rivadeneira, M. (2012). Relación entre el canto de anuncio y variables ecológicas y ambientales de especies del clado Epipedobates (Anura: Dendrobatidae) (*Bachelor's thesis*, Quito, 2012).

Black IV, W. C., Baer, C. F., Antolin, M. F., & DuTeau, N. M. (2001). Population Genomics: Genome-Wide Sampling of Insect Populations. *Annual Review of Entomology*, *46*(1), 441–469. https://doi.org/10.1146/annurev.ento.46.1.441

Bolnick, D. I. & Fitzpatrick, B. M. (2007). Sympatric Speciation: Models and Empirical Evidence. *Annual Review of Ecology, Evolution, and Systematics*, *38*(1), 459-487. https://doi.org/10.1146/annurev.ecolsys.38.091206.095804

Brumfield, R. T., Beerli, P., Nickerson, D. A. & Edwards, S. V. (2003). The utility of single nucleotide polymorphisms in inferences of population history. *Trends in Ecology & Evolution*, *18*(5), 249-256. https://doi.org/10.1016/s0169-5347(03)00018-1

Brummitt, N., & Lughadha, E. (2003). Biodiversity: where's hot and where's not. *Conservation Biology*, *17*(5), 1442-1448. http://www.jstor.org/stable/3588969

Cadena, C. D., Kozak, K. H., Gómez, J. P., Parra, J. L., McCain, C. M., Bowie, R. C. K., Carnaval, A. C., Moritz, C., Rahbek, C., Roberts, T. E., Sanders, N. J., Schneider, C. J., VanDerWal, J., Zamudio, K. R. & Graham, C. H. (2011). Latitude, elevational climatic zonation and speciation in New World vertebrates. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1726), 194-201. https://doi.org/10.1098/rspb.2011.0720

Carvajal-Castro, J. D., Vargas-Salinas, F., Casas-Cardona, S., Rojas, B. & Santos, J. C. (2021). Aposematism facilitates the diversification of parental care strategies in poison frogs. *Scientific Reports*, *11*(1). https://doi.org/10.1038/s41598-021-97206-6

Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W., and Postlethwait, J. H. (2011). Stacks: building and genotyping loci de novo from short-read sequences. *G3: Genes| genomes| genetics*, *1*(3), 171-182.

Catchen, J., Bassham, S., Wilson, T., Currey, M., O'Brien, C., Yeates, Q., and Cresko, W. A. (2013a). The population structure and recent colonization history of Oregon threespine stickleback determined using restriction-site associated DNA-sequencing. *Molecular Ecology*, *22*(11), 2864-2883. https://doi.org/10.1111/mec.12330

Catchen, J., Bassham, S., Wilson, T., Currey, M., O'Brien, C., Yeates, Q., & Cresko, W. A. (2013b). The population structure and recent colonization history of Oregon threespine stickleback determined using restriction-site associated DNA-sequencing. *Molecular Ecology*, *22*(11), 2864-2883.

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013c). Stacks: an analysis tool set for population genomics. *Molecular Ecology*, *22*(11), 3124–3140. https://doi.org/10.1111/mec.12354

Chazot, N., Willmott, K. R., Condamine, F. L., De-Silva, D. L., Freitas, A. V. L., Lamas, G., Morlon, H., Giraldo, C. E., Jiggins, C. D., Joron, M., Mallet, J., Uribe, S., & Elias, M. (2016). Into the Andes: multiple independent colonizations drive montane diversity in the Neotropical clearwing butterflies Godyridina. *Molecular Ecology*, *25*(22), 5765–5784. https://doi.org/10.1111/mec.13773

Clark, M. I., Bradburd, G. S., Akopyan, M., Vega, A., Rosenblum, E. B., & Robertson, J. M. (2022). Genetic isolation by distance underlies colour pattern divergence in red-eyed treefrogs (*Agalychnis callidryas*). *Molecular Ecology*, *31*(6), 1666-1681. https://doi.org/10.1111/mec.16350

Coloma, L. A., Frenkel, C. and Ron, S. R. (2018). Epipedobates anthonyi En: Ron, S. R., Merino-Viteri, A. Ortiz, D. A. (Eds). Anfibios del Ecuador. Version 2019.0. Museo de Zoología, Pontificia Universidad Católica del Ecuador. Recuperado el martes, 24 de noviembre de 2020 de https://bioweb.bio/faunaweb/amphibiaweb/FichaEspecie/Epipedobates%20anthonyi

