



**CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS
AVANZADOS DEL INSTITUTO POLITÉCNICO
NACIONAL**

UNIDAD IRAPUATO

UNIDAD DE GENÓMICA AVANZADA

**Especiación de salamandras del género *Pseudoeurycea* en el Eje
Neovolcánico de México**

Tesis que presenta

María Guadalupe Segovia Ramírez

Para Obtener el Grado de

Maestro en Ciencias

En la Especialidad de Biotecnología de Plantas

Director de la Tesis

Dr. Sean Michael Rovito

Irapuato, Guanajuato

Agosto, 2017

EL presente trabajo fue elaborado en el Laboratorio de Genómica y Biodiversidad de Vertebrados en la Unidad de Genómica Avanzada, Cinvestav Unidad Irapuato bajo la dirección del Dr. Sean M. Rovito en colaboración con El Dr. Ke Bi (Computational Genomics Resource Laboratory (CGRL), California Institute for Quantitative Biosciences (QB3), University of California, Berkeley, Berkeley) y fungiendo como sinodales los doctores Angélica Cibrian Jaramillo y Martin Heil adscritos a la Unidad de Genómica Avanzada y Departamento de Ingeniería Genética respectivamente en Cinvestav Irapuato. Agradezco al Consejo Nacional de Ciencia y Tecnología (CONACyT) por la beca de maestría número 396123 otorgada en el periodo de septiembre del 2014 a agosto del 2016.

Agradecimientos

Al Consejo Nacional de Ciencia y Tecnología (CONACyT) por la beca de maestría número 396123 otorgada en el periodo de septiembre del 2014 a agosto del 2016.

Al pueblo mexicano que responsablemente paga sus impuestos porque de ahí viene la beca otorgada por CONACyT que me sostuvo

Al Dr. Sean M. Rovito por aceptar ser mi asesor durante este largo tiempo, dirigir este hermoso proyecto y por enseñarme a diario como ser un buen profesionista y persona

Al grupo de trabajo del laboratorio de Genómica y Biodiversidad de Vertebrados a Cargo del Dr. Rovito

A los miembros del comité sinodal Dra. Angélica Cibrián Jaramillo y Dr. Martin Heil por asesorarme y guiarme durante la realización de este proyecto

Al Dr. Ke Bi por su colaboración y ayuda en algunos aspectos bioinformáticos

Al Dr. Cei Abreu por el apoyo incondicional y charlas sobre el proyecto

A Obed Ramírez, Roberto Bermúdez, Alejandro Álvarez, Pablo de la Rosa y Araceli Fernández por la ayuda y asesoría bioinformática

A mis amigos (los viejos y los nuevos)

A la naturaleza que me ha permitido estar aquí y admirarla

Especialmente a mi hijo Itzamná por ser mi motor desde hace 4 años

Dedicatoria

A mi hijo Itzamná y mi familia
A mis amigos
A todas las personas honestas, trabajadoras, conscientes y justas de mi país

ÍNDICE GENERAL

SIGLAS Y ABREVIATURAS	9
RESUMEN	10
ABSTRACT	11
INTRODUCCIÓN	12
MARCO TEÓRICO	13
1. GEOGRAFÍA DE LA ESPECIACIÓN	13
1.1. <i>Especiación parapátrica y simpátrica</i>	13
1.2. <i>Especiación alopátrica</i>	15
2. ESTUDIO DE LA ESPECIACIÓN GEOGRÁFICA.....	17
2.1. <i>Estudios clásicos de especiación geográfica</i>	17
2.2. <i>Estudios recientes de especiación geográfica</i>	18
3. EJE NEOVOLCÁNICO MEXICANO COMO ESCENARIO PARA LA ESPECIACIÓN.....	20
3.1. <i>Estudio de la especiación geográfica en el Eje Neovolcánico</i>	21
4. PLETHODONTIDAE	21
4.1 <i>Diversidad de Plethodontidae: Pseudoeurycea en el ENV</i>	23
HIPÓTESIS	25
OBJETIVO GENERAL	25
OBJETIVOS PARTICULARES	25
MÉTODO	26
1. MUESTREO POBLACIONAL	26
2. EXTRACCIÓN DE ADN Y AMPLIFICACIÓN DE CYTB	26
3. EXTRACCIÓN DE ARN, PREPARACIÓN DE LIBRERÍAS DE ADNC Y SECUENCIACIÓN	26
4. PROCESAMIENTO DE DATOS CRUDOS (RNA-SEQ)	27
4.1 <i>Análisis de calidad, limpieza, ensamblado y anotación de transcriptomas</i>	27
4.2 <i>Mapeo y llamado de genotipos</i>	28
4.3 <i>Edición y selección de ortólogos</i>	28
5. DISEÑO DE MODELOS DEMOGRÁFICOS DE ESPECIACIÓN	29
5.1 <i>Generación de árbol filogenético con cytb</i>	30
6. PRUEBA DE HIPÓTESIS	30
6.1 <i>Clado Veracruz</i>	30
RESULTADOS	35
1. CLADO EDOMEX.....	36
2. CLADO VERACRUZ	37
2.1 <i>Modelo más probable de especiación</i>	38
DISCUSIÓN	42
1. ESPECIACIÓN GEOGRÁFICA DEL CLADO VERACRUZ.....	42
3. DATOS MASIVOS COMO FUENTE DE MARCADORES MOLECULARES	46
4. OBTENCIÓN DE SNPs Y ORTÓLOGOS	47
CONCLUSIONES	48
PERSPECTIVAS	49
LITERATURA CITADA	50

ÍNDICE DE CUADROS

CUADRO 1: DESIGNACIÓN DE PRIORES PARA CADA MODELO DE ESPECIACIÓN CORRESPONDIENTE AL CLADO DE VERACRUZ.....	30
Cuadro 2: ESTIMACIÓN DE LA TASA DE SUBSTITUCION.....	33
CUADRO 3: NÚMERO PROMEDIO DE TRASNCRITOS ENSAMBLADOS (NO ANOTADOS Y ANOTADOS) POR ESPECIE. ...	36
CUADRO 4: NÚMERO DE LOCI CON SITIOS SEGREGANTES (Y NÚMERO SITIOS SEGREGANTES) EN CADA POBLACIÓN Y CON SITIOS SEGREGANTES EN TODAS LAS POBLACIONES PARA EL CLADO EDOMEX.....	36
CUADRO 5: MEDIA Y VARIANZA DE LOS ESTADÍSTICOS DE RESUMEN OBSERVADOS EN EL CLADO EDOMEX CONSIDERANDO 20 LOCI CON SITIOS SIGREGANTES EN TODAS LAS POBLACIONES.....	37
CUADRO 6: ESTADÍSTICOS DE RESUMEN CORRESPONDIENTES A LOS DATOS OBSERVADOS EMPLEADOS EN LA COMPARACIÓN ENTRE MODELOS.....	38
CUADRO 7: MATRIZ DE CONFUSIÓN COMO RESULTADO DE LA VALIDACIÓN CRUZADA.....	38
CUADRO 8: PARÁMETROS POSTERIORES CORRESPONDIENTES AL MODELO ALOPÁTRICO.....	41

ÍNDICE DE FIGURAS

FIGURA 1: FILOGENIA DEL GÉNERO <i>PSEUDOEURYCEA</i>	23
FIGURA 2: DISTRIBUCIÓN DE SALAMANDRAS PLETODÓNTIDAS EN EL CENTRO-ESTE DEL EJE NEOVOLCÁNICO.....	24
FIGURA 3: REFORMATO DE LOS GENOTIPOS RESULTANTES DE ANGSD AL FORMATO MSABC.....	29
FIGURA 4: ESQUEMA DE LOS MODELOS DEMOGRÁFICOS DEL CLADO VERACRUZ.....	31
FIGURA 5: ÁRBOL FILOGENÉTICO ESTIMADO CON CYTB CORRESPONDIENTE A LAS CINCO ESPECIES MICROENDÉMICAS DEL TMVB.....	35
FIGURA 6: DENSIDAD POSTERIOR	40
FIGURA 7: DENSIDAD POSTERIOR (LÍNEA CONTINUA) DEL TIEMPO DE DIVERGENCIA ENTRE <i>P. MELANOMOLGA</i> Y <i>P. GADOVII</i>	41

Siglas y Abreviaturas

ABC	Approximate Bayesian computation (por su nombre y siglas en inglés)
ADN	Ácido desoxirribonucleico
ADNc	Ácido desoxirribonucleico complementario
ARN	Ácido ribonucleico
ca.	cerca
CdPV	Cofre de Perote Volcano (por su nombre y siglas en inglés)
COI	cytochrome c oxidase subunit I (por su nombre y siglas en inglés)
COII	cytochrome c oxidase subunit II (por su nombre y siglas en inglés)
CVF	Chichinautzin Volcanic Field (por su nombre y siglas en inglés)
cytb	citocromo b
ENV	Eje Neovolcánico
Gb	Gigabase
Ka	mil años
Kb	Kilobase
m	Tasa de migracion
Ma	Millones de años
ND1	NADH-ubiquinone oxidoreductase chain 1 (por su nombre y siglas en inglés)
NdTV	Nevado de Toluca Volcano (por su nombre y siglas en inglés)
N_e	Tamaño efectivo poblacional
NGS	Next Generation Sequencing (por su nombre y siglas en inglés)
N_o	Tamaño poblacional
p.e	Por ejemplo
pb	Pares de bases
POMC	proopiomelanocortin (por su nombre y siglas en inglés)
RAG1	recombination activating gene 1 (por su nombre y siglas en inglés)
S	Estadísticos de resumen
SLC8A3	solute carrier family 8 member 3 (por su nombre y siglas en inglés)
SMOc	Sierra Madre Occidental
SMOr	Sierra Madre Oriental
SMR	Sean Michael Rovito
SMS	Sierra Madre del Sur
SNPs	Single nuclear polymorphisms (por su nombre y siglas en inglés)
S_{obs}	Estadísticos de resumen observados
S_{sims}	Estadísticos de resumen simulados
TMVB	Trans-Mexican Volcanic Belt
tol	tolerancia
UMG	Último máximo glacial
ZVF	Zitacuaro Volcanic Field (por su nombre y siglas en inglés)

Resumen

Las características de las poblaciones actuales son el resultado de diferentes eventos demográficos ocurridos en el pasado. Cada población almacena en el ADN huellas sobre su pasado demográfico y evolutivo y cada gen cuenta su propia historia. Si pretendemos elucidar la historia demográfica de una población y la historia de la formación de especies, tenemos que buscar estas huellas en su acervo génico en un contexto multilocus. Actualmente la secuenciación masiva y las herramientas bioinformáticas nos permiten generar y utilizar miles de marcadores moleculares en especies no modelo. Historicamente la alopatría ha sido considerada como el modo de especiación más común en animales. Dentro de la alopatría, la especiación vicariante ha sido identificada como modo de especiación más frecuente mientras que la especiación peripátrica ha sido controversial debido a la dificultad que representa identificar un evento fundador y al efecto que tiene este evento en las poblaciones. En México el Eje Neovolcánico (ENV) es un arco volcánico del Neogeno compuesto por más de 4000 volcanes cuya edad varia desde los 23 Ma hasta el Holoceno. Las salamandras plethodóntidas representan más del 60% de las especies de caudados en el mundo, dentro de estas, la tribu Bolitoglossini es exclusiva a la zona Neotropical y representa la mayor radiación de salamandras en los trópicos. En el ENV habitan cinco especies microendémicas de salamandras plethodóntidas del género *Pseudoeurycea* propias a cinco volcanes (*P. altamontana*, *P. longicauda*, *P. robertsi*, *P. melanomolga* y *P. gadovii*) divididas en dos clados (Veracruz y EdoMex de acuerdo en la relación filogenética entre ellas) representando un sistema natural para estudiar la especiación geográfica. Utilizando 27 transcriptomas individuales (ensamblados y anotados *de novo*) para obtener sets de ortólogos de más de 10,000 loci por clado y generando millones de simulaciones coalescentes utilizando estadística pseudobayesiana bajo tres modelos de especiación alopátrica, concluyo que las especies del clado Veracruz (*P. melanomolga* y *P. gadovii*) se originaron por un evento vicariante hace menos de 200 ka y que el tiempo de divergencia es correspondiente al tiempo de formación de los volcanes donde habitan. Sugiero para el clado EdoMex un posible evento peripátrico que dio origen a *P. robertsi* y *P. altamontana*. Los resultados obtenidos demuestran la influencia de la historia geológica del ENV en las poblaciones de las salamandras plethodóntidas.

Abstract

Characteristics of extant populations are the result of different demographic events that occurred in the past. Each population stores in its DNA signatures of its evolutionary and demographic history and each gene has its own independent history. If we aim to elucidate the demographic history of a population and the history of species formation we have to look for these signatures in its gene pool in a multilocus context. Next Generation Sequencing (NGS) and bioinformatic tools allow us to generate thousands of molecular markers in non-model species. Allopatric speciation has been considered as the most common process by which species of animals form. Within allopatric speciation, vicariant speciation has been identified as the most common mode of speciation while peripatric speciation has been controversial because of the difficulty of identifying founder events and the effect of these events on populations. In Mexico, the Trans-Mexican Volcanic-Belt (TMVB) is a Neogene volcanic arc composed for more than 4000 volcanoes that arose from 23 Ma to the Holocene. Plethodontid salamanders represent more than 60% of caudate species, within these the Tribe Bolitoglossini is exclusive to the Neotropics and represents the largest radiation of salamanders in the tropics. In the TMVB five microendemic plethodontid species of the genus *Pseudoeurycea* occur, each of them found only on one volcano (*P. altamonatana*, *P. longicauda*, *P. robertsi*, *P. melanomolga* and *P. gadovii*). They are divided into two clades (Veracruz and EdoMex, based on phylogenetics relationships between them) representing a natural island system for studying geographic speciation. Using 27 individual transcriptomes (assembled and annotated *de novo*) to get sets of orthologs with more than 10,000 loci per clade and generating millions of coalescent simulations using Approximate Bayesian Computation under three models of allopatric speciation, I conclude that species of the Veracruz clade (*P. melanomolga* and *P. gadovii*) are the result of a vicariant event that occurred 200 ka and that the divergence time corresponds to the formation of the volcanoes where they occur. For the EdoMex clade, I suggest possible peripatric events that gave rise to *P. robertsi* and *P. altamontana*. My results demonstrate the influence of the geological history of the TMVB in these species of plethodontid salamanders.

Introducción

Los modos de especiación geográfica (alopátrico, parapátrico y simpátrico) han sido ampliamente estudiados en diferentes tipos de organismos con el objetivo de proporcionar indicios sobre el papel que juegan diferentes factores como el aislamiento geográfico (barreras físicas), la migración, la selección, la deriva génica y la adaptación en los organismos (Coyne & Orr, 2004). De estos modos de especiación ha sido la la vicarianza (alopatría) el modo que más común que se reporta como causa de divergencia en las poblaciones animales (Barraclough & Vogler, 2000; Freitas et al., 2016; Macher et al., 2015). Por otro lado, la especiación peripátrica ha sido menos estudiada y probada debido a la dificultad que representa la detección de un evento fundador ocasionado por un cuello de botella en el pasado de las poblaciones (Coyne & Orr, 2004; Harrison, 1991). El presente estudio está enfocado en elucidar mediante la selección de diferentes modelos de especiación alopátrica el modo de especiación (vicarianza o peripatría) que ha originado las especies de salamandras pletodóntidas microendémicas del este del Eje Neovolcánico (ENV). Estas salamandras (*P. altamontana*, *P. longicauda*, *P. gadovii*, *P. melanomolga*, y *P. robertsi*) están restringidas a los volcanes desde la frontera entre Michoacán y el Estado de México en el oeste hasta el Pico de Orizaba y Cofre de Perote en el este, formando un sistema natural de “islas” en la zona continental para el estudio de la especiación. Los últimos dos de estos volcanes se formaron hace menos de 1.3 Ma (Carrasco-Nunez, 2000; Carrasco-Nuñez & Nelson, 1998).

Los procesos demográficos que ocurren durante un evento vicariante y peripátrico son diferentes, la principal diferencia recae en la reducción poblacional (cuello de botella) que ocurre en un evento peripátrico mientras que en la vicarianza aunque pueden experimentarse reducciones en el tamaño poblacional estas no deben ocasionar un evento fundador en las poblaciones. Dichos eventos demográficos dejan huellas en el acervo genético de las poblaciones. Obteniendo transcriptomas individuales para 10 individuos de *P. gadovii* y *P. melanomolga* (5 individuos por población) obtuve más de 10 mil ortólogos para probar diferentes escenarios de especiación geográfica utilizando calculo pseudobayesiano bajo un esquema coalescente. Considerando 3 especies más de salamandras microendémicas del mismo género (*P. altamontana*, *P. longicauda* y *P. robertsi*) restringidas a las zonas volcánicas de Lagunas de Zempoala, Zitácuaro y el Nevado de Toluca respectivamente, generé 17 transcriptomas (5, 6 y 6 individuos respectivamente por población) y obtuve diferentes estadísticos de resumen que sugieren un posible cuello de botella y un origen reciente para *P. robertsi*. Concluyo que el efecto la historia del ENV influye en la diversificación de especies montanas en los trópicos.