Cortázar-Chinarro, M., Lattenkamp, E. Z., Meyer-Lucht, Y., Luquet, E., Laurila, A., & Höglund, J. (2017). Drift, selection, or migration? Processes affecting genetic differentiation and variation along a latitudinal gradient in an amphibian. *BMC Evolutionary Biology*, *17*(1), 604–611. https://doi.org/10.1186/s12862-017-1022-z

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G. & Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156-2158. https://doi.org/10.1093/bioinformatics/btr330

De Queiroz, K. (2007). Species Concepts and Species Delimitation. *Systematic Biology*, *56*(6), 879–886. https://doi.org/10.1080/10635150701701083

Dellinger, A. S., Paun, O., Baar, J., Temsch, E. M., Fernández-Fernández, D. & Schönenberger, J. (2022). Population structure in Neotropical plants: Integrating pollination biology, topography and climatic niches. *Molecular Ecology*, *31*(8), 2264-2280. https://doi.org/10.1111/mec.16403

Eguiarte, L. E., Aguirre-Liguori, J. A., Jardón-Barbolla, L., Aguirre-Planter, E. & Souza, V. (2013). Genómica de Poblaciones: Nada en Evolución va a tener sentido si no es a la luz de la Genómica y nada en Genómica tendrá sentido si no es a la luz de la Evolución. *TIP*, *16*(1), 42-56. https://doi.org/10.1016/s1405-888x(13)72077-1

Frost, Darrel R. (2021). Amphibian Species of the World: an Online Reference. Version 6.1 (*Date of access*). Electronic Database accessible at https://amphibiansoftheworld.amnh.org/index.php. American Museum of Natural History, New York, USA. doi.org/10.5531/db.vz.0001

Funk, W. C., Blouin, M. S., Corn, P. S., Maxell, B. A., Pilliod, D. S., Amish, S., & Allendorf, F. W. (2005). Population structure of Columbia spotted frogs (*Rana luteiventris*) is strongly affected by the landscape. *Molecular Ecology*, *14*(2), 483–496. https://doi.org/10.1111/j.1365-294x.2005.02426.x

Funk, W. C., McKay, J. K., Hohenlohe, P. A. and Allendorf, F. W. (2012). Harnessing genomics for delineating conservation units. *Trends in Ecology & Evolution*, *27*(9), 489-496. https://doi.org/10.1016/j.tree.2012.05.012

Funk, W. C., Murphy, M. A., Hoke, K. L., Muths, E., Amburgey, S. M., Lemmon, E. M. & Lemmon, A. R. (2015). Elevational speciation in action? Restricted gene flow associated with adaptive divergence across an altitudinal gradient. *Journal of Evolutionary Biology*, *29*(2), 241-252. https://doi.org/10.1111/jeb.12760

Ghalambor, C. K., Huey, R. B., Martin, P. R., Tewksbury, J. J., & Wang, G. (2006). Are mountain passes higher in the tropics? Janzen's hypothesis revisited. *Integrative and comparative biology*, *46*(1), 5-17. https://doi.org/10.1093/icb/icj003

Guarnizo, C. E. & Cannatella, D. C. (2013). Genetic divergence within frog species is greater in topographically more complex regions. *Journal of Zoological Systematics and Evolutionary Research*, *51*(4), 333-340. https://doi.org/10.1111/jzs.12027

Guo, B., Lu, D., Liao, W. B. & Merilä, J. (2016). Genomewide scan for adaptive differentiation along altitudinal gradient in the Andrew's toad *Bufo andrewsi*. *Molecular Ecology*, *25*(16), 3884-3900. https://doi.org/10.1111/mec.13722

Hartl, D. L. (2006). *Principles of Population Genetics* (4th Revised ed.). Sinauer Associates Is an Imprint of Oxford University Press.

Hutter, C. R., Lambert, S. M., & Wiens, J. J. (2017). Rapid Diversification and Time Explain Amphibian Richness at Different Scales in the Tropical Andes, Earth's Most Biodiverse Hotspot. *The American Naturalist*, *190*(6), 828–843. https://doi.org/10.1086/694319

Iturralde-Pólit, P., Dangles, O., Burneo, S. F., & Meynard, C. N. (2017). The effects of climate change on a mega-diverse country: predicted shifts in mammalian species richness and turnover in continental Ecuador. *Biotropica*, *49*(6), 821–831. https://doi.org/10.1111/btp.12467

Jombart, T. (2008). Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*(11), 1403– 1405. https://doi.org/10.1093/bioinformatics/btn129

Jombart, T., & Ahmed, I. (2011). Adegenet 1.3- 1: New tools for the analysis of genome-wide SNP data. *Bioinformatics*, *27*(21), 3070– 3071. https://doi.org/10.1093/bioinformatics/btr521

Jones, M. R., Forester, B. R., Teufel, A. I., Adams, R. V., Anstett, D. N., Goodrich, B. A., Landguth, E. L., Joost, S., & Manel, S. (2013). Integrating landscape genomics and spatially explicit approaches to detect loci under selection in clinal populations. *Evolution*, *67*(12), 3455–3468. https://doi.org/10.1111/evo.12237

Kitada, S., Nakamichi, R., & Kishino, H. (2021). Understanding population structure in an evolutionary context: population-specific F_{ST} and pairwise F_{ST}. *G3 Genes|Genomes|Genetics*, *11*(11). https://doi.org/10.1093/g3journal/jkab316

Koch, P. B., & Behnecke, B. (2000). The molecular basis of melanism and mimicry in a swallowtail butterfly. *Current Biology*, *10*(10), 591-594.

Luikart, G., England, P. R., Tallmon, D., Jordan, S., & Taberlet, P. (2003). The power and promise of population genomics: from genotyping to genome typing. *Nature Reviews Genetics*, *4*(12), 981–994. https://doi.org/10.1038/nrg1226

Luu, K., Bazin, E., & Blum, M. G. (2017). Pcadapt: an R package to perform genome scans for selection based on principal component analysis. *Molecular Ecology Resources*, *17*(1), 67-77. https://doi.org/10.1111/1755-0998.12592

Mastretta-Yanes, A., Arrigo, N., Alvarez, N., Jorgensen, T. H., Piñero, D., & Emerson, B. C. (2014). Restriction site-associated DNA sequencing, genotyping error estimation andde novoassembly optimization for population genetic inference. *Molecular Ecology Resources*, *15*(1), 28–41. https://doi.org/10.1111/1755-0998.12291

Mayr, E. (1970). Populations, species, and evolution: an abridgment of animal species and evolution (Vol. 19). Harvard University Press.

Muir, A. P., Biek, R., Thomas, R. & Mable, B. K. (2014). Local adaptation with high gene flow: temperature parameters drive adaptation to altitude in the common frog (*Rana temporaria*). *Molecular Ecology*, *23*(3), 561 - 574. https://doi.org/10.1111/mec.12624

Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A. B., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature*, *403*(6772), 853–858. https://doi.org/10.1038/35002501

Nali, R. C., Becker, C. G., Zamudio, K. R. & Prado, C. P. A. (2020). Topography, more than land cover, explains genetic diversity in a Neotropical savanna tree frog. *Diversity and Distributions*, *26*(12), 1798-1812. https://doi.org/10.1111/ddi.13154

Orme, C. D. L., Davies, R. G., Burgess, M., Eigenbrod, F., Pickup, N., Olson, V. A., Webster, A. J., Ding, T. S., Rasmussen, P. C., Ridgely, R. S., Stattersfield, A. J., Bennett, P. M., Blackburn, T. M., Gaston, K. J. & Owens, I. P. F. (2005). Global hotspots of species richness are not congruent with endemism or threat. *Nature*, *436*(7053), 1016-1019. https://doi.org/10.1038/nature03850

Páez-Vacas, M. I., Trumbo, D. R. & Funk, W. C. (2022). Contrasting environmental drivers of genetic and phenotypic divergence in an Andean poison frog (*Epipedobates anthonyi*). *Heredity*, *128*(1), 1-12. https://doi.org/10.1038/s41437-021-00481-2

Paris, J. R., Stevens, J. R. & Catchen, J. M. (2017). Lost in parameter space: a road map for stacks. *Methods in Ecology and Evolution*, *8*(10), 1360-1373. https://doi.org/10.1111/2041-210x.12775

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S. & Hoekstra, H. E. (2012). Double Digest RADseq: An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and Non-Model Species. *PLoS ONE*, *7*(5), e37135. https://doi.org/10.1371/journal.pone.0037135