Marco Teórico

1. Geografía de la especiación

Un evento de especiación (origen de nuevas especies) ocurre cuando una población acumula suficientes cambios a través del tiempo que le impiden el flujo génico y la reproducción con cualquier otra población, generando dos linajes con historia evolutiva independiente. La velocidad a la que pueden acumularse estos cambios no es uniforme y depende de factores como el tamaño efectivo poblacional, interacciones intra- e interespecíficas como la selección sexual y la temporalidad de eventos reproductivos, presión de selección impuesta por las condiciones medioambientales y el nivel del flujo genético entre poblaciones (Futuyma, 2014).

Los eventos de especiación pueden ocurrir de diferentes formas, y para su estudio se han clasificado bajo diferentes criterios siendo el geográfico el más usado (Futuyma, 2014). Los modos de especiación geográfica son definidos considerando el aislamiento geográfico, la distribución, el tamaño efectivo poblacional y el aislamiento o flujo génico de las poblaciones. La influencia que ejerce el aislamiento geográfico en las poblaciones recae en la dificultad o imposibilidad para la ocurrencia del flujo génico (migración) entre ellas. El aislamiento génico (producido por una barrera física) entre las poblaciones ocasiona que cada población continúe con su propia historia evolutiva sin influencia de otras poblaciones.

El tamaño efectivo poblacional (N_e) refiere al número de individuos necesarios para mantener el nivel de deriva observado en una población ideal (Wright, 1930). El N_e determina la intensidad con la que la deriva génica y la selección (fuerzas evolutivas) actúan sobre las poblaciones (Coyne & Orr, 2004; Futuyma, 2014). En poblaciones pequeñas la deriva se vuelve una fuerza mucho más fuerte que la selección ya que un alelo raro puede fijarse en la población más rápidamente por efecto del muestreo. En poblaciones grandes la selección es una fuerza predominante esto significa que las frecuencias alélicas de los individuos con mayor adecuación (éxito reproductivo) aumentarán mientras que los alelos que no están presentes en los individuos con mayor adecuación bajarán su frecuencia en la población (Futuyma, 2014; Soler, 2002).

De acuerdo al N_e de las poblaciones ancestrales, la distribución y el aislamiento geográfico se consideran tres modos básicos de especiación

1. Alopatría
 - a. Vicarianza
 - b. Peripatría
2. Parapatría
3. Simpatria

1.1. Especiación Parapátrica y Simpátrica

Los modos de especiación parapátrico y simpátrico se caracterizan por que no existe una separación geográfica total entre las poblaciones (Soler, 2002; Coyne & Orr, 2004; Futuyma, 2014)

En la especiación parapátrica, la distribución de la población es continua y existe la oportunidad para flujo génico a través de la distribución de la población. En este modo de especiación, la adaptación local juega un papel importante ya que debido a esta se comienzan a originar demes con aislamiento reproductivo aunque aún con flujo génico entre ellos (Soler, 2002; Futuyma, 2014). Un deme es un grupo de individuos de la misma especie con suficiente proximidad para compartir un sistema de apareamiento y reproducirse entre sí (Templeton, 2006). Teóricamente la especiación parapátrica puede ocurrir si el flujo génico entre demes adyacentes es menor que la fuerza de selección y la adaptación local dentro de las poblaciones (Coyne & Orr, 2004). En términos de genética de poblaciones la especiación parapátrica es un intermedio entre la especiación alopátrica y simpátrica ya que el nivel de flujo génico oscila entre 0 (alopatría) y 0.5 (simpatría) (Coyne & Orr, 2004).

Los modelos de especiación parapátrica pueden agruparse en dos tipos:

A. Modelo Clinal

Refiere a la adaptación local de las poblaciones a un gradiente ecológico (Coyne & Orr, 2004; Fisher, 1930). (Fisher, 1930) argumenta que la heterogeneidad del ambiente podría favorecer la adaptación local mediante una reducción de la migración previniendo el movimiento de individuos hacia zonas donde no están adaptados. Sin embargo, (Balkau & Feldman, 1973) consideran que la reducción de la migración per sé no conduce hacia la especiación sino que además debe existir una fuerte diferenciación genética. Posteriormente Fisher propone que el apareamiento selectivo puede evolucionar en poblaciones diferenciadas para prevenir hibridación no adaptativa, esto significa que los individuos adaptados localmente a un gradiente ambiental prefieren aparearse con individuos de fenotipo similar permitiendo la divergencia por comportamiento de las poblaciones (Coyne & Orr, 2004).

B. Modelo Stepping-stone

Este modelo considera a las subpoblaciones como unidades discretas en vez de considerarlas como un continuo en un gradiente ambiental (modelo clinal) (Coyne & Orr, 2004). Cada subpoblación mantiene migración pero sólo con las subpoblaciones más cercanas, es decir, incorpora el criterio de aislamiento por distancia (Coyne & Orr, 2004). Felsenstein (1981) propuso un modelo basado en tres loci, dos de ellos involucrados en la adaptación local y uno en el apareamiento. Bajo este modelo, la migración entre poblaciones se reduciría con el tiempo, lo cual conduciría al aislamiento reproductivo e independencia evolutiva de las subpoblaciones.

La especiación simpátrica ocurre cuando las poblaciones que divergen se encuentran en el mismo espacio geográfico (Futuyma, 2014), lo anterior implica que no existe una barrera física para el flujo génico (Soler, 2002; Futuyma, 2014). La divergencia de demes o poblaciones puede ser resultado de la especialización de nicho o de eventos de hibridación (Futuyma, 2014).

La divergencia por especialización ocurre cuando un grupo de organismos colonizan un nuevo nicho, p.e. consumir un nuevo recurso, poder colonizar un nuevo micro hábitat, colonizar un nuevo hospedero (en el caso de parásitos) o seleccionar parejas para apareamiento con características muy específicas (Soler, 2002; Futuyma, 2014). Estas adaptaciones locales con el tiempo resultaran en el aislamiento reproductivo de demes que comparten un mismo espacio geográfico (Soler, 2002).

El primer caso (divergencia por especialización) considera que la selección disruptiva actúa en al menos dos loci: el primero para la adaptación a un recurso como lo es el poder comer una planta que antes no se consumía en la población y el segundo que influye algún rasgo etológico del organismo como el apareamiento (Coyne & Orr, 2004; Futuyma, 2014). Sin embargo, no se conocen muchos casos definitivos de la especiación simpátrica, debido a la dificultad en descartar la posibilidad de que había un periodo de alopatria en el pasado entre las especies hermanas (Coyne & Orr, 2004).

El segundo caso (divergencia por hibridación) se basa en el apareamiento interespecífico entre especies estrechamente cercanas o hermanas que comparten una parte o la totalidad de sus distribución. Bajo este escenario, se crean las zonas de hibridación (zona geográfica de apareamiento entre especies) (Soler, 2002; Coyne & Orr, 2004; Futuyma, 2014). Es necesario que los híbridos no sean estériles y que desarrollen independencia evolutiva y su adecuación no sea baja en comparación con sus parentales (especies que hibridan), además pueden poseer características que les confieran ventaja frente a sus parentales (Soler, 2002). Si los híbridos son poliploides, se genera de manera instantánea el aislamiento reproductivo don las especies parentales (Coyne & Orr, 2004).

Cualquier escenario de especiación simpátrica contiene dos aspectos controversiales. El primero es el antagonismo entre la selección y la migración, ya que mientras la selección tiende a separar a las poblaciones, el flujo génico que ocurren entre migrantes mantienen la baja estructura y diferenciación poblacional (Felsenstein, 1981; Coyne & Orr, 2004). El segundo aspecto es la coexistencia de las poblaciones; se requiere que las poblaciones hayan desarrollado diferenciación y especialización ecológica suficiente durante y después del desarrollo de las barreras reproductivas (Coyne & Orr, 2004).

Modelos recientes sobre especiación simpátrica se basan en la selección sexual (elección de pareja para el apareamiento) y la selección natural (competencia por recursos) (Coyne & Orr, 2004). Al no existir una barrera física entre las poblaciones la selección actúa como principal agente que moldea la divergencia entre las poblaciones. Las especies hermanas que comparten una distribución geográfica pueden ser el resultado de alguno de estos procesos de especiación o bien también puede ser el resultado de una especiación de tipo alopátrico con contacto secundario (Soler, 2002; Coyne & Orr, 2004; Futuyma, 2014).

1.2. Especiación Alopátrica

La especiación alopátrica es el proceso de divergencia de poblaciones geográficamente separadas debido al aislamiento genético ocasionado por una barrera física (Futuyma, 2014).

En sentido estricto, la alopatria es definida no por la distancia geográfica entre las poblaciones, sino por el impedimento de la migración entre individuos (flujo génico) (Futuyma, 2014). Para su estudio, los procesos de especiación alopátrica han sido divididos en eventos vicariantes y peripátricos basados principalmente en el tamaño de la población fundadora (Coyne & Orr, 2004; Futuyma, 2014).

Un evento vicariante ocurre cuando una población ancestral es dividida en dos o más poblaciones debido a la aparición de una barrera geográfica o ecológica (surgimiento de montañas o desiertos que separan zonas húmedas) ocasionando el aislamiento génico entre las poblaciones provocando que cada población genere independencia evolutiva (Soler, 2002; Coyne & Orr, 2004; Futuyma, 2014). Por otro lado, la especiación peripátrica ocurre cuando un grupo pequeño de individuos se aísla al margen de la distribución de la población ancestral y se vuelve inexistente el flujo génico entre la población periférica y la población ancestral. Debido al aislamiento génico y al efecto fundador ocasionado por el cuello de botella (población periférica de pequeño tamaño) la deriva génica y la presión de selección ocasionan la divergencia de esta población periférica (Mayr, 1954; Soler, 2002; Coyne & Orr, 2004; Futuyma, 2014). Un cuello de botella refiere a una reducción en el tamaño poblacional produciendo cambios en las frecuencias alélicas por efecto de la deriva génica (Mayr, 1954; Futuyma, 2014). En la especiación peripátrica generalmente un cuello de botella ocasiona un evento fundador que refiere a un evento de colonización por un grupo pequeño de individuos para establecer una nueva población (Mayr, 1954; Coyne & Orr, 2004).

Los eventos de colonización poseen características especiales que conducen a una especiación rápida en comparación con un evento vicariante (Coyne & Orr, 2004). Por ejemplo, la colonización es frecuentemente seguida por una adaptación rápida a hábitats nuevos incluso si estos son completamente diferentes (p.e. Archipiélago Galapagos y Sudamérica) (Coyne & Orr, 2004). Debido a que la población colonizadora es pequeña y consiste en una muestra restringida de la variación genética de la población ancestral, aunque la selección puede ser idéntica en ambos hábitats, los alelos sujetos a selección pueden diferir. Estas diferencias genéticas pueden formar la base de la divergencia. Finalmente, la deriva génica en la población colonizadora puede ocasionar la fijación de alelos raros o novedosos (aunque deletéreos a veces) mismos que pueden promover incompatibilidades entre la población ancestral y descendiente (Coyne & Orr, 2004). Aunque la población que ha divergido experimente cambios en el N_e como expansiones o reducciones, los cambios en las frecuencias alélicas con respecto de la población ancestral resultado del cuello de botella y efecto fundador permanecerán a través de las generaciones (Mayr, 1954; Futuyma, 2014).

A pesar de que la especiación alopátrica ha sido el modo de especiación más ampliamente documentado, múltiples estudios realizados sobre diferentes tipos de organismos han concluido que la vicarianza es el modo de especiación más probable en cada uno de ellos (Macher et al., 2015; Freitas et al., 2016) basados principalmente en los patrones de distribución de la mayoría de las especies de plantas y animales y a que la especiación ocurre mayoritariamente en ausencia de flujo génico (Coyne & Orr, 2004). La observación de patrones de distribución que no corresponden propiamente al resultado de un evento vicariante ha resultado en la inferencia de la divergencia por peripatria de grupos de especies hermanas (Mayr, 1942; Rovito, 2010; Ikeda et al., 2012; Taylor et al., 2014; Lawson et al.,

2015; Rittmeyer & Austin, 2015). Sin embargo, la especiación peripátrica así como la importancia de los efectos fundadores en un proceso peripátrico ha sido controversial durante mucho tiempo (Harrison, 1991; Coyne & Orr, 2004).

Las características que pueden indicar un posible evento de especiación peripátrica son (Coyne & Orr, 2004):

- a) Las especies hermanas tienen rangos de distribución muy asimétricos.
- b) El efecto fundador como resultado de un cuello de botella causa que la variación genética poblacional sea más baja en las especies que divergen por eventos peripátricos en comparación con especies hermanas de amplia distribución.
- c) Debido a que la especiación peripátrica puede ocurrir rápidamente, es posible que especies "recientes" que cumplan con el inciso a sean resultado de un evento peripátrico.

2. Estudio de la especiación geográfica

La especiación geográfica considera para su estudio el área de distribución y propiedades demográficas y genéticas de las poblaciones en cuestión, como el tamaño efectivo poblacional (N_e) y la tasa de migración (m), el nivel de estructura poblacional, y la diversidad genética (referencia). Los análisis de estructura filogenética de las poblaciones que se desean estudiar también pueden proporcionar indicios el proceso de especiación más probable (Coyne & Orr, 2004).

2.1. Estudios clásicos de especiación geográfica

Antes de la disponibilidad de las técnicas modernas de secuenciación masiva, el estudio de la especiación geográfica utilizaba los patrones de bandeo cromosómico, las aloenzimas o la amplificación de uno o varios genes mitocondriales y nucleares y la estructura filogenética (basada en uno o pocos loci) como fuente de variabilidad y información genética en la inferencia de procesos especiación. Específicamente, la alopatria ha sido el modo de especiación geográfico más documentado, por tanto existen múltiples ejemplos que demuestran que la vicarianza es el modo de especiación más probable en la divergencia de diferentes especies de animales como aves de la familia Cracidae (Pereira & Baker, 2014), krill (*Euphausia spp.*) (Patarnello et al., 1996) y mariposas, libélulas y luciérnagas (Osozawa et al., 2013).

Un caso específico de la especiación peripátrica es el correspondiente al género *Drosophila* en el archipiélago de Hawaii (Coyne & Orr, 2004). En el archipiélago de Hawaii habitan aproximadamente 400 especies endémicas de *Drosophila*. Varios autores (Carson, 1970; Carson & Templeton, 1984; DeSalle, 1995) demostraron que la colonización de este género casi siempre consistió en el movimiento de islas de origen más antiguo hacia islas de origen más reciente implicando que la formación de especies ocurre con mayor facilidad en islas con nichos vacíos. Posteriormente Coyne & Orr (2004) al combinar estos resultados con el estudio

de Carson & Templeton (1984) donde demuestra mediante un análisis de patrones de bandeo cromosómico que fueron necesarios al menos 45 eventos de colonización entre islas hawaianas para explicar la formación de las 97 especies de *Drosophila* de alas coloreadas concluyen que al menos el 50% de las especies tuvo un origen peripátrico (Coyne & Orr, 2004). Estudios sobre el género *Laupala* de grillos no voladores y las arañas del género *Tetragnatha*, ambos distribuidos en Hawaii, también sugieren el origen peripátrico de las especies (Otte & Society, 1994; Gillespie & Croom, 1995; Shaw, 2002). Cuando las islas de un archipiélago se originan sucesivamente a través del tiempo (como en Hawaii) se puede realizar una primera inferencia acerca de la especiación peripátrica correlacionando el origen de las islas y la divergencia de las especies (Coyne & Orr, 2004).

En zonas continentales, además de los endemismos, la asimetría en los rangos de distribución de especies hermanas puede indicar que la especie más joven es resultado de un evento peripátrico. Sin embargo, esta asimetría también puede ser el resultado de un evento vicariante seguido de la reducción o expansión de una de las poblaciones (Coyne & Orr, 2004). *Clarkia biloba* y *C. lingulata* (también conocidas como guirnalda de flores) son una par de especies hermanas que tienen una asimetría en el rango de distribución y diferencias cromosómicas. *Clarkia biloba* se distribuye por el cañón del Río Merced (E.U.A) y tiene 8 pares de cromosomas mientras que *C. lingulata* esta restringida al sur de la distribución de *C. biloba* y posee 9 pares de cromosomas. Además, *C. lingulata* tiene una heterocigocidad 30% más baja y muestra menor variación entre individuos a comparación de *C. biloba* según estudios de aloenzimas (Lewis & Roberts, 1956; Lewis, 1973). Las características anteriores de este par de especies sugieren que *C. lingulata* divergió rápidamente debido a la acumulación de rearrreglos cromosómicos a partir de una pequeña población de *C. biloba* genéticamente aislada debido a la dispersión limitada de semillas y a alguna autocompatibilidad (H. Lewis & Roberts, 1956; Lewis, 1973; Gottlieb, 1974). Se considera este caso como el mejor ejemplo de especiación peripátrica en la zona continental (Coyne & Orr, 2004).

2.2. Estudios recientes de especiación geográfica

En ausencia de genomas de referencia, la generación y ensamblaje de transcriptomas *de novo* resulta una estrategia adecuada cuando se pretende obtener el mayor número de marcadores moleculares. En el caso específico de especies no modelo (como las salamandras) son pocos o nulos los recursos de referencia para el ensamblaje y o anotación de transcriptomas además de que durante el proceso de ensamblaje hasta la anotación y llamado de variantes y genotipos se pueden obtener diferentes artefactos como resultado de errores en los diferentes procesos de análisis de datos (Singhal, 2013). En la actualidad se cuenta con diversas herramientas que permiten el ensamblaje, mapeo, anotación y búsqueda de ortólogos de organismos sin necesidad de una referencia previa (Singhal, 2013).