Pintanel, P., Tejedo, M., Ron, S. R., Llorente, G. A. & Merino-Viteri, A. (2019). Elevational and microclimatic drivers of thermal tolerance in Andean *Pristimantis* frogs. *Journal of Biogeography*, *46*(8), 1664-1675. https://doi.org/10.1111/jbi.13596

Raj, A., Stephens, M. & Pritchard, J. K. (2014). fastSTRUCTURE: Variational Inference of Population Structure in Large SNP Data Sets. *Genetics*, *197*(2), 573-589. https://doi.org/10.1534/genetics.114.164350

Richards, E. J., Servedio, M. R., & Martin, C. H. (2019). Searching for Sympatric Speciation in the Genomic Era. *BioEssays*, *41*(7), 1900047. https://doi.org/10.1002/bies.201900047

Roesti, M., Hendry, A. P., Salzburger, W., & Bender, D. (2012). Genome divergence during evolutionary diversification as revealed in replicate lake-stream stickleback population pairs. *Molecular Ecology*, *21*(12), 2852–2862. https://doi.org/10.1111/j.1365-294x.2012.05509.x

Roland, A. B., Santos, J. C., Carriker, B. C., Caty, S. N., Tapia, E. E., Coloma, L. A. & O'Connell, L. A. (2017). Radiation of the polymorphic Little Devil poison frog (*Oophaga sylvatica*) in Ecuador. *Ecology and Evolution*, *7*(22), 9750-9762. https://doi.org/10.1002/ece3.3503

Ron, S. R., Merino-Viteri & Ortiz, D. A. 2019. Anfibios del Ecuador. Version 2019.0. Museo de Zoología, Pontificia Universidad Católica del Ecuador. Recuperado el martes, 24 de marzo de 2021 de https://bioweb.bio/faunaweb/amphibiaweb>

Salazar-Valenzuela, D., Kuch, U., Torres-Carvajal, O., Valencia, J. H. & Gibbs, H. L. (2019). Divergence of tropical pitvipers promoted by independent colonization

events of dry montane Andean habitats. *Journal of Biogeography*, *46*(8), 1826-1840. https://doi.org/10.1111/jbi.13661

Santos, J. C., Coloma, L. A., Summers, K., Caldwell, J. P., Ree, R. & Cannatella, D. C. (2009). Amazonian Amphibian Diversity Is Primarily Derived from Late Miocene Andean Lineages. *PLoS Biology*, *7*(3), e1000056. https://doi.org/10.1371/journal.pbio.1000056

Santos, J. C., Baquero, M., Barrio-Amorós, C., Coloma, L. A., Erdtmann, L. K., Lima, A. P. & Cannatella, D. C. (2014). Aposematism increases acoustic diversification and speciation in poison frogs. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1796), 20141761. https://doi.org/10.1098/rspb.2014.1761

Savolainen, O., Lascoux, M. & Merilä, J. (2013). Ecological genomics of local adaptation. *Nature Reviews Genetics*, *14*(11), 807-820. https://doi.org/10.1038/nrg3522

Smith, B. T., McCormack, J. E., Cuervo, A. M., Hickerson, M. J., Aleixo, A., Cadena, C. D., Pérez-Emán, J., Burney, C. W., Xie, X., Harvey, M. G., Faircloth, B. C., Glenn, T. C., Derryberry, E. P., Prejean, J., Fields, S. & Brumfield, R. T. (2014). The drivers of tropical speciation. *Nature*, *515*(7527), 406-409. https://doi.org/10.1038/nature13687

Schlötterer, C. (2004). The evolution of molecular markers — just a matter of fashion? *Nature Reviews Genetics*, *5*(1), 63-69. https://doi.org/10.1038/nrg1249

Schluter, D. (2000). *The ecology of adaptive radiation*. OUP Oxford.

Schluter, D. (2001). Ecology and the origin of species. *Trends in Ecology & Evolution*, *16*(7), 372-380. https://doi.org/10.1016/s0169-5347(01)02198-x

Shaffer, H. B., Gidiş, M., McCartney-Melstad, E., Neal, K. M., Oyamaguchi, H. M., Tellez, M. & Toffelmier, E. M. (2015). Conservation Genetics and Genomics of Amphibians and Reptiles. *Annual Review of Animal Biosciences*, *3*(1), 113-138. https://doi.org/10.1146/annurev-animal-022114-110920

Tarvin, R. D., Powell, E. A., Santos, J. C., Ron, S. R., & Cannatella, D. C. (2017). The birth of aposematism: High phenotypic divergence and low genetic diversity in

a young clade of poison frogs. *Molecular Phylogenetics and Evolution*, *109*, 283-295. https://doi.org/10.1016/j.ympev.2016.12.035

Thia, J. A. (2022). Guidelines for standardizing the application of discriminant analysis of principal components to genotype data. *Molecular Ecology Resources*. https://doi.org/10.1111/1755-0998.13706

Trumbo, D. R., Epstein, B., Hohenlohe, P. A., Alford, R. A., Schwarzkopf, L. & Storfer, A. (2016). Mixed population genomics support for the central marginal hypothesis across the invasive range of the cane toad (*Rhinella marina*) in Australia. *Molecular Ecology*, *25*(17), 4161-4176. https://doi.org/10.1111/mec.13754

Trumbo, D. R., Salerno, P. E., Logan, K. A., Alldredge, M. W., Gagne, R. B., Kozakiewicz, C. P., Kraberger, S., Fountain-Jones, N. M., Craft, M. E., Carver, S., Ernest, H. B., Crooks, K. R., VandeWoude, S. & Funk, W. C. (2019). Urbanization impacts apex predator gene flow but not genetic diversity across an urban-rural divide. *Molecular Ecology*, *28*(22), 4926-4940. https://doi.org/10.1111/mec.15261

Van Buskirk, J., & Jansen van Rensburg, A. (2020). Relative importance of isolation-by-environment and other determinants of gene flow in an alpine amphibian. *Evolution*, *74*(5), 962–978. https://doi.org/10.1111/evo.13955

Whitlock, M. C., & Lotterhos, K. E. (2015). Reliable detection of loci responsible for local adaptation: Inference of a null model through trimming the distribution of F ST. *The American Naturalist*, *186*(S1), S24-S36. https://doi.org/10.1086/682949

Wollenberg Valero, K. C., Marshall, J. C., Bastiaans, E., Caccone, A., Camargo, A., Morando, M., Niemiller, M. L., Pabijan, M., Russello, M. A., Sinervo, B., Werneck, F. P., Sites, J. W., Wiens, J. J., and Steinfartz, S. (2019). Patterns, Mechanisms and Genetics of Speciation in Reptiles and Amphibians. *Genes*, *10*(9), 646. https://doi.org/10.3390/genes10090646

Yang, W., Qi, Y., & Fu, J. (2016). Genetic signals of high-altitude adaptation in amphibians: a comparative transcriptome analysis. *BMC Genetics*, 17(1). https://doi.org/10.1186/s12863-016-0440-z

SUPPLEMENTARY INFORMATION

Table S1.
Pairwise population genetic differentiation, fixation index (F_{ST}).