En la última década el desarrollo de tecnologías de secuenciación masiva ha permitido la generación de miles de marcadores moleculares que pueden ser utilizados para elucidar patrones y procesos evolutivos con mayor precisión (Rittmeyer & Austin, 2015). Para obtener una mejor estimación de los parámetros demográficos como el tamaño efectivo poblacional, tiempo de divergencia, y la tasa de migración es necesario el uso de un gran número de marcadores moleculares (Carling & Brumfield, 2007). Con miles de marcadores moleculares

en estudios de especiación geográfica, se pueden inferir con mayor precisión eventos demográficos en el pasado de las poblaciones para distinguir entre modelos diferentes de especiación.

El uso de marcadores moleculares múltiples e independientes en adición con métodos coalescentes permitirá probar estadísticamente diferentes escenarios de especiación para la estimación de parámetros demográficos como el tiempo de divergencia y cambios en el tamaño efectivo poblacional propios en los procesos de especiación (Rosenberg & Nordborg, 2002; Knowles, 2004). En términos evolutivos, el coalescente (Kingman, 1982) es un proceso inverso a la replicación del ADN (Templeton, 2006). Puntualmente podemos referir a un evento coalescente como la fusión de dos cadenas de ADN en el pasado, esto significa que dos o más cadenas de ADN que observamos en el presente pueden coalescer en algún momento en pasado (Templeton, 2006). Al considerar un proceso coalescente debemos asumir que:

- a) no sabemos cuantos alelos existían en el pasado,
- b) no existe variación genética en un evento coalescente,
- c) no conocemos que eventos evolutivos ocurrieron antes de la aparición del ancestro común y finalmente
- d) no conocemos cuantos alelos diferentes del locus más existieron en el momento de la última coalescencia.

El uso de la teoría del coalescente ha permitido probar hipótesis sobre procesos en el pasado de las poblaciones que pueden explicar los patrones de variación genética que observamos en la poblaciones actuales (Fu & Li, 1999). El uso de múltiples loci en un proceso coalescente implica que el cálculo de la verosimilitud se vuelva complejo o casi imposible bajo modelos demográficos complejos. Una estrategia que nos permite utilizar cientos e incluso miles de loci de manera eficaz y bajo cualquier modelo demográfico en un proceso coalescente es el Cálculo de Aproximación Bayesiana (ABC por sus siglas en inglés de ahora en adelante) (Beaumont et al., 2002; Beaumont, 2010; Csilléry et al., 2010; Sunnåker et al., 2013). Un esquema de trabajo basado en la utilización del ABC permite la utilización de miles de marcadores moleculares que pueden ser simulados bajo modelos demográficos complejos de cualquier grupo de poblaciones naturales (Beaumont et al., 2002; Beaumont, 2010; Katalin Csilléry et al., 2010; Sunnåker et al., 2013).

El ABC sustituye el valor de verosimilitud por un algoritmo de rechazo que consiste en la comparación de estadísticos de resumen obtenidos en cada simulación bajo los diferentes modelos utilizados con respecto a los valores observados de los mismos estadísticos (Tavaré et al., 1997; Beaumont et al., 2002; Hudson, 2002; Wegmann et al., 2009; Beaumont, 2010; Bertorelle et al., 2010; Wegmann & Excoffier, 2010; Sunnåker et al., 2013). Los estadísticos de resumen (S) refieren a los valores de diferentes propiedades demográficas como el número de alelos, D de Tajima, F_{ST} , diversidad nucleotídica, migración, etc. (Beaumont et al., 2002; Beaumont, 2010). S representa para cada propiedad el valor obtenido bajo una simulación considerando una combinación de valores para cada parámetro dado un modelo y dado una distribución a priori de los parámetros (Tavaré et al., 1997; Beaumont et al., 2002; Beaumont, 2010) . Por tanto, resulta más viable y menos complejo calcular el valor de las propiedades

demográficas que la verosimilitud para cada simulación (Tavaré et al., 1997; Beaumont et al., 2002; Beaumont, 2010;).

Los estudios recientes sobre la especiación geográfica combinan el uso de un gran número de marcadores moleculares y el ABC para la realización de simulaciones coalescentes bajo diferentes escenarios demográficos con el objetivo de elucidar los eventos demográficos que pueden explicar las poblaciones actuales y sus parámetros demográficos con mayor precisión y poder de discriminación entre modelos. Un ejemplo detallado acerca del empleo de miles de marcadores moleculares en el área de la historia demográfica de las poblaciones es un estudio realizado por Rittmeyer y Austin en 2015.

En este estudio (Rittmeyer & Austin, 2015) integran la secuenciación masiva (Illumina), análisis morfológicos y el ABC para la delimitación de especies, la inferencia de parámetros demográficos y procesos que pudieron conducir a la especiación de eskinco (*Tribolonotus spp.*) en las islas de Buka y Bouganville en el Archipiélago Solomon al sureste de Papua Nueva Guinea. En este estudio, los autores prueban nueve modelos demográficos diferentes que varían en complejidad desde el modelo que no implica una divergencia entre las poblaciones de las islas hasta modelos que incluyen cambios en el tamaño poblacional y en las tasas de migración entre las poblaciones de Buka y Bouganville después de la divergencia. El resultado de este estudio demostró que aunque el uso de pocos marcadores (4 loci independientes) puede arrojar estimaciones sobre el tamaño poblacional y el tiempo de divergencia similares a los obtenidos utilizando miles de loci (1526), los intervalos de confianza son mucho más amplios y la capacidad para identificar eventos migratorios es limitada (Rittmeyer & Austin, 2015).

3. Eje Neovolcánico Mexicano como escenario para la especiación

El Eje Neovolcánico (ENV) es un arco volcánico del Neogeno (~23 Ma) que se extiende de este a oeste en el centro de México. Posee una longitud de aproximadamente 1000 km y un área aproximada de 160,000 Km² abarcando los estados de Colima, Guanajuato, Estado de México, Ciudad de México, Jalisco, Michoacán, Morelos, Puebla, Oaxaca, Tlaxcala y Veracruz (Ferrari et al., 2012). Esta provincia biogeográfica compuesta por más de 8000 formaciones de origen volcánico ha sido dividida en tres grandes zonas : Este (100° 30'W–Golfo de México), Centro (103° 00'W–100° 30'W) y Oeste (West of 103° 00'W hasta el Golfo de California) (Ferrari et al., 2012; García-Palomo et al., 2000). Entre las formaciones volcánicas del ENV, se encuentran los estratovolcanes Nevado de Toluca (~2.6 Ma) (García-Palomo et al., 2002), Complejo volcánico de Chichinautzin (3.8–1.8 Ma) (Osete et al., 2000), Citlaltépetl o Pico de Orizaba (0.29Ma–4Ka) cuya formación ha sido dividida en 3 etapas principales (Carrasco-Núñez, 2000), Cofre de Perote (1.3Ma–11Ka) formado en 5 etapas principales (Carrasco-Núñez & Nelson, 1998), Complejo volcánico Zitácuaro (12 Ma con actividad postcaldera cada 0.5 Ma al menos en tres ocasiones) (Capra et al., 1997).

Debido a que muchas de las formaciones volcánicas del ENV alcanzan elevaciones por encima de los 2000 msnm, sus zonas montañosas se caracterizan por el clima templado y bosques de pino-encino (Metcalf, 2006). Por su complejidad geológica y climática el ENV se ha considerado como un “hotspot” para la biodiversidad, especialmente de vertebrados

(Morrone et al., 2007).

3.1. Estudio de la especiación geográfica en el Eje Neovolcánico

Diversos estudios han demostrado la influencia de la compleja historia geológica de zonas particulares del ENV en la historia demográfica y divergencia de las poblaciones actuales de varios organismos, incluyendo peces de agua dulce (Mateos, 2005), anfibios (Mulcahy et al., 2006; Parra-Olea et al., 2012) y angiospermas (Ruiz-Sanchez & Specht, 2013). Parra-Olea et al. (2012) demostraron que la historia demográfica de las poblaciones de salamandras apulmonadas de *P. leprosa* ha sido altamente influenciada por la historia geológica del ENV. En este estudio, se demostró la correspondencia entre los eventos migratorios (tiempo y sentido), contracciones y expansiones poblacionales de *P. leprosa* y la historia geológica en el centro y este del ENV. Además, se concluyó que *P. leprosa* divergió en este del ENV y después hubieron múltiples eventos migratorios concordantes con la actividad máfica en el centro y este del ENV.

De los resultados derivados del estudio de Parra-Olea et al. (2012), Velo-Antón et al. (2013) realizaron un estudio del efecto de la variabilidad espacial y temporal en los eventos de dispersión y migración sobre la diferenciación genética de las poblaciones actuales de *P. leprosa* por medio del modelado de nicho ecológico y la utilización de 11 microsatélites y ADN mitocondrial. Los resultados de este estudio indican que las poblaciones de noreste y sureste del ENV se mantienen aisladas del resto de las poblaciones además de identificar cinco poblaciones genéticas en el centro del ENV. Los ensayos de modelado de nicho ecológico revelan que después del último máximo glacial (UMG) *P. leprosa* ha sufrido una reducción en su rango de distribución y en la conectividad inter-poblacional. Los autores concluyeron que los cambios climáticos y de hábitat durante y después del UMG influyeron directamente en la dispersión y estructuración de las poblaciones de *P. leprosa* (Velo-Antón et al., 2013). Tanto el estudio de Parra-Olea et al. (2012) como el realizado por Velo-Antón et al. (2013) consideran el efecto de la complejidad e historia geológica del ENV en la historia natural y demográfica de las poblaciones de *P. leprosa*.

4. Plethodontidae

Del total de especies de salamandras reportadas (704 spp.), la familia Plethodontidae esta compuesta por 466 spp. representando un 66% de la diversidad total de salamandras (AmphibiaWeb, 2017). Esta familia se caracteriza por llevar a cabo el 100% de su respiración a través de la piel, poseer un surco nasolabial para quimiorrecepción y narinas externas, siendo casi exclusiva del continente americano, con unas pocas especies en Europa y Korea del Sur (AmphibiaWeb, 2017). Habitan en una gran variedad de hábitats que van desde el bosque templado o de niebla hasta zonas tropicales, cálidas y húmedas, y incluyen especies terrestres y acuáticas (Wake, 1966). La morfología de los plethodóntidos también presenta grandes variaciones en el tamaño (*T. infernalis* cuya es LHC=19 mm a *I. gigantea* con una LHC máxima de 161 mm) (Parra-Olea et al., 2016; AmphibiaWeb, 2017), forma de las extremidades y la coloración principalmente.

Los pletodóntidos se dividen en dos grandes subfamilias: Hemidactylinae (369 spp.) y Plethodontinae (97 spp.) (Wake, 2012; AmphibiaWeb, 2017). Los pletodóntidos son la única familia de salamandras de amplia representación en los trópicos, y todas las especies tropicales pertenecen a un solo linaje: la tribu Bolitoglossini (Hemidactylinae) (Wake, 2012). Esta tribu está compuesta por 14 géneros, todos de distribución mesoamericana y sudamericana (Rovito et al., 2015) y se caracterizan por ser de hábitos terrestres y desarrollo directo (Lombard & Wake, 1986; AmphibiaWeb, 2017). De los 14 géneros, el género *Pseudoeurycea* (39 spp) es el segundo género más diverso de la subfamilia sólo por debajo del género *Bolitoglossa* (132 spp.) (AmphibiaWeb, 2017). El género *Pseudoeurycea* es monofilético y está compuesto por dos grandes clados, los clado *P. leprosa* y clado *P. gadovii* (Parra-Olea, 2002). Se cree que la divergencia del género ocurrió hace aproximadamente 20 Ma (Rovito et al., 2015).

Aunque el sitio de origen de los pletodóntidos permanece incierto, se sugiere que la divergencia de la familia ocurrió hace aproximadamente 66 Ma en el Paleoceno temprano en el Noroeste de Norte América (Shen et al., 2015) y que el ancestro de la tribu Bolioglossini ocurrió hace 41 a 53 Ma en ambos lados del Istmo de Tehuantepec (Rovito et al., 2015). Dentro de la familia, la divergencia entre los géneros *Bolitoglossa* y *Pseudoeurycea* ocurrió hace 24 a 31 Ma (Rovito et al., 2015). Se considera además que la distribución de todas las salamandras tropicales ancestrales abarcó Centroamérica y el ENV, y que la ruta de dispersión del ancestro pudo ser desde California hasta el Nevado de Colima o por la costa Este hasta Centroamérica y el ENV hasta la Sierra Madre Oriental (Rovito et al., 2015).

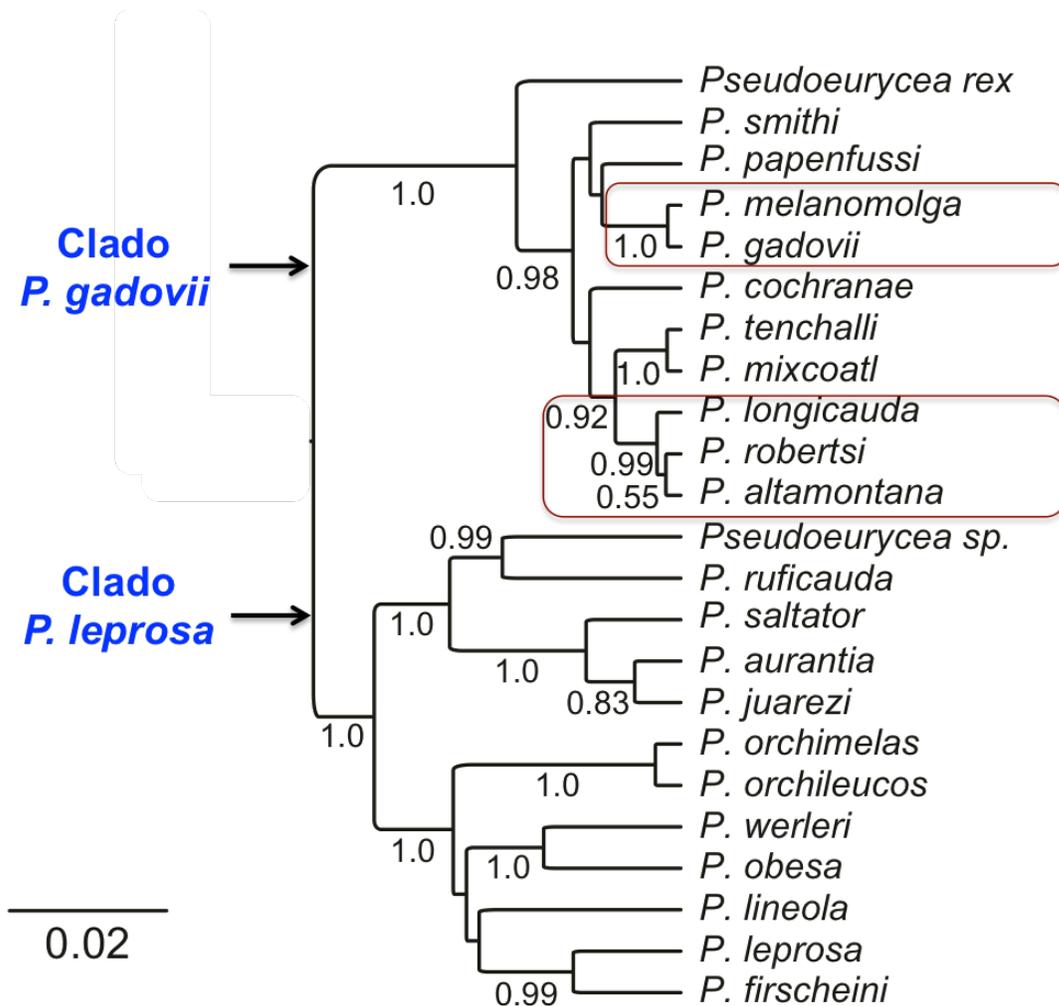


Figura 1: Filogenia del género *Pseudoeurycea*. Árbol de especies estimado con seis genes mitocondriales (cytb, 12S, 16S, ND1, COI y COII) y tres nucleares (POMC, RAG1, SLC8A3) utilizando BEAST. Números en la ramas indican la probabilidad posterior del clado. Flechas indican los 2 clados principales; especies microendémicas incluidas en este estudio enmarcadas en rojo. Modificado de Rovito et al., 2015.

4.1 Diversidad de Plethodontidae: *Pseudoeurycea* en el ENV

En el ENV habitan 41 especies de salamandras de la familia Plethodontidae pertenecientes a siete géneros, de los cuales *Pseudoeurycea* es el más diverso (Rovito et al., 2015). Del total de 41 especies, 13 especies pertenecen al género *Pseudoeurycea* (Rovito et al., 2015). De estas, *P. leprosa* se distribuye desde el este hacia el centro del ENV (Parra-Olea et al., 2012) mientras que *P. gadovii*, *P. melanomolga*, *P. altamontana*, *P. robertsi* y *P. longicauda* son microendémicas para los volcanes Pico de Orizaba, Cofre de Perote, Complejo volcánico Chichinautzin (Zempoala), Nevado de Toluca y Complejo volcánico de Zitácuaro respectivamente (AmphibiaWeb, 2017). Los resultados de Rovito et al. (2015), Parra-Olea (2012, 2016) y Velo-Antón *et al.* (2013) sugieren que la historia geológica en el ENV ha tenido un gran impacto en la diversificación y distribución de las salamandras plethodontidas.

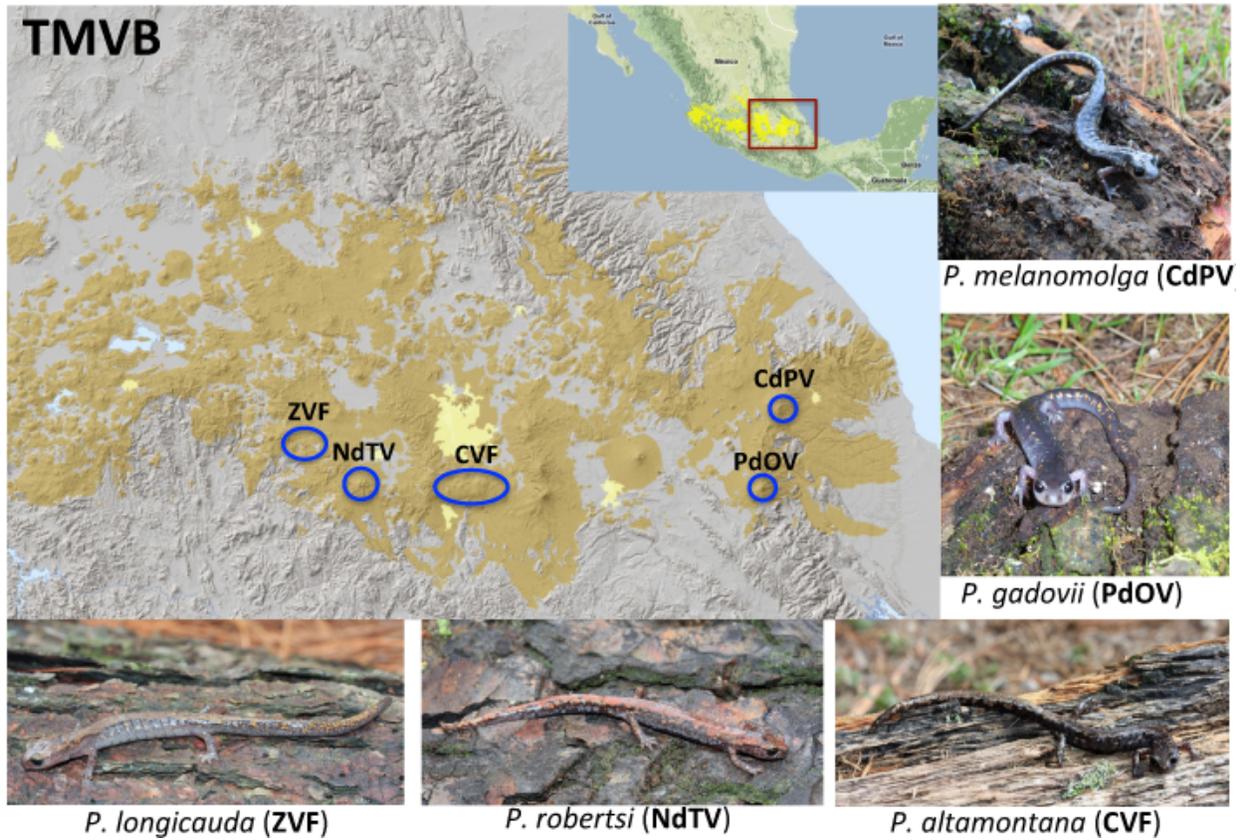


Figura 2: Distribución de salamandras pletodóntidas en el Centro-Este del Eje Neovolcánico. Complejo Volcánico Zitácuaro (ZVF), Volcán Nevado de Toluca (NdTV), Complejo volcánico Chichinautzin (CVF), volcán Pico de Orizaba (PdOV) y volcán Cofre de Perote (CdPV). Modificado de Ferrari et al., 2012. Fotos: Sean M. Rovito.

Hipótesis

Cada vez que surge un volcán en el centro-este del Eje Neovolcánico existe un grupo de salamandras que lo coloniza y mediante un proceso de especiación peripátrica se originan las especies actuales que habitan cada sitio.

Objetivo general

Inferir sobre los procesos de especiación de las salamandras *P. altamontana*, *P. longicauda*, *P. robertsi*, *P. gadovii* y *P. melanomolga*. Probar la correspondencia entre los tiempos de formación de los volcanes y los tiempos de divergencia de las especies. Probar la presencia de un cuello de botella asociado con la formación de las especies microendémicas.

Objetivos particulares

- 1) Obtener el transcriptoma individual y el gen *cytb* de las especies de *Pseudoeurycea* en cuestión.
- 2) Obtener los genotipos ortólogos y SNPs para cada uno de los clados en el estudio.
- 3) Diseñar modelos demográficos para diferentes escenarios geográficos de especiación.
- 4) Realizar pruebas de hipótesis integrando la información de los transcriptomas y modelos demográficos mediante el uso de Approximate Bayesian Computation.

Método

1. Muestreo poblacional

Para la obtención de ejemplares de cada población en cuestión se muestrearon cinco zonas volcánicas (Complejo Volcánico Zitácuaro, Complejo Volcánico Chichinautzin (Lagunas de Zempoala), Nevado de Toluca, Cofre de Perote y Citalatépetl (Figura 2) en los meses de mayo y julio de 2015. En cada zona volcánica se identificaron zonas de bosque en buen estado de conservación o poco perturbadas y con humedad visible; una vez identificadas las zonas se procedió a la búsqueda y colecta de salamandras en nichos específicos como lo son debajo de troncos caídos y húmedos en estado de descomposición, debajo de la corteza e incluso entre la madera de los mismos, debajo de rocas, y por laderas muy húmedas. Una vez localizado un ejemplar se colectó con las manos y se guardó en una bolsa plástica para su fácil traslado hasta el laboratorio. El anexo 1 resume el número de ejemplares colectados por sitio, especie y número de catálogo asignado dentro de la colección de anfibios y reptiles SMR.

Los ejemplares fueron anesteciados con una solución de clorotone, sacrificados y disectados con el objetivo de obtener tejidos para la extracción de ADN (hígado) y ARN (corazón, bazo, cola, hígado e intestino). Posteriormente, los ejemplares fueron depositados en la colección SMR para futuras consultas. Todos los tejidos fueron preservados en una solución de ARNlater y almacenados a -70°C

2. Extracción de ADN y amplificación de Cytb

Cada extracción de ADN se realizó a partir de tejido de hígado realizando una extracción por individuo de acuerdo a (Aljanabi & Martinez, 1997) (Anexo 3). A partir de la cuantificación en NanoDrop ND-1000 (Thermo Fisher) se generaron alícuotas de cada muestra de ADN a una concentración de $20\text{ ng}/\mu\text{l}$ para ser utilizadas en la amplificación del gen mitocondrial cytb. Para la amplificación de cytb se utilizaron los oligos MVZ15 (F) y MVZ16 (R) (Moritz et al., 1992) (Anexos). Los productos de PCR fueron purificados mediante el método con Exosap-IT (Affimetrix-USB). La secuenciación capilar de cada uno de los productos fue realizada por Servicios Genómicos Langebio (<http://langebio.cinvestav.mx/labsergen/>). Las secuencias fueron editadas y ensambladas *de novo* utilizando el software Geneious R8 (Kearse et al., 2012).

3. Extracción de ARN, preparación de librerías de ADNc y secuenciación

Para cada individuo se realizaron dos extracciones de ARN total: una con hígado y corazón y otra con cola, bazo e intestino. La extracción de ARNtotal se realizó modificando el protocolo de Chomczynski (1993) (Anexo 4). La concentración de cada extracción se midió utilizando un Nanodrop ND-1000. Dichas concentraciones de cada extracción fueron utilizadas para calcular el volumen de cada una de las extracciones (a y b) necesario para obtener $1\text{ }\mu\text{g}$ de cada extracción y poder iniciar la preparación de las librerías a partir de $2\text{ }\mu\text{g}$ de ARN total de

cada ejemplar.

En total se elaboraron 27 librerías de acuerdo a los ejemplares colectados (Anexo 1) de cDNA a partir de ARN total modificando el protocolo de Illumina TruSeq RNA Library Preparation Kit v2 (Anexo 7). La concentración, tamaño y calidad de cada librería fueron determinadas por Servicios Genómicos Langebio (<http://langebio.cinvestav.mx/labsergen/>) mediante un Bioanalyzer DNA1000 (Agilent Technologies). Las librerías fueron secuenciadas en dos líneas 100PE en la plataforma HiSeq4000 Illumina (QB3 Vincent J. Coates Genomics Sequencing Laboratory, UC Berkeley, CA).

4. Procesamiento de datos crudos (RNA-seq)

4.1 Análisis de calidad, limpieza, ensamblado y anotación de transcriptomas

Los datos crudos (raw data) o reads obtenidos de la secuenciación fueron analizados con el software FastQC v0.11.2 (Andrews, 2011-2015) con el objetivo de determinar la calidad promedio de los reads por librería. Debido a la cantidad de datos a analizar (tamaño de archivo y número de reads) desde el análisis de calidad hasta el proceso de mapeo y llamado de genotipos/SNPs fue realizado en el cluster MAZORKA en la Unidad de Genómica Avanzada en Cinvestav Irapuato.

Una vez conocida la calidad de los reads, la limpieza (eliminación de secuencias de baja complejidad, baja calidad, contaminantes y adaptadores), recorte, ensamblado y anotación de los transcriptomas fue realizado de acuerdo al pipeline desarrollado por Bi et al. (2012) y Singhal (2013). Debido a que no existen genomas y transcriptomas disponibles de salamandras para la anotación de los transcriptomas, se utilizaron como referencia los genomas, transcriptomas y proteomas de pez (*Danio rerio*), rana (*Xenopus tropicalis*), lagartija (*Anolis carolinensis*) y humano (*Homo sapiens*) (disponibles en <http://www.ensembl.org/info/data/ftp/index.html>). Para cada librería se obtuvieron cuatro transcriptomas anotados (basados en las cuatro referencias anteriores).

La anotación basada en cuatro referencias de animales pertenecientes a cuatro de las cinco clases de vertebrados (peces, anfibios, reptiles y mamíferos) se realizó con el objetivo de maximizar el número de transcritos únicos que pudieran anotarse. Debido al nivel de relación filogenética (Chordata: Vertebrata) de las salamandras con las referencias utilizadas, la elección del porcentaje de identidad (60%) estuvo basada en evitar la anotación errónea de transcritos al mismo tiempo que se optimizaba la anotación de transcritos con variaciones considerables debido a la lejanía filogenética. Los transcriptomas anotados para cada librería fueron colapsados de acuerdo a Singhal (2013) con el objetivo de obtener un transcriptoma por librería con la mayor cantidad de transcritos únicos posibles.

4.2 Mapeo y llamado de genotipos

Para la obtención de transcritos ortólogos, diseño de modelos demográficos y prueba de hipótesis, los individuos colectados de las especies en cuestión fueron agrupadas en dos clados: 1) clado EdoMex (*P. longicauda*, *P. robertsi* y *P. altamontana*) y 2) clado Veracruz (*P. gadovii* y *P. melanomolga*) de acuerdo a la filogenia de Rovito et al. (2015). El mapeo de los reads se realizó utilizando el software Bowtie2 v. 2.3.0 (Langmead & Salzberg, 2012) utilizando los reads resultantes del proceso de limpieza de librerías. Como referencia, se utilizó el transcriptoma con el mayor número de transcritos para cada clado. Para el mapeo, se consideró sólo conservar los reads que tuvieran hit en un solo sitio y que fueran pareados. El resultado del mapeo fue reformateado y ordenado utilizando distintas herramientas disponibles en SAMTools v.1.3.1 (Li et al., 2009). Propiedades como la profundidad de cada sitio en el mapeo fueron determinadas también para la elección de parámetros en el proceso de genotipificación.

La obtención de genotipos se realizó utilizando el software ANGSD v.0.917 (Korneliussen et al., 2014) considerando como referencia el mismo transcriptoma utilizado para el mapeo con Bowtie2. De acuerdo a una metodología diseñada por Ke Bi y Mathew Fujita (comunicación personal), se realizó el llamado de genotipos y SNPs para cada clado. Una vez obtenidos los genotipos y SNPs para cada clado se localizaron e integraron los SNPs en los mismos contigs, sitios e individuos y cualquier otro sitio segregante sin correspondencia con un SNP fue reemplazado con la base (ATCG) más común entre los individuos para cada sitio. De esta forma los SNPs filtrados fueron conservados en los genotipos (dos alelos para cada individuo) como filtro de calidad en los loci a utilizar. Liga para descargar los scripts utilizados en el presente proyecto: (<https://drive.google.com/open?id=0B7LRALUJMfG-Snp2NUloekpGdms>).

4.3 Edición y selección de ortólogos

Los ortólogos (genotipos) obtenidos de la integración de SNPs a los genotipos fueron modificados con el objetivo de calcular diferentes estadísticos de resumen (Anexo 8) en el software msABC (Pavlidis *et al.*, 2010). Los cambios de formato de la salida de genotipos resultado de ANGSD a la salida del software msABC se ilustran en la Figura 3.

Se generó un archivo fasta por cada locus y sus ortólogos. Los ortólogos fueron elegidos considerando aquellos contigs que se encontraran en el 100% los individuos y que tuvieran reads exclusivos (sólo mapeados en un sitio y pareados). Para el llamado de SNPs se consideró una profundidad mínima de 6 (basada en el la profundidad promedio del mapeo para retener el 75% de los reads), que la probabilidad mínima de que sea un polimorfismo real sea de $1e^{-06}$ (pre-establecida en el programa ANGSD) y al igual que en el llamado de genotipos que el contig estuviera presente en todos los individuos del clado.

Con el objetivo de evitar recombinación intralocus, se calculó la longitud de cada ortólogo y en caso de que la longitud fue mayor a 1kb se localizó la posición central del locus y se conservaron 459 bases antes y 500 bases después de la posición central para obtener una longitud máxima de 1kb por locus. Se estimó la longitud promedio real y con loci "recortados" para su utilización en la estimación posterior de los parámetros.

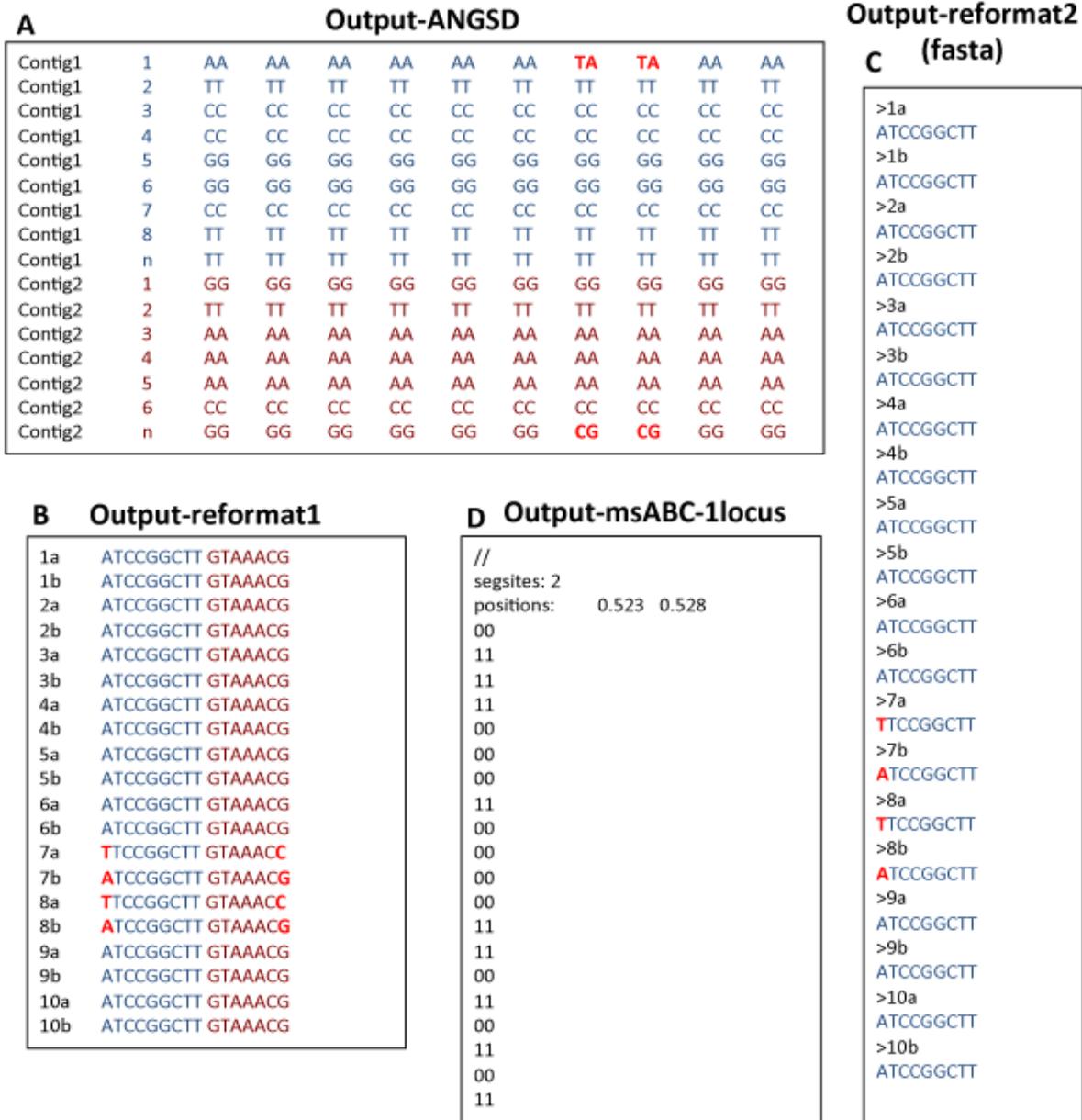


Figura 3: Reformato de los genotipos resultantes de ANGSD al formato msABC para el cálculo de los estadísticos de resumen. A) Archivo de salida de genotipos con el software ANGSD (a alelos por posición por individuo), B) conversión de columnas a filas y asignación de ID's (cada locus se parado por un espacio, en la imagen se ilustran por color); C) Conversión de B a formato fasta para usarse en msABC y D) archivo de salida del software msABC, se indica el número de sitios segregantes para un locus, las coordenadas y el estado de estos sitios en los individuos.

5. Diseño de modelos demográficos de especiación

El diseño de los modelos demográficos de especiación se realizó por separado para cada clado considerando aspectos demográficos como el tamaño efectivo poblacional, tamaño efectivo de la población ancestral y tiempo de divergencia. La estimación del árbol de cytb fue con el

objetivo de diferenciar cada una de las poblaciones y para el caso del clado EdoMex fue como base para la topología de los modelos demográficos

5.1 Generación de árbol de genes con cytb

Utilizando un modelo de partición por posición de codón obtenido del análisis del alineamiento de las secuencias de cytb con el software Partitionfinder v.1.1.1 (Lanfear et al., 2012), se estimó un árbol por el método de máxima verosimilitud en RAxML (Stamatakis, 2006) en el portal CIPRES Science Gateway (Miller et al., 2010) con 1000 bootstraps. La visualización y edición del árbol de genes se realizó con ayuda del software figtree v.1.4.3 (Rambaut, 2006-2016).

6. Prueba de hipótesis

6.1 Clado Veracruz

A. Selección de priores

Cada uno de los priores empleados en los modelos demográficos fue asignado bajo una distribución uniforme para que cada uno de los valores de cada prior tuviera la misma probabilidad de ser simulado bajo cada modelo. Los parámetros incluidos en los modelos demográficos se muestran en la Figura 4 y los rangos de los priores se muestran en el Cuadro 1

Designación de priores para el clado de Veracruz					
Modelo Alopátrico		Modelo Peripátrico para <i>P. melanomolga</i>		Modelo Peripátrico para <i>P. gadovii</i>	
N_e pmel	0–1.5	N_e pmel	0–2	N_e pmel	0–1
N_e pgad	0–2.0	N_e pgad	0–0.4	N_e pgad	0–2
N_a	0–1	N_{bpmel}	0–1.0	N_{bpgad}	0–1
T_d	0–1	T_{bnpmel}	0–2	T_{bnpmel}	0–1
theta	0-3	theta	0–5	theta	0–5
		T_d	0–2.5	T_d	0–1.2

Cuadro 1: Designación de priores para cada modelo de especiación correspondiente al clado de Veracruz. N_e refiere al tamaño efectivo poblacional, pmel: *Pseudoeurycea melanomolga*, pgad: *Pseudoeurycea gadovii*, N_a : tamaño efectivo poblacional del ancestro, T_d : tiempo de divergencia y T_{bn} : duración del cuello de botella. Una vez obtenido el valor estimado de la mediana de theta, este valor se divide entre 4μ *longitud promedio de los transcritos para obtener el valor de N_o que a su vez es utilizado como factor de conversión para cada uno de los priores (ver fórmula 4).

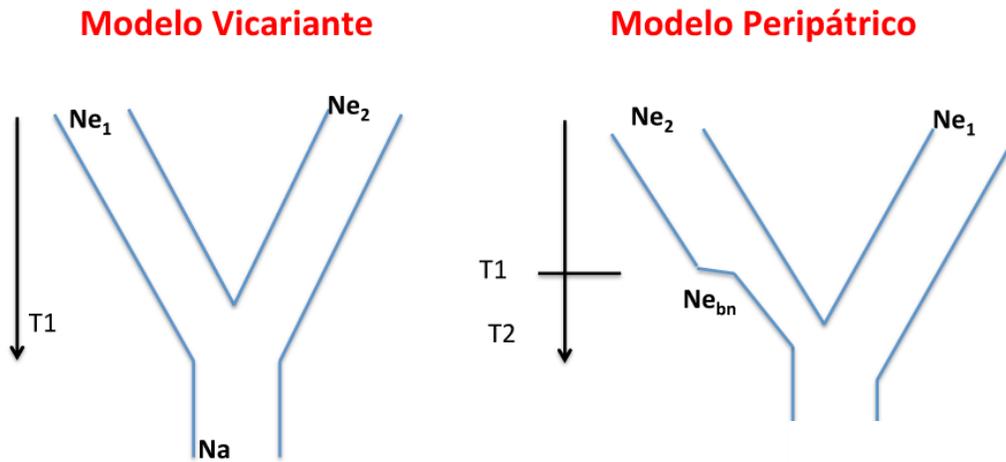


Figura 4: Esquema de los modelos demográficos del clado Veracruz. Modelo vicariante (izquierda) y modelo peripátrico (derecha).

B. Obtención de Estadísticos de Resumen

El diseño de los modelos demográficos de especiación sólo se realizó para el clado Veracruz (*P. melanomolga* y *P. gadovii*) debido a la complejidad, costo computacional y tiempo que fue requerido para la optimización de los priores para cada uno de los modelos. El set final de priores para los modelos del clado Veracruz estuvo basada en aproximadamente 20 rondas previas de simulaciones que mostraban la insuficiencia de la cobertura de los priores (datos no incluidos). Al considerar una distribución uniforme para cada uno de los priores en los modelos del clado Veracruz aseguramos que cada valor del prior tuviera la misma probabilidad de ser simulado. Finalmente el rango de cada prior fue optimizado para que no existiese solapamiento entre modelos y se permitiera la distinción entre estos para realizar una elección de modelo discriminatoria.

La obtención de los estadísticos de resumen de los datos observados (S_{obs}) y simulados (S_{sims}) (marcadores moleculares obtenidos de los transcriptomas) se realizó con la ayuda del software msABC (Pavlidis et al., 2010). Para la comparación y selección de modelos se consideraron los siguientes estadísticos (media y varianza) para cada población:

- D de Tajima (prueba de neutralidad)
- θ_w (la theta de Watterson)
- π (diversidad nucleotídica π)
- número de sitios segregantes

C. Selección del modelo más probable

Todo el proceso de análisis de datos y selección del mejor modelo se realizó utilizando el paquete abcvignette v.2.1 (Csilléry et al., 2012). El primer aspecto a considerar en la prueba de hipótesis/selección de modelo es la visualización de los S_{sims} , para esto se realizaron gráficos de cajas (boxplots) de cada uno de los estadísticos de cada modelo a considerar en la prueba de hipótesis.

El primer análisis considerado previo a la elección del mejor modelo fue la validación cruzada para probar que los modelos eran mutuamente excluyentes y diferentes entre sí. Posteriormente se realizó una prueba de bondad de ajuste de los datos simulados para cada modelo con respecto a los datos observados. Posteriormente se realizó la estimación de las distribuciones de densidad posterior para cada parámetro bajo cada modelo con respecto de la distribución prior con el objeto de visualizar si los rangos de los priores eran suficientes para la estimación de los parámetros. Dependiendo de los resultados de la validación cruzada, la bondad de ajuste y la visualización de las distribuciones de densidad posterior se reajustaron los priores en los modelos y se realizaron nuevas simulaciones. Cada ronda de simulaciones consistió en una paralelización de mil procesos simultáneos por modelo con un mínimo de 200 mil simulaciones por modelo utilizando el cluster Abacus-I del Laboratorio de Matemática Aplicada y Computo de Alto Rendimiento del Cinvestav debido al costo computacional requerido para los procesos de simulación.

Una vez que los datos simulados de cada modelo fueron cercanos a los datos observados (resultado de la prueba de bondad de ajuste) y que los modelos fueron distinguibles entre sí y las curvas de densidad mostraban la suficiencia de los priores, se procedió a realizar la elección del modelo más probable, para ello se determinó la distancia adecuada a considerar entre los datos observados y los datos simulados. Dicha distancia (umbral) fue basada en la retención de al menos 500 simulaciones más cercanas utilizando el algoritmo de rechazo implementado en el abc (fórmula 1).

La elección de utilizar el software msABC (Pavlidis et al., 2010) para la realización de las simulaciones coalescentes pseudobayesianas (ABC) bajo cada modelo se basó principalmente en la naturaleza de los datos (genotipos completos) y la cantidad de los mismos (3962 loci con 20 secuencias por cada uno). La realización de la prueba de hipótesis en abc vignette (paquete de lenguaje R) consideró los aspectos esenciales para la selección de modelos ya que se consideró la exclusión mutua entre modelos y la prevalencia del modelo más probable a diferentes valores de tolerancia. Para la selección de una tolerancia adecuada se realizó una prueba de estimación de error entre los datos observados y los datos simulados de cada modelo a diferentes tolerancias considerando como una distancia adecuada al valor de la tolerancia que menor probabilidad de error obtuvo.

Considerando las simulaciones retenidas (más cercanas) correspondientes a todos los modelos la selección del modelo más probable se realizó considerando el número de simulaciones para un modelo dado con respecto de la suma total de las simulaciones de todos los modelos (formula 2) (Beaumont, 2002).

$$(1) (S_i, S_{obs}) \leq \epsilon$$

Donde: ρ es la distancia entre un valor simulado y su equivalente en los estadísticos de resumen observados y ϵ es el umbral o tolerancia máxima entre los S_i y S_{obs} .

$$(2) \frac{\rho(M_1 | X)}{\rho(M_2 | X)} \div \frac{\rho(M_1)}{\rho(M_2)} = \frac{\rho(X | M_1)}{\rho(X | M_2)}$$

Donde X representa los S_{obs} y M_1 y M_2 representan S_{sims} de dos modelos diferentes. Básicamente el modelo más probable resulta ser el que más simulaciones retenidas produjo con respecto del total entre los modelos.

Una vez que se seleccionó el mejor modelo se realizó una prueba de bondad de ajuste de cada uno de los estadísticos de resumen retenidos (posteriores) para este modelo con respecto a los observados. De cada uno de los parámetros del modelo seleccionado se construyó la distribución de densidad posterior con respecto del prior y se calcularon las medidas de tendencia central para cada set de valores en los parámetros posteriores.

Utilizando la mediana de cada parámetro (debido a que esta medida es menos afectada por valores extremos en la distribución posterior) se calcularon los intervalos "reales" para cada parámetro utilizando un factor de conversión N_0 (Pavlidis et al., 2010) de la siguiente forma:

- Se obtuvo la media geométrica en unidades de sustituciones/sitio/año de las tasas de sustitución estimadas por Shen et al. (2015) **Cuadro 2**.
- La media geométrica resultó en $4.529587e-10$ subs/sitio/año y utilizando un tiempo generacional de 2 años (Houck, 1975) se obtuvo una tasa de $9.059174e-10$ subs/sitio/gen

Partición	MCMCTREE(independent-rate, clock=2)	BEAST
1st codon	0.0325	0.0320
2nd codon	0.0172	0.0155
3rd codon	0.178	0.1750

Cuadro 2: Estimación de la tasa de sustitución en unidades de sustituciones/sitio/100Ma tomadas de Shen et al., 2015.

$$(3) \frac{\theta}{4\mu(\bar{X})}$$

Donde: θ es la mediana de los valores posteriores de theta; μ es la tasa de sustitución $9.059174e-10$ subs/sitio/gen y \bar{X} es la longitud promedio de los loci simulados (745.609763b)

Como resultado se obtuvo un N_0 de 66917.28

- De acuerdo al manual msABC (Pavlidis et al., 2010) el tamaño efectivo poblacional y la duración del cuello de botella y tiempo de divergencia se obtuvieron de la siguiente forma:

$$(4) N_e = x * N_o \quad T = x * 4N_o$$

En este caso x corresponde a la mediana de los parámetros aceptados del mejor modelo.

Para la estimación de los intervalos de confianza al 95% de los parámetros se consideraron los valores de los cuantiles al 2.5% de los datos y al 97.5% y los valores del N_e y N_o se determinaron de la misma forma en la que se utilizó la mediana.

Resultados

El árbol de genes de las especies micro endémicas de *Pseudoeurycea* muestra una clara definición de cada una de las poblaciones, lo cual indica que la topología resultante para cada clado incluye individuos de una sola especie para cada clado (Figura 5)

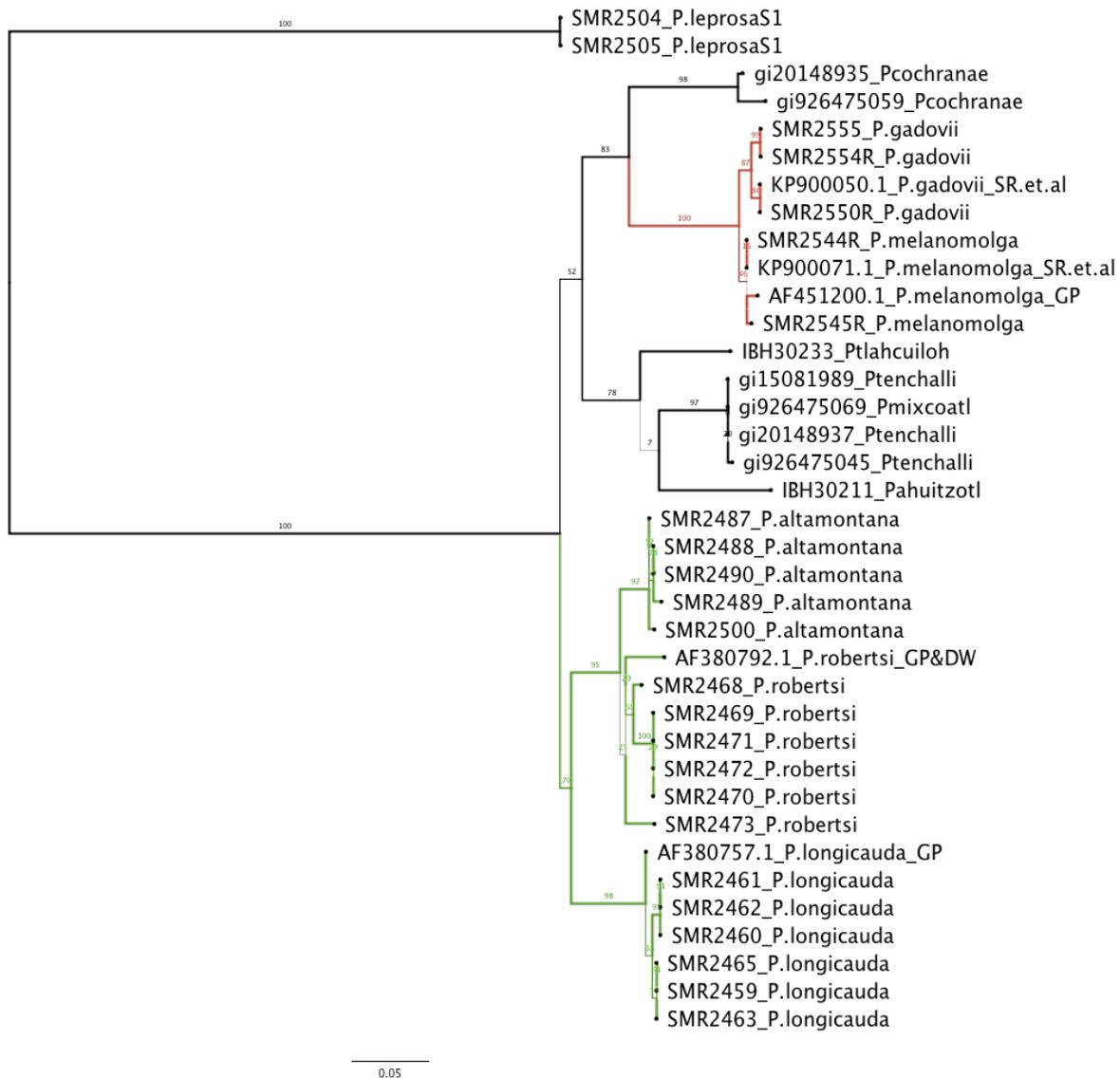


Figura 5: Árbol filogenético estimado con cytb correspondiente a las cinco especies microendémicas del TMVB. Cada color corresponde a un clado, Veracruz (rojo) y EdoMex (verde). El grosor de la rama indica el valor del bootstrap (también indicado con los valores sobre las ramas).

El número de reads (datos crudos) obtenido en promedio por especie varía entre los 12 y 15 millones como resultado de la secuenciación en la plataforma HiSeq4000. El número promedio por especie de transcritos ensamblados y anotados se muestra en el **Cuadro 3**:

Número promedio de transcritos por especie					
Transcritos/ reads	Clado Veracruz			Clado EdoMex	
	<i>P. melanomolga</i>	<i>P. gadovii</i>	<i>P. altamontana</i>	<i>P. longicauda</i>	<i>P. robertsi</i>
reads (datos crudos)	147,80,292	15,3306,25	12,813,639	13,051,673	12,194,380
T. no anotados	71,696	64,509	49,273	54,347	57,991
T. anotados	10,259	9,813	8,942	9,885	9,425
Porcentaje de la anotación (%)	14.3	15.21	18.14	18.18	16.26

Cuadro 3: Número promedio de transcritos ensamblados (no anotados y anotados) por especie.

En este estudio siguiendo el pipeline desarrollado por (Bi et al., 2012; Singhal, 2013; Mathew Fujita (comunicación personal)) se logró ensamblar, mapear, anotar y generar ortólogos de genotipos completos para 27 individuos pertenecientes a cinco especies de salamandras plethodontidas micro endémicas. La longitud promedio de los transcritos en este estudio varía entre los 876 y 1325 pb. Con un promedio de 59,563 transcritos no anotados por individuo y 9,665 transcritos anotados se logró anotar un promedio de 16.3 % de los transcritos.

1. Clado EdoMex

El el clado EdoMex el número total de ortólogos fue de 10536. De este conjunto al calcular los estadísticos de resumen se encontró que el número de loci que tenía sitios segregantes en todas las poblaciones fue de 20. El número de loci con sitios segregantes y el total de sitios segregantes para cada población fue:

Número de Loci	<i>P. altamontana</i>	<i>P. longicauda</i>	<i>P. robertsi</i>	Total
Con sitios segregantes	3,489	3,654	20	7,163
Número total de sitios segregantes	15,800	18,759	69	34,628
Sitios segregantes en loci comunes en todas las poblaciones	496	587	69	1,152

Cuadro 4: Número de loci con sitios segregantes (y número sitios segregantes) en cada población y con sitios segregantes en todas las poblaciones para el clado EdoMex.

La longitud promedio considerando los 20 loci que contenían sitios segregantes en las tres poblaciones del clado EdoMex, fue de 995 pb, mientras que los estadísticos de resumen observados considerando los mismos loci se muestran en el **Cuadro 5**.

Población	Población 1 (<i>P. melanomolga</i>)		Población 2 (<i>P. gadovii</i>)		Watterson (θ_w)		nucleotídica (π)	
	Med	Var	Med	Var	Med	Var	Med	Var
<i>P. altamontana</i>	0.0157	$1.5 e^{-04}$	-0.0139	0.896	0.00555	0.000018	0.00578	$23.7 e^{-06}$
<i>P. longicauda</i>	0.0179	$1.6 e^{-04}$	0.386	1.136	0.00593	0.000017	0.00689	$3 e^{-05}$
<i>P. robertsi</i>	0.00183	$3 e^{-06}$	-1.434	0.104	0.00061	$3.183e^{-07}$	0.00032	$1.1 e^{-07}$

Cuadro 5: Media (Med) y varianza (Var) de los estadísticos de resumen observados en el clado EdoMex considerando 20 loci con sitios segregantes en todas las poblaciones.

Además de los estadísticos de resumen mostrados en el **Cuadro 5**, el estadístico F_{ST} (Wright, 1943) fue calculado para cada par de poblaciones. F_{ST} es un índice de aislamiento poblacional, y mide la proporción de la heterocigosidad que se distribuye entre poblaciones comparado con la heterocigosidad total de todas las poblaciones. Un F_{ST} cerca de 0 indica panmixia, mientras que un F_{ST} cercano a 1 indica aislamiento poblacional total. Se obtuvieron los siguientes valores en cada caso:

- F_{ST} *P. altamontana*–*P. longicauda*: media=0.0482, var=0.00488,
- F_{ST} *P. altamontana*–*P. robertsi*: media=0.101, var=0.00199,
- F_{ST} *P. longicauda*–*P. robertsi*: media=0.221, var=0.0126

Aunque en el clado EdoMex se encontró un set de 10536 ortólogos en las tres poblaciones, sólo 20 loci contenían sitios variables (SNPs) en las tres poblaciones (*P. altamontana*, *P. longicauda* y *P. robertsi*). El número limitado de loci con sitios segregantes en la totalidad de las poblaciones es debido a que la población de *P. robertsi* sólo presentó este número de loci con sitios variables. Dentro del clado EdoMex, considerando la totalidad del set de ortólogos (10536) la población de *P. longicauda* presentó el mayor número de sitios segregantes (18,800 sitios aproximadamente) en el mayor número de loci mientras que *P. altamontana* presentó aproximadamente tres mil sitios segregantes menos que *P. longicauda* contenidos en aproximadamente 200 loci menos.

2. Clado Veracruz

El número total de ortólogos para el clado Veracruz fue de 10457; sin embargo, los ortólogos con sitios segregantes en ambas poblaciones fueron 3962 cuya longitud promedio (considerando una longitud máxima de 1 kb) fue de 745.6 pb. El número total de sitios segregantes obtenidos para ambas poblaciones fue de 25,393. El promedio y la varianza de los estadísticos de resumen de los datos observados (3962 loci) se muestra en el **Cuadro 6**.

Dentro del clado Veracruz, la población de *P. melanomolga* mostró menos sitios variables (16,732) mientras que *P. gadovii* contuvo 19,282 sitios segregantes distribuidos en los 3962 loci en común.

resumen	Promedio	Varianza	Promedio	Varianza
Sitios segregantes	0.00736	0.0000871	0.00845	0.0000999
D de Tajima	-0.1035	0.815	-0.308	0.697
θ_w	0.00260	0.0000109	0.00299	0.0000125
π	0.00257	0.0000129	0.00278	0.0000135

Cuadro 6: Estadísticos de resumen correspondientes a los datos observados empleados en la comparación entre modelos. θ_w (theta de Watterson), π (diversidad nucleotídica, pi).

2.1 Modelo más probable de especiación

Utilizando $1e^6$ simulaciones para cada modelo la validación cruzada basada en la elección de cinco simulaciones aleatorias de cada modelo se obtuvo la siguiente matriz de confusión (Cuadro 7).

La matriz de confusión obtenida de la validación cruzada indican que los tres modelos son mutuamente excluyentes y distinguibles entre sí. La prueba de bondad de ajuste demostró que los S_{ims} obtenidos son cercanos a los S_{obs} . Para cada modelo la distancia promedio entre los datos simulados y los datos observados fue de:

Tol $5e^{-06}$	Alopátrico	Peri_ <i>P. melanomolga</i>	Peri_ <i>P. gadovii</i>
Alopátrico	5	0	0
Peri_ <i>P. melanomolga</i>	0	5	0
Peri_ <i>P. gadovii</i>	0	1	4

Cuadro 7: Matriz de confusión como resultado de la validación cruzada.

- 6.6 para el modelo alopátrico,
- 16.7 para el modelo peripátrico para *P. melanomolga* y
- 7.47 para el modelo de origen peripátrico para *P. gadovii*.

Basado en el algoritmo de rechazo utilizando una tolerancia de $5e^{-4}$ (fórmula 1) y en la prueba de selección de modelos el modelo alopátrico es el modelo más probable ($p = 0.8460$) para explicar el origen de *P. melanomolga* y *P. gadovii*. A continuación se resume la prueba de hipótesis para la selección del modelo más probable.

Tolerancia : **0.0005**

Simulaciones totales por modelo: **$1e^6$**

Simulaciones retenidas (3 modelos) : **1500**

Proporción de simulaciones aceptadas=(probabilidad posterior):

- 1) **Alopátrico : 1269=0.8460**
- 2) **Peripátrico *P. melanomolga* : 119=0.0793**
- 3) **Peripátrico *P. gadovii* : 112=0.0747**

Factores de Bayes:

	Alopátrico	Peri_ <i>P.melanomolga</i>	Peri_ <i>P.gadovii</i>
Alopátrico	1.0000	10.6639	11.3304
Peri_ <i>P.melanomolga</i>	0.0938	1.0000	1.0625
Peri_ <i>P.gadovii</i>	0.0883	0.9412	1.0000

La prueba de bondad de ajuste para los valores posteriores (S_{sims}) del modelo implementando el método de regresión utilizando 50 valores aleatorios de las 500 simulaciones retenidas ($\text{tol}=5e^{-4}$) demuestra una distancia de 0.7 ± 0.5 con respecto a los valores observados (S_{obs}). La distribución de densidad posterior en comparación con la distribución de cada prior para los datos retenidos del tamaño efectivo poblacional (en número de individuos) de *P. melanomolga*, *P. gadovii* así como el valor de theta puede observarse en la Figura 6. El tiempo de divergencia en años y la edad de los volcanes Pico de Orizaba y Cofre de Perote puede observarse en la Figura 7.

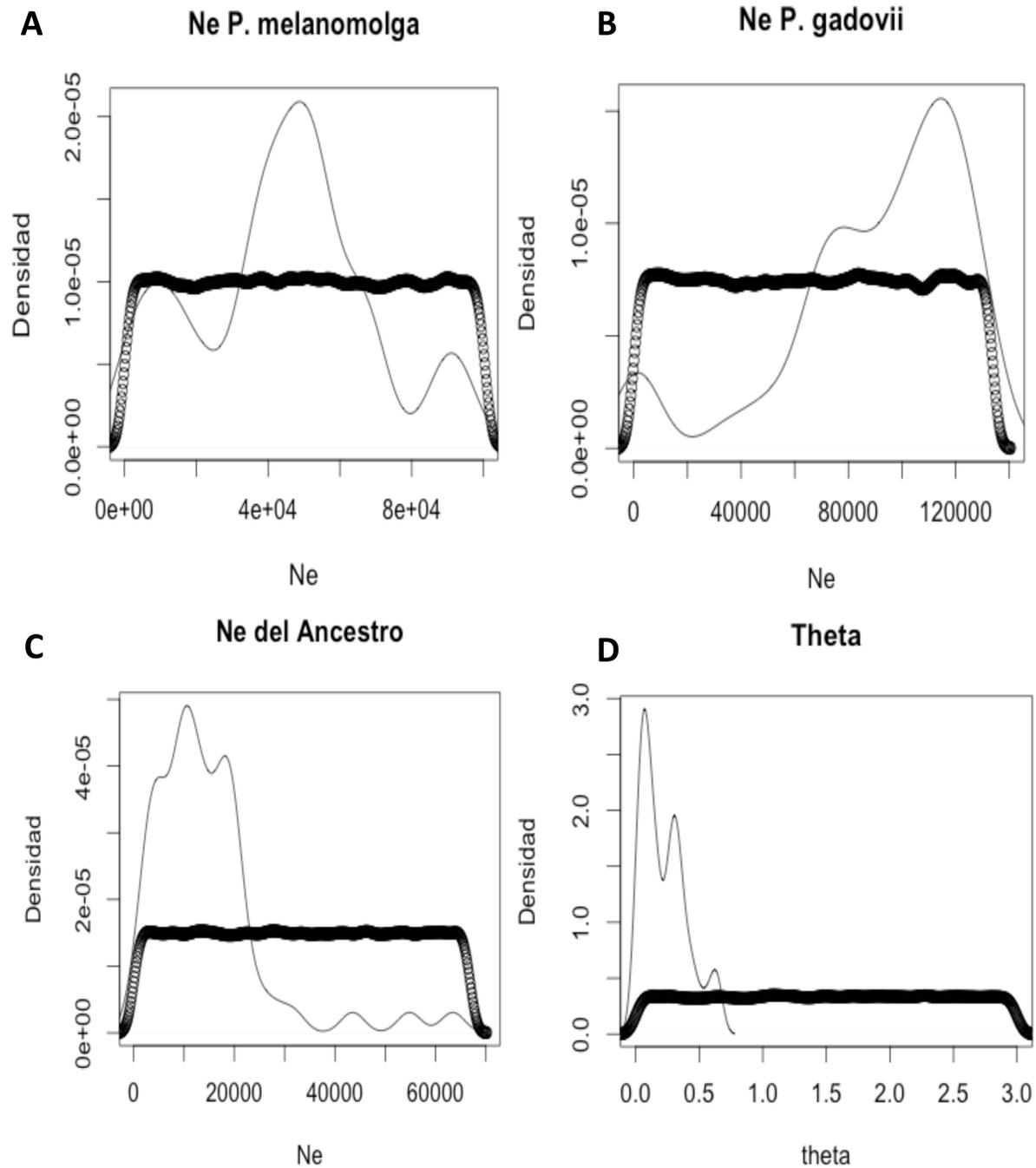


Figura 6: Densidad posterior (línea continua) del tamaño efectivo poblacional con respecto del prior (línea punteada) (N_e en número de individuos) correspondiente a *P. melanomolga* (A), *P. gadovii* (B), ancestro (C) y theta (D), $tol=0.0005$ considerando 500 simulaciones retenidas.

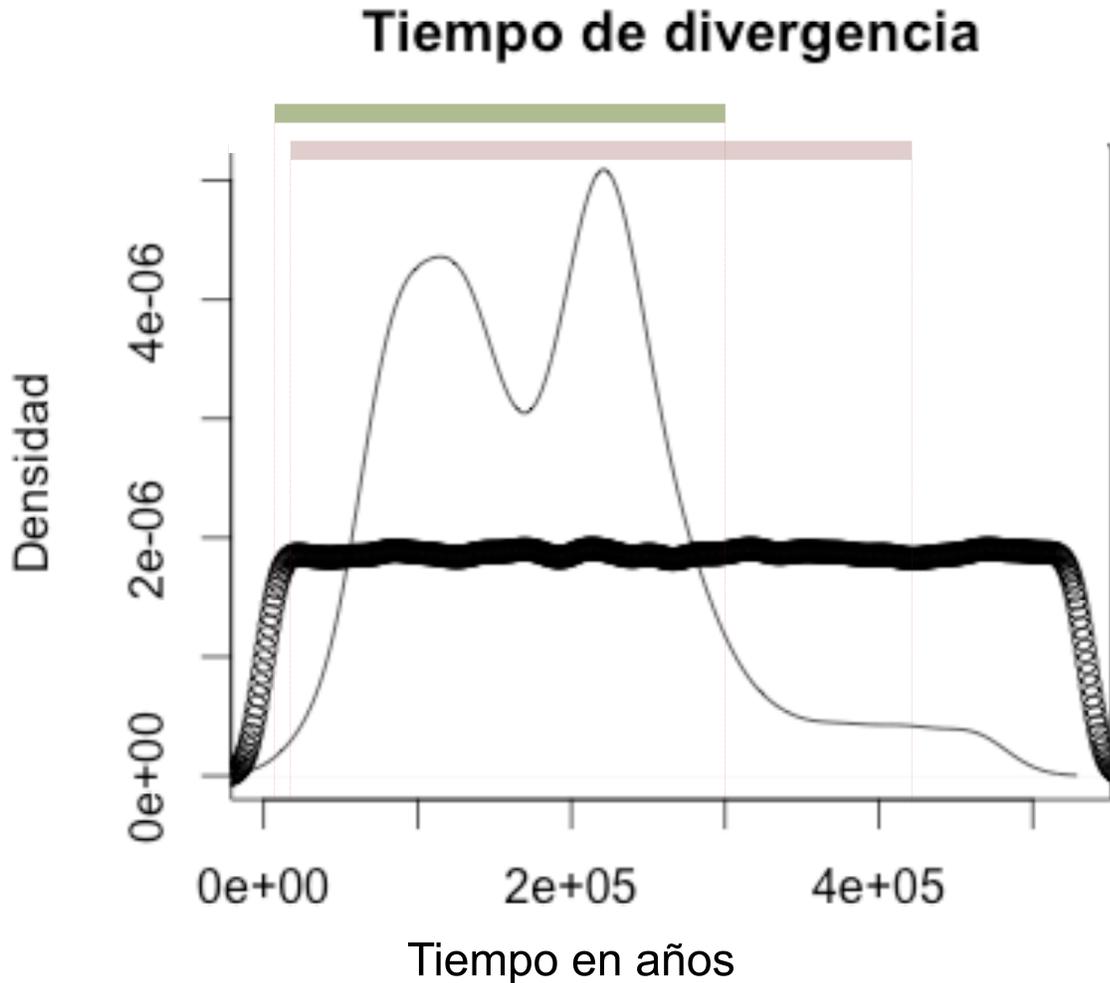


Figura 7: Densidad posterior (línea continua) del tiempo de divergencia entre *P. melanomolga* y *P. gadovii* en años a una tolerancia de 0.0005 considerando 500 simulaciones retenidas. La distribución del prior se muestra en puntos negros. La barra verde representa la formación del Pico de Orizaba (9–290 ka) mientras que la barra rosa representa la formación del volcán Cofre de Perote (13–420 ka).

El **Cuadro 8** resume los valores posteriores obtenidos para cada parámetro del modelo alopátrico.

Parámetro	Mediana	Media	Intervalo de confianza para el 95% de los datos
N_e <i>P. melanomolga</i>	46,130	43,760	25,820–56,260
N_e <i>P. gadovii</i>	100,600	90,710	7,921–115,300
N_e Ancestro	12,080	14,690	6,731–18,830
Tiempo de divergencia	185,500	186,400	116,500–231,500

Cuadro 8: Parámetros posteriores correspondientes al modelo alopátrico. Los valores correspondientes a la mediana, media e intervalos de confianza (CI) al 95% fueron calculados a partir de las 500 simulaciones más cercanas. El N_e se expresa en número de individuos y el tiempo de divergencia en años.

Discusión

1. Especiación geográfica del clado Veracruz

El modelo de especiación más soportado estadísticamente para el clado de Veracruz fue el modelo vicariante, con una probabilidad posterior de 0.85. En este estudio el modelo vicariante considera que no existieron reducciones severas (cuello de botella) en ninguna de las poblaciones al momento de la divergencia, lo que representa que las poblaciones divergieron probablemente en aislamiento geográfico y no experimentaron un efecto fundador (como ocurre en la especiación peripátrica). El modelo vicariante no considera migración ya que para este estudio asumí que las poblaciones debían estar aisladas debido a la restricción geográfica que representa la franja boscosa en los volcanes Cofre de Perote y Pico de Orizaba. Considero incluir la migración como parámetro adicional en modelos alternativos para el clado Veracruz en pruebas posteriores con el objetivo de probar la existencia o ausencia de migrantes entre estas dos poblaciones.

Los intervalos del confianza para el 95% de los parámetros posteriores obtenidos en el modelo vicariante para el clado Veracruz arrojaron que el N_e actual de *P. gadovii* puede ser mayor (por un factor de 2) que el N_e de *P. melanomolga*. Incluso considero que en el caso de *P. gadovii* es necesario ampliar el rango del prior ya que la curva de densidad posterior muestra una tendencia hacia tamaños poblacionales más grandes. Por otro lado, el modelo vicariante arrojó que el tamaño efectivo poblacional del ancestro fue siempre más pequeño con respecto al tamaño efectivo poblacional actual de las dos especies de *Pseudoeurycea*. Por la diferencia existente entre el tamaño poblacional del ancestro y las poblaciones actuales considero como opción incluir cambios en el tamaño poblacional (reducciones y expansiones) post-divergencia en modelos posteriores para poder comparar y comprobar la existencia o ausencia de estos eventos demográficos en la historia del clado Veracruz.

La curva de densidad posterior correspondiente al tiempo de divergencia en años para el clado Veracruz (Figura 7) muestra una alta correspondencia con el tiempo de formación de los volcanes Pico de Orizaba y Cofre de Perote hasta los últimos eventos máficos y volcánicos registrados para estos volcanes. Parra-Olea et al. (2012) sugirieron que la distribución y expansión de las poblaciones de *P. leprosa* distribuidas desde el este hasta el centro del ENV corresponde con la historia geológica de estas zonas. Concluyeron que las poblaciones más antiguas corresponden a la región Este del Eje mientras que las poblaciones más recientes son las que habitan en el centro del ENV. En el presente estudio el clado Veracruz presenta una alta correspondencia entre el tiempo de divergencia y el surgimiento e historia geológica de los volcanes Pico de Orizaba y Cofre de Perote.

La diferencia existente entre los tiempos de colonización de *P. leprosa* en el estudio de Parra-Olea (2012) (3.8 en el Este y 0.7 ca. en el Centro del TMVB) sugiere que las poblaciones de *P. leprosa* ya habitaban en el ENV mucho antes de la divergencia del clado Veracruz, además el estudio con *P. leprosa* estimó los tiempos de aislamiento entre poblaciones utilizando ADN mitocondrial (cytb). Debido a la naturaleza de los marcadores utilizados por Parra-Olea (2012) y los utilizados en el presente estudio la diferencia en la estimación de los tiempos puede ser debida al tipo de marcadores utilizados. Los mecanismos y fuerzas ambientales y evolutivos

que actuaron para que las poblaciones de *P. leprosa* sigan sin diferenciarse (en especies diferentes) (a pesar de su divergencia antes de la formación de los volcanes) mientras que las especies de *P. gadovii* y *P. melanomolga* hayan sido el producto del efecto ocasionado por el aislamiento geográfico al surgir los volcanes donde habitan permanecen desconocidos. Considero que uno de los factores es que quizá el número de migrantes de *P. leprosa* que colonizó cada uno de estos volcanes fue mayor que el del ancestro del clado Veracruz y que la migración durante los periodos interglaciares pudo ser considerable para mantener el flujo génico entre las poblaciones.

El estudio de Parra-Olea et al. (2012) no ha sido el único que demuestra la correspondencia entre la historia geológica del ENV y la distribución de las poblaciones. Ruiz-Sanchez y Specht (2013) utilizando un locus nuclear y dos loci de cloroplasto en 28 poblaciones de *Nolina parviflora* encontraron una alta correspondencia entre los eventos geológicos del ENV y la distribución de las poblaciones de esta planta. En su estudio, los mismo autores concluyeron que la geología del ENV influyó de manera significativa en la diversificación de poblaciones de *N. parviflora* y los eventos geológicos y orogénicos en el ENV crearon barreras físicas que imposibilitaron el flujo génico entre las poblaciones. Considerando los estudios de Parra-Olea et al. (2012) y Ruiz-Sanchez y Specht (2013) descritos anteriormente, sugiero que al igual que en el caso de *P. leprosa* y *Nolina parviflora* la dinámica geológica del ENV ha influido directamente en la divergencia del clado Veracruz, impidiendo el flujo génico entre las poblaciones ancestrales mediante el aislamiento geográfico que representa el habitar en zonas montañas cuya franja de bosque esta limitada por un valle.

Caballero et al. (2010) propusieron que las fluctuaciones climáticas ocurridas durante el Último Máximo Glacial en el Eje Neovolcánico pudieron contribuir a los procesos de especiación en este complejo volcánico. En el mismo estudio se sugiere que el descenso de las franjas de vegetación de hasta 1000 m ca. contribuyeron de manera notable en la diversificación de los géneros *Pinus* y *Quercus* en el ENV y en el resto de México. Dicha diversificación pudo ser causada además por los cambios en precipitación y temperatura en el ENV durante el Pleistoceno tardío ya que estas variaciones causaron la fragmentación y ampliación de los rangos de distribución de estos géneros de plantas.

Considerando la estimación para el tiempo de divergencia, el tamaño estimado de la población ancestral, y el modelo vicariante y considerando que las salamandras del género *Pseudoeurycea* en el ENV están restringidas a zonas boscosas que normalmente están compuestas por pino-encino y oyamel mayoritariamente en zonas montañosas no es posible atribuir la divergencia del clado Veracruz al evento del aislamiento geográfico debido al surgimiento de los volcanes sin considerar que las variaciones climáticas y fragmentación y reducción de bosque en estas zonas haya influido en el proceso de divergencia. La fragmentación del bosque y las variaciones climáticas pudieron evitar el flujo génico entre las poblaciones por restricción del hábitat.

Los estadísticos de resumen estimados para los datos observados (**Cuadro 6**) soportan la posibilidad de que las poblaciones hayan sufrido una reducción poblacional en el pasado basado en los resultados de la prueba de D de Tajima (Carlson et al., 2005; Tajima, 1989). Dentro del clado, ambas poblaciones mostraron valores similares para los diferentes

estadísticos de resumen calculados. La similitud en los resultados de los estadísticos de resumen obtenidos puede reflejar que las poblaciones de *P. melanomolga* y *P. gadovii* han experimentado historias demográficas similares.

Mastretta-Yanes et al. (2015) realizaron una compilación de estudios de patrones filogeográficos en las zonas montañosas mayores de México ((ENV), Sierra Madre Oriental (SMOr), Sierra Madre Occidental (SMOc), Sierra Madre del Sur (SMS) y Tierras altas de Chiapas-Guatemala) y concluyeron que los taxa exclusivos de zonas montañosas han experimentado una dinámica “sky-island” (aislamiento en zonas montañosas) y variaciones geológicas y climáticas a través del tiempo. Los mismos autores indican que estas variaciones climáticas y geológicas han promovido eventos recientes de divergencia y especiación en el ENV. El presente estudio revela la correlación entre la historia geológica del Este del ENV y la diversificación de especies en zonas montañosas (Cofre de Pero y Pico de Orizaba).

Los estudios descritos sobre la influencia de la historia geológica del ENV en la diversificación de especies aunado a los resultados obtenidos para el clado Veracruz indican que una de las causas principales de la alta diversidad en zonas montañosas en los trópicos es la complejidad de la historia geológica de la zona. En el presente estudio se demuestra que las poblaciones de *P. gadovii* y *P. melanomolga* presentan un patrón descrito en otras especies de baja dispersión con *Nolina parviflora* (Ruiz-Sanchez & Specht, 2013), *Poeciliopsis* y *Poeciliae* (Mateos, 2005). También se sugiere que la dinámica geológica ha sido una de las causas de la alta diversidad de pleurodontos en los trópicos. Al respecto Rovito et al. (2015) indica que el ENV alberga 41 especies de salamandras pleurodontas correspondientes a siete géneros (*Thorius*, *Chiropterotriton*, *Parvimolge*, *Aquiloerycea*, *Isthmura*, *Pseudoeurycea* y *Bolitoglossa*). De estos géneros, *Pseudoeurycea* es el más diverso con 13 especies (representando el 30% de especies totales del género).

Esta alta diversidad de *Pseudoeurycea* con respecto al resto de los géneros puede indicar la influencia que ha tenido la geología de esta región en la radiación de las Bolitoglossini. De manera general se sugiere que el aislamiento puede causar la divergencia de las especies en tiempos relativamente cortos (en este caso <200 ka.) en otras poblaciones de organismos. El estudio de Mateos (2005) concluye que las especies de los géneros *Poeciliopsis* y *Poeciliae* divergieron entre los 5 y 0.6 Ma en el ENV, un estudio realizado por He et al. (2016) demuestra el origen reciente del topo montañoso *Anourosorex squamipes* (0.6–0.1Ma) en las montañas del suroeste de China y que los periodos glaciares e interglaciares influyeron en la diversificación de las poblaciones de esta especie. Los trabajos de Mateos y He demuestran que en zonas montañosas, la divergencia de las poblaciones puede ocurrir en tiempos cortos (menores incluso a 1 Ma).

En resumen, el estudio de los modos geográficos de especiación nos permite elucidar la influencia que han tenido diferentes factores (como la dinámica geológica) en la diversificación de los organismos. Además, nos permite elucidar la ausencia o presencia de eventos demográficos en el pasado de las poblaciones que explique la variabilidad genética que observamos actualmente.

1.1 Escenarios alternativos de especiación alopátrica (peripátría) en el clado Veracruz

Debido a que la probabilidad posterior para el modelo vicariante no fue de 1.0, en este estudio considero la posibilidad de escenarios demográficos de especiación alternativos para las poblaciones de *P. melanomolga* y *P. gadovii*. Estos modelos alternativos describen el origen peripátrico para *P. melanomolga* a partir de una población de *P. gadovii* ($p=0.0793$) y viceversa ($p=0.0747$). En otras palabras, considero la posibilidad de que las especies en cuestión son el resultado de un aislamiento de un grupo reducido de individuos al margen de la distribución de la población ancestral (*P. melanomolga* o *P. gadovii* según sea el modelo) que debido a la fragmentación de bosque causada por las fluctuaciones de vegetación en el Pleistoceno tardío (Caballero et al., 2010) originaron un aislamiento posterior y como resultado del tamaño reducido de la población ancestral se originó un efecto fundador que causó el origen de las especies actuales micro endémicas de *Pseudoeurycea*. Sin embargo la prueba de hipótesis realizada en el presente estudio muestra una probabilidad posterior baja para cada uno de los modelos peripátricos.

2. Propiedades demográficas y posibles escenarios de especiación en el clado EdoMex

El cálculo del estadístico D de Tajima depende de la diversidad nucleotídica (π) y de la theta de Watterson (que a su vez depende del número de sitios segregantes y del número de muestras). Una baja diversidad nucleotídica o cercana a cero debido a que existen pocas diferencias en promedio entre las secuencias mientras que un valor bajo en el estimador de la theta de Watterson indica que existe un número bajo de sitios segregantes. El resultado de la prueba de D de Tajima muestra valores negativos para las poblaciones de *P. altamontana* y *P. robertsi* (especies hermanas de acuerdo al árbol filogenético estimado a partir del gen *cytb*). Valores negativos en este estadístico sugieren que la población ha experimentado un cuello de botella (Tajima, 1989). Entre este par de especies, la población de *P. robertsi* obtuvo un valor para la prueba de D de Tajima menor que la población de *P. altamontana* cuyo valor es más cercano a cero indicando que *P. robertsi* ha experimentado un cuello de botella reciente (Tajima, 1989).

Aunque no se realizaron simulaciones coalescentes para el clado EdoMex, los estadísticos de resumen obtenidos para este conjunto de poblaciones (*P. altamontana*, *P. longicauda* y *P. robertsi*) indican que la población de *P. robertsi* puede ser el resultado de un evento fundador como consecuencia de un cuello de botella, o que sufrió un cuello de botella muy severo después de su formación. Esta población muestra el menor número de loci ortólogos con sitios segregantes y con un menor número de sitios segregantes (**Cuadro 5**). Contrario a los valores obtenidos por las poblaciones hermanas de *P. altamontana* y *P. robertsi*, la población de *P. longicauda* obtuvo un valor positivo para el mismo estadístico, sugiriendo que esta población no ha experimentado una reducción significativa (cuello de botella) recientemente (Carlson et al., 2005; Tajima, 1989).

Considerando los valores obtenidos de las pruebas de F_{ST} pareadas para las poblaciones del clado EdoMex, la menor diferenciación poblacional es la existente entre las poblaciones de *P. altamontana* y *P. robertsi* mientras que el mayor grado de diferenciación entre poblaciones pertenece a *P. longicauda* y *P. robertsi*. La falta de diferenciación poblacional observada entre el primer par de especies sugiere una divergencia reciente entre estas poblaciones.

Basado en la historia geológica del centro del ENV propongo que al menos la divergencia (que sugiero que fue causada por un evento fundador) de *P. robertsi* pudo ocurrir en correspondencia con la formación y eventos volcánicos recientes del Nevado de Toluca (desde hace aproximadamente 2.3 Ma hasta el Holoceno). Sin embargo, considero que al tomar en cuenta 20 loci para el cálculo de los estadísticos de resumen para el clado EdoMex puede limitar la capacidad de inferencia sobre posibles escenarios demográficos. Además, en el presente estudio aunque el escrutinio para la selección del set de ortólogos a utilizar fue detallada (4.2 Mapeo y llamado de genotipos, 4.3 Edición y selección de ortólogos) no descarto la posibilidad de un posible sesgo en la identificación y llamado de SNPs y genotipos en la población de *P. robertsi*.

3. Datos masivos como fuente de marcadores moleculares

La generación de datos masivos que pueden ser usados como fuentes de marcadores moleculares para la resolución de escenarios e historias demográficas ha permitido indagar con mayor precisión eventos demográficos como cambios en el tamaño poblacional, migración o aislamiento entre poblaciones y divergencia o fusión de poblaciones (Carling & Brumfield, 2007; Knowles, 2004; Rittmeyer & Austin, 2015; Rosenberg & Nordborg, 2002). En comparación con el uso de uno o pocos marcadores moleculares, el uso de miles de marcadores moleculares independientes permite una estimación más precisa de los parámetros demográficos (Nadachowska-Brzyska et al., 2016; Rittmeyer & Austin, 2015; Robinson et al., 2014). En especies no modelo la generación de datos de escala genómica ha incrementado debido al desarrollo de tecnologías y herramientas que permiten el procesamiento de millones de datos moleculares (Robinson et al., 2014).

En el caso específico de salamandras, el tamaño del genoma oscila entre 10.12Gb (*Gyrinophilus porphyriticus*) y 120.60Gb (*Necturus lewesi*) (Gregory, 2017). Debido al gran tamaño y cantidad de secuencias repetitivas en los genomas de las salamandras, la secuenciación de genomas enteros del grupo no es factible todavía. Por lo tanto, la generación de transcriptomas como fuente de múltiples marcadores moleculares resulta una estrategia adecuada para la resolución de historias demográficas de las poblaciones. Para la especie del clado Veracruz y clado EdoMex, los transcriptomas ensamblados arrojaron un total de 25,000 y 34,600 sitios segregantes respectivamente. Este gran número de SNPs contenidos en más de 3000 ortólogos en el clado Veracruz permitió la estimación de los estadísticos de resumen (D de Tajima, theta de Watterson, pi) y la inferencia precisa (Robinson, Bunnefeld, Hearn, Stone, & Hickerson, 2014b) de los parámetros demográficos del modelo vicariante. Los transcriptomas obtenidos en el presente estudio pueden ser utilizados para la búsqueda de firmas de selección y para la selección de nuevos genes candidatos para el estudio de

inferencias filogenéticas en salamandras (e incluso vertebrados) diferentes a los utilizados actualmente.

Robinson et al. (2014) realizaron un estudio simulando datasets variando el número y longitud de los loci y el número de individuos para detectar el efecto de estos factores en la estimación de parámetros demográficos. Los autores observan que la precisión de la estimación mejora cuando se aumenta la longitud y el número de loci mientras que el aumento de número de individuos no mejora la estimación de los parámetros. En el presente estudio para el clado Veracruz utilicé un set de 3962 loci ortólogos con una longitud máxima de 1000 pb para un total de cinco individuos diploides por población para la estimación de parámetros demográficos. Los resultados de Robinson et al. (2014) la probabilidad posterior de elegir el modelo correcto alcanza una asíntota a partir de la utilización de 2000 loci con una longitud media de 1000 pb para cinco individuos por población. Tomando en cuenta el estudio anterior, considero que la elección de mejor modelo y la estimación de los parámetros es precisa y que el aumentar el número de individuos o el número de loci no afectará de manera significativa los valores obtenidos en el presente estudio.

4. Obtención de SNPs y ortólogos

Durante el llamado de genotipos (que incluye secuencias completas de los alelos de cada individuo) al utilizar ANGSD se puede obtener una tasa elevada de falsos negativos y falsos positivos (Singhal, 2013). Por otro lado cuando se pretende identificar únicamente los SNPs en un conjunto de secuencias la información acerca de los sitios invariables y su longitud permanece desconocida ya que los genotipos no son llamados (Singhal, 2013). En el presente estudio, al realizar el llamado de genotipos para el clado Veracruz se obtuvieron aproximadamente 65 mil sitios segregantes (datos no incluidos) en total mientras que al realizar el llamado de SNPs, el número de sitios variables disminuyó a aproximadamente 25 mil. En el caso del clado EdoMex, al realizar el llamado de genotipos se obtuvieron un total de 118,400 sitios segregantes mientras que al sólo incluir los sitios segregantes resultado del filtrado de SNPs el número total se redujo a un aproximado de 34,600.

La estrategia de llamar genotipos completos y SNPs por separado para después colapsarlos en un solo archivo que contenga genotipos completos y sólo sitios variables que se hayan detectado en el llamado de SNPs permite disminuir la tasa de error de tipo I ya que se evita conservar regiones aparentemente variables con una baja probabilidad (error tipo I). Debido a las grandes diferencias obtenidas durante el llamado de genotipos y de SNPs considero que es necesario probar la eficiencia de la integración de SNPs a genotipos descartando cualquier otro sitio segregante en los últimos para elucidar si la variación obtenida en los clado Veracruz y EdoMex es particular a este estudio o si la integración propuesta en este trabajo mitiga de manera general el error de tipo I (falsos positivos) en otros datasets derivados de estudios independientes.

Conclusiones

El clado Veracruz muestra una alta correspondencia entre la formación de los volcanes Cofre de Perote y Pico de Orizaba y la divergencia de las especies *P. melanomolga* y *P. gadovii* y que además estas especies pudieron originarse por un evento vicariante ($p=0.8460$) hace aproximadamente 185,000 años debido al surgimiento de los volcanes donde habitan. Sin embargo, es incierto si la divergencia de las poblaciones del clado Veracruz fue debida al aislamiento geográfico ocasionado por el surgimiento de los volcanes donde habitan. Un escenario alternativo sugiere que las poblaciones ancestrales recién aisladas en cada uno de los volcanes divergieron por efecto de la fragmentación del bosque y fluctuaciones climáticas ocurridas durante los periodos interglaciares ocurridos desde el Pleistoceno tardío al Holoceno y el Último Máximo Glacial.

La correspondencia del tiempo de divergencia de las salamandras del clado Veracruz con el tiempo de formación de los volcanes Cofre de Perote y Pico de Orizaba soporta la idea del efecto de los eventos geológicos sobre la divergencia de especies ocurrida por modos geográficos de especiación siendo en este estudio la vicarianza.

La utilización de datos masivos de escala genómica como marcadores moleculares permitió en el presente estudio una estimación precisa sobre los parámetros demográficos y el modo de especiación geográfica que originó las especies de salamandras del clado Veracruz basado en la diferencia de variaciones genéticas y estadísticas de resumen.

La población de *P. robertsi*, a pesar de compartir más de 10 mil transcritos con las poblaciones *P. altamontana* y *P. longicauda*, mostró sólo 20 loci con sitios segregantes lo que reflejó una baja diversidad nucleotídica. Además, los estadísticos de resumen observados para el clado EdoMex sugieren un posible origen peripátrico debido a un evento fundador como producto de un cuello de botella de *P. robertsi* a partir de *P. altamontana* y un tiempo de divergencia reciente.

Perspectivas

Con base en los resultados obtenidos en el presente estudio, considero como perspectivas el diseño y la optimización de los modelos demográficos para el clado EdoMex y la inclusión de los parámetros de migración y cambios en el tamaño efectivo poblacional en los modelos demográficos correspondientes al clado Veracruz y EdoMex para la realización de pruebas de hipótesis sobre el modo de especiación y eventos demográficos post-divergencia para ambos clados. La inclusión de los parámetros de migración y de cambios en el tamaño poblacional nos permitirá detallar más los procesos demográficos ocurridos en las especies de *Pseudoeurycea* que han originado la variabilidad genética observada en las poblaciones actuales.

Pseudoeurycea leprosa es una especie que cohabita en los sitios de las salamandras consideradas en el presente estudio (excepto con *P. longicauda*). Por tanto, esta especie ha experimentado los mismos eventos geológicos y climatológicos que las especies microendémicas. Esta especie de amplia distribución ha experimentado por tanto múltiples eventos de migración y colonización sin causar eventos de especiación aparentes. Debido a la historia demográfica de *P. leprosa*, considero generar datos transcriptómicos para seleccionar sets de ortólogos entre las poblaciones que cohabitan con las especies microendémicas del estudio y realizar pruebas de hipótesis bajo diferentes modelos demográficos. El uso poblaciones de *P. leprosa* para la realización de pruebas de hipótesis será con el objetivo de utilizar los resultados obtenidos como un tipo de control positivo de eventos de aislamiento poblacional que no producen divergencia en comparación con los eventos demográficos y geográficos que originaron a las especies microendémicas del presente estudio.

Para mitigar el efecto de acotar y utilizar el set de ortólogos con anotación, considero utilizar el transcriptoma sin anotar con mayor número de transcritos de clada clado como referencia en el proceso de búsqueda y selección de genotipos y ortólogos para la generación de datos simulados y la realización de prueba de hipótesis. Del mismo modo considero seleccionar sets aleatorios de diferentes números de ortólogos con el objetivo de probar el efecto de la utilización de 1 a miles de ortólogos en la realización de pruebas de hipótesis sobre modos de especiación geográfica y estimación de parámetros demográficos.

Actualmente existen diferentes herramientas para la inferencia demográfica de las poblaciones, por lo tanto considero la utilización de los software dadi (Gutenkunst, 2008) que utiliza el espectro de frecuencias de sitios (SFS por sus siglas en inglés) de loci bialélicos independientes y G-PhoCS que utiliza loci cortos con fase gamética indeterminada (“unphased haplotypes”) para realizar las mismas pruebas de hipótesis bajo los mismos modelos realizadas con ABC con el objetivo de reforzar y comprobar los resultados obtenidos.

Literatura Citada

- Aljanabi, S. M., & Martinez, I. (1997). Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Research*, 25(22), 4692–4693. <https://doi.org/10.1093/nar/25.22.4692>
- Andrews, S. (n.d.). FastQC A Quality Control Tool for High Throughput Sequence Data. Retrieved from <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>
- Balkau, B. J., & Feldman, M. W. (1973). Selection for migration modification. *Genetics*, 74(1), 171–174.
- Barracough, T., & Vogler, A. (2000). Detecting the Geographical Pattern of Speciation from Species-Level Phylogenies. *The American Naturalist*, 155(4), 419–434. <https://doi.org/10.1086/303332>
- Beaumont, M. A. (2010). Approximate Bayesian Computation in Evolution and Ecology. *Annual Review of Ecology, Evolution, and Systematics*, 41(1), 379–406. <https://doi.org/10.1146/annurev-ecolsys-102209-144621>
- Beaumont, M. A., Zhang, W., & Balding, D. J. (2002). Approximate Bayesian computation in population genetics. *Genetics*, 162(4), 2025–2035. [https://doi.org/Genetics December 1, 2002 vol. 162 no. 4 2025-2035](https://doi.org/Genetics%20December%201,%202002%20vol.%20162%20no.%204%202025-2035)
- Bertorelle, G., Benazzo, A., & Mona, S. (2010). ABC as a flexible framework to estimate demography over space and time: Some cons, many pros. *Molecular Ecology*, 19(13), 2609–2625. <https://doi.org/10.1111/j.1365-294X.2010.04690.x>
- Bi, K., Vanderpool, D., Singhal, S., Linderoth, T., Moritz, C., & Good, J. M. (2012). Transcriptome-based exon capture enables highly cost-effective comparative genomic data collection at moderate evolutionary scales. *BMC Genomics*, 13(1), 403. <https://doi.org/10.1186/1471-2164-13-403>
- Caballero, M., Lozano-García, S., Vázquez-Selem, L., & Ortega, B. (2010). Evidencias de cambio climático y ambiental en registros glaciales y en cuencas lacustres del centro de México durante el último máximo glacial. *Boletín de La Sociedad Geológica Mexicana*, 62(3), 359–377.
- Capra, L., Macías, J. L., & Garduño, V. H. (1997). The zitácuaro volcanic complex, michoacán, Mexico: Magmatic and eruptive history of a resurgent caldera. *Geofísica Internacional*, 36(3), 161–179.
- Carling, M. D., & Brumfield, R. T. (2007). Gene sampling strategies for multi-locus population estimates of genetic diversity (theta). *PloS One*, 2(1), e160. <https://doi.org/10.1371/journal.pone.0000160>

- Carlson, C. S., Thomas, D. J., Eberle, M. a, Swanson, J. E., Livingston, R. J., Rieder, M. J., & Nickerson, D. a. (2005). Genomic regions exhibiting positive selection identified from dense genotype data Genomic regions exhibiting positive selection identified from dense genotype data. *Genome Research*, *15*, 1553–1565. <https://doi.org/10.1101/gr.4326505>
- Carrasco-Nunez, G. (2000). Structure and proximal stratigraphy of Citlaltepētł Volcano (Pico de Orizaba), Mexico. *Special Paper - Geological Society of America*, *334*(JANUARY 2000), 247–262. <https://doi.org/10.1130/0-8137-2334-5.247>
- Carrasco-Nuñez, G., & Nelson, S. (1998). Edad y tasa de crecimiento del volcán Cofre de Perote (resumen). In *Primera Reunión Nacional de Ciencias de la tierra* (p. 41). Ciudad de México.
- Carson, H. L. (1970). Chromosome Tracers of the Origin of Species. *Science*, *168*, 1414–1418. <https://doi.org/10.1126/science.168.3938.1414>
- Carson, H. L., & Templeton, A. R. (1984). Genetic Revolutions in Relation to Speciation Phenomena: The Founding of New Populations. *Annual Review of Ecology and Systematics*, *15*(1984), 97–131.
- Chomczynski, P. (1993). A reagent for the single-step simultaneous isolation of RNA, DNA and proteins from cell and tissue samples. *BioTechniques*, *15*(3), 532–537. <https://doi.org/10.1038/news.2010.498>
- Coyne, J. A., & Orr, H. A. (2004). *Speciation*. Sunderland, MA: Sinauer Associates.
- Csilléry, K., Blum, M. G. B., Gaggiotti, O. E., & François, O. (2010). Approximate Bayesian Computation (ABC) in practice. *Trends in Ecology and Evolution*, *25*(7), 410–418. <https://doi.org/10.1016/j.tree.2010.04.001>
- Csilléry, K., François, O., & Blum, M. (2012). Approximate Bayesian Computation (ABC) in R: A Vignette. *202.162.217.53*, 1–21. Retrieved from <ftp://202.162.217.53/CRAN/web/packages/abc/vignettes/abcvignette.pdf>
- Darwin, C. (1859). El Origen De Las Especies, 1–463. Retrieved from [http://www.uruguayeduca.edu.uy/Userfiles/P0001%5CFile%5COrigen de las especies.pdf](http://www.uruguayeduca.edu.uy/Userfiles/P0001%5CFile%5COrigen%20de%20las%20especies.pdf)
- DeSalle, R. (1995). Molecular approaches to biogeographic analysis of Hawaiian Drosophilidae. In W. L. Wagner & V. . Funk (Eds.), *Hawaiian biogeography: Evolution on a hot spot Archipelago* (pp. 72–89). Smithsonian.
- Felsenstein, J. (1981). Skepticism towards santa rosalia, or why are there so few kinds of animals?, *35*(1), 124–138.
- Ferrari, L., Orozco-Esquivel, T., Manea, V., & Manea, M. (2012). The dynamic history of the

- Trans-Mexican Volcanic Belt and the Mexico subduction zone. *Tectonophysics*, 522–523, 122–149. <https://doi.org/10.1016/j.tecto.2011.09.018>
- Fisher, R. a. (1930). The Genetical Theory of Natural Selection. *Genetics*, 154, 272. <https://doi.org/10.1038/158453a0>
- Freitas, S., Vavakou, A., Arakelyan, M., Drovetski, S. V., Crnobrnja-isailović, J., Kidov, A. a., ... Carretero, M. a. (2016). Cryptic diversity and unexpected evolutionary patterns in the meadow lizard, *Darevskia praticola* (Eversmann, 1834). *Systematics and Biodiversity*, 2000(January), 1–14. <https://doi.org/10.1080/14772000.2015.1111267>
- Fu, Y. X., & Li, W. H. (1999). Coalescing into the 21st century: An overview and prospects of coalescent theory. *Theoretical Population Biology*, 56(1), 1–10. <https://doi.org/10.1006/tpbi.1999.1421>
- Futuyma, D. (2014). *Evolution*. SINAUER ASSOCIATES, INC. <https://doi.org/10.1007/s13398-014-0173-7.2>
- García-Palomo, A., Macías, J. L., Arce, J. L., Capra, L., Garduño, V. H., & Espíndola, J. M. (2002). Geology of Nevado de Toluca Volcano and surrounding areas , central Mexico. *Geological Society of America Map and Chart Series, MCH089*, 26.
- García-Palomo, A., Macías, J. L., & Garduño, V. H. (2000). Miocene to recent structural evolution of the Nevado de Toluca volcano region, Central Mexico. *Tectonophysics*, 318(1–4), 281–302. [https://doi.org/10.1016/S0040-1951\(99\)00316-9](https://doi.org/10.1016/S0040-1951(99)00316-9)
- Gillespie, R. ., & Croom, H. . (n.d.). Hawaiian Biogeography: Evolution on a Hot-Spot Archipelago. In W. L. Wagner & V. . Funk (Eds.). Washington, D.C: Smithsonian Institution Press.
- Gottlieb, L. (1974). Genetic confirmation of the origin of *Clarkia lingulata*. *Evolution*, 28(2), 244–250. <https://doi.org/10.2307/2407325>
- Gregory, R. T. (2017). Animal Genome Size Database. Retrieved from <http://www.genomesize.com>
- Harrison, R. G. (1991). Molecular Changes at Speciation. *Annual Review of Ecology and Systematics*, 22(1), 281–308. <https://doi.org/10.1146/annurev.es.22.110191.001433>
- He, K., Hu, N.-Q., Chen, X., Li, J.-T., & Jiang, X.-L. (2016). Interglacial refugia preserved high genetic diversity of the Chinese mole shrew in the mountains of southwest China. *Heredity*, 116(1), 23–32. <https://doi.org/10.1038/hdy.2015.62>
- Houck, L. D. (1975). *Reproductive Patterns in Neotropical Salamanders*. University of California Berkeley.

- Hudson, R. R. (2002). Generating samples under a Wright-Fisher neutral model of genetic variation. *Bioinformatics*, *18*(2), 337–338. <https://doi.org/10.1093/bioinformatics/18.2.337>
- Ikeda, H., Carlsen, T., Fujii, N., Brochmann, C., & Setoguchi, H. (2012). Pleistocene climatic oscillations and the speciation history of an alpine endemic and a widespread arctic-alpine plant. *New Phytologist*, *194*(2), 583–594. <https://doi.org/10.1111/j.1469-8137.2012.04061.x>
- Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., ... Drummond, A. (2012). Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, *28*(12), 1647–1649. <https://doi.org/10.1093/bioinformatics/bts199>
- Kingman, J. F. C. (1982). The coalescent. *Stochastic Processes and Their Applications*, *13*(3), 235–248. [https://doi.org/10.1016/0304-4149\(82\)90011-4](https://doi.org/10.1016/0304-4149(82)90011-4)
- Knowles, L. L. (2004). The burgeoning field of statistical phylogeography. *Journal of Evolutionary Biology*, *17*(1), 1–10. <https://doi.org/10.1046/j.1420-9101.2003.00644.x>
- Korneliussen, T. S., Albrechtsen, A., & Nielsen, R. (2014). ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics*, *15*(1), 356. <https://doi.org/10.1186/s12859-014-0356-4>
- Lanfear, R., Calcott, B., Ho, S. Y. W., & Guindon, S. (2012). PartitionFinder: Combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution*, *29*(6), 1695–1701. <https://doi.org/10.1093/molbev/mss020>
- Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat Methods*, *9*(4), 357–359. <https://doi.org/10.1038/nmeth.1923>
- Lawson, L. P., Bates, J. M., Menegon, M., & Loader, S. P. (2015). Divergence at the edges: peripatric isolation in the montane spiny throated reed frog complex. *BMC Evolutionary Biology*, *15*, 128. <https://doi.org/10.1186/s12862-015-0384-3>
- Lewis, H. (1973). The origin of diploid neospecies in *Clarkia*. *American Naturalist*, *