						m3M5n5						
	EA10	**EA11**	**EA14**	**EA15**	**EA16**	**EA22**	**EA2**	**EA31**	**EA32**	**EA35**	**EA6**	**EA8**
EA10	-	**0.0149223**	0.0508808	0.0471733	0.0523646	0.0542347	0.0563526	0.0644376	0.0562336	0.0750967	0.0791441	0.0400703
EA11	-	-	0.0290395	0.0360132	0.0324504	0.0421958	0.0505249	0.0426556	0.0348444	0.0492687	0.0454775	0.0252761
EA14	-	-	-	0.0597782	0.0905157	0.0443216	0.0644492	0.0705119	0.0502522	0.100557	**0.127945**	0.0710105
EA15	-	-	-	-	0.0360751	0.0329985	0.0244421	0.0398757	0.0518065	0.0378697	0.0438747	0.0390217
EA16	-	-	-	-	-	0.0417744	0.0329064	0.0571152	0.0688195	0.0637926	0.0884605	0.0568358
Ea22	-	-	-	-	-	-	0.0369144	0.0308453	0.0406819	0.0404325	0.0463714	0.0493189
EA2	-	-	-	-	-	-	-	0.0398356	0.0615117	0.0251154	0.0229818	0.0306255
EA31	-	-	-	-	-	-	-	-	0.0493052	0.0514332	0.0634823	0.0622375
EA32	-	-	-	-	-	-	-	-	-	0.0762859	0.0902798	0.0690194
EA35	-	-	-	-	-	-	-	-	-	-	0.0564427	0.0632317
EA6	-	-	-	-	-	-	-	-	-	-	-	0.0697452
						m4M5n5						
	EA10	**EA11**	**EA14**	**EA15**	**EA16**	**EA22**	**EA2**	**EA31**	**EA32**	**EA35**	**EA6**	**EA8**
EA10	-	**0.0149357**	0.0522061	0.0464024	0.0596273	0.055409	0.0540899	0.0648434	0.0556104	0.0718113	0.0733618	0.0394977
EA11	-	-	0.0285028	0.0339125	0.0332606	0.0413285	0.0477464	0.041548	0.0326058	0.0448907	0.0405882	0.0254356
EA14	-	-	-	0.0602467	0.103232	0.0434187	0.0627126	0.0692398	0.054383	0.0978657	**0.124082**	0.0729981
EA15	-	-	-	-	0.0387525	0.0364698	0.0246688	0.0431361	0.0507607	0.0377868	0.0452882	0.0390261
EA16	-	-	-	-	-	0.0445146	0.0314168	0.0641412	0.0738526	0.0639391	0.103401	0.0629739
Ea22	-	-	-	-	-	-	0.0401096	0.0319587	0.0394254	0.0423031	0.0509798	0.0521185
EA2	-	-	-	-	-	-	-	0.0419156	0.0580247	0.0257314	0.023437	0.0291143
EA31	-	-	-	-	-	-	-	-	0.0513379	0.0550263	0.069865	0.0639487

EA32	-	-	-	-	-	-	-	-	-	0.0758774	0.0897577	0.0663272
EA35	-	-	-	-	-	-	-	-	-	-	0.0645383	0.0614648
EA6	-	-	-	-	-	-	-	-	-	-	-	0.0650574

Table S2.
Pairwise population genetic differentiation, fixation index (F_{ST}) in the outlier SNP and the neutral SNP matrices.

m3M5n5 - 69 (outlier SNPs)												
	EA10	**EA11**	**EA14**	**EA15**	**EA16**	**EA22**	**EA2**	**EA31**	**EA32**	**EA35**	**EA6**	**EA8**
EA10	-	0.00245515	0.00283304	0.00287791	0.00299713	0.00259496	0.00277814	0.00271555	0.00265914	0.00308385	0.00336026	0.00293553
EA11	-	-	0.00266156	0.00274795	0.00262648	0.00265469	0.00266762	0.00256719	0.00247155	0.00292187	0.0029234	0.00256184
EA14	-	-	-	0.00298748	0.00316447	0.00288885	0.00274776	0.00296168	0.00297266	0.00299449	0.0034304	0.00305221
EA15	-	-	-	-	0.00290087	0.00250744	0.00244511	0.00262383	0.00262666	0.00274271	0.00282219	0.00284014
EA16	-	-	-	-	-	0.00266357	0.00255689	0.00285927	0.00288063	0.0031052	0.0036486	0.0029037
EA22	-	-	-	-	-	-	0.00238922	0.00268591	0.00264944	0.00255612	0.0027214	0.00258891
EA2	-	-	-	-	-	-	-	0.00232901	0.0025221	0.00245308	0.0025122	0.00254336
EA31	-	-	-	-	-	-	-	-	0.00264422	0.00270348	0.00306218	0.00257653
EA32	-	-	-	-	-	-	-	-	-	0.00282185	0.00310641	0.00273746
EA35	-	-	-	-	-	-	-	-	-	-	0.00310422	0.00283253
EA6	-	-	-	-	-	-	-	-	-	-	-	0.00313884
m3M5n5 - 2076 (neutral SNPs)												
	EA10	**EA11**	**EA14**	**EA15**	**EA16**	**EA22**	**EA2**	**EA31**	**EA32**	**EA35**	**EA6**	**EA8**
EA10	-	0.000814627	0.0293823	0.0389364	0.0299064	0.0554481	0.0687241	0.0642739	0.0509923	0.0796182	0.078748	0.0218115
EA11	-	-	0.025357	0.0448453	0.0329394	0.0558315	0.0689807	0.0605446	0.0431161	0.0740006	0.0721456	0.0219001
EA14	-	-	-	0.0409195	0.0422559	0.0237719	0.0902981	0.0496238	0.0132652	0.0897846	0.0884997	0.0492678
EA15	-	-	-	-	-0.0090081	0.0177921	0.0110412	0.0188537	0.0403972	0.0114824	0.000288334	0.0171478
EA16	-	-	-	-	-	0.0155652	0.0164251	0.0226712	0.0455782	0.0257061	0.0227842	0.0215089

EA22	-	-	-	-	-	-	0.0372041	0.0093125	0.0280519	0.0240936	0.0192597	0.0436092
EA2	-	-	-	-	-	-	-	0.0388322	0.0778056	0.00695446	-0.0167376	0.0210419
EA31	-	-	-	-	-	-	-	-	0.0304555	0.027533	0.0252633	0.0511411
EA32	-	-	-	-	-	-	-	-	-	0.0707535	0.0749938	0.0636113
EA35	-	-	-	-	-	-	-	-	-	-	0.000971947	0.0487334
EA6	-	-	-	-	-	-	-	-	-	-	-	0.0385695

Table S3.

Dataset of genetic differentiation F_{ST} and topographic distance (km) between populations (x-y) of *Epipedobates anthonyi*.

x-y	F_{ST} (m3M5n5)	TopoDist (km)
EA8-EA2	0.030625	3529.040388
EA22-EA2	0.036914	15419.23847
EA15-EA2	0.024442	14632.78763
EA6-EA2	0.022982	10360.99856
EA16-EA2	0.032906	15880.64665
EA31-EA2	0.039836	28490.79045
EA32-EA2	0.061512	47818.60985
EA35-EA2	0.025115	65031.85751
EA14-EA2	0.064449	54544.27471
EA10-EA2	0.056353	51409.8733
EA11-EA2	0.050525	39963.0615
EA22-EA8	0.049319	16321.59289
EA15-EA8	0.039022	12418.8524
EA6-EA8	0.069745	8657.172871
EA16-EA8	0.056836	16506.50074
EA31-EA8	0.062238	30079.00476
EA32-EA8	0.069019	48502.85599
EA35-EA8	0.063232	65373.4566
EA14-EA8	0.07101	53212.75122
EA10-EA8	0.04007	49404.0674
EA11-EA8	0.025276	37160.90503
EA15-EA22	0.032999	12620.69609
EA6-EA22	0.046371	11139.3753
EA16-EA22	0.041774	830.7756776
EA31-EA22	0.030845	14352.05493
EA32-EA22	0.040682	32430.13722
EA35-EA22	0.040433	49658.08662
EA14-EA22	0.044322	42236.84437
EA10-EA22	0.054235	40845.05028
EA11-EA22	0.042196	31141.62318
EA6-EA15	0.043875	4388.689435
EA16-EA15	0.036075	12234.87733
EA31-EA15	0.039876	26508.95817
EA32-EA15	0.051807	41079.0618

EA35-EA15	0.03787	55334.31295
EA14-EA15	0.059778	40799.63683
EA10-EA15	0.047173	37031.78213
EA11-EA15	0.036013	25285.40234
EA16-EA6	0.08846	11018.91856
EA31-EA6	0.063482	25417.85601
EA32-EA6	0.09028	41920.47212
EA35-EA6	0.056443	57604.0354
EA14-EA6	0.12795	44887.66531
EA10-EA6	0.079144	41040.67158
EA11-EA6	0.045477	29702.8304
EA31-EA16	0.057115	14462.91024
EA32-EA16	0.06882	32196.87711
EA35-EA16	0.063793	49145.58221
EA14-EA16	0.090516	41450.24702
EA10-EA16	0.052365	40029.22925
EA11-EA16	0.03245	30259.0882
EA32-EA31	0.049305	22104.56199
EA35-EA31	0.051433	41918.86016
EA14-EA31	0.070512	42755.08912
EA10-EA31	0.064438	44184.63581
EA11-EA31	0.042656	37500.51813
EA35-EA32	0.076286	20032.71353
EA14-EA32	0.050252	32473.94927
EA10-EA32	0.056234	37851.44333
EA11-EA32	0.034844	37532.3398
EA14-EA35	0.10056	27275.60504
EA10-EA35	0.075097	35115.7528
EA11-EA35	0.049269	41635.24668
EA10-EA14	0.050881	7941.32216
EA11-EA14	0.029039	17650.92826
EA11-EA10	0.014922	11974.16712
EA31-EA8	0.062238	30079.00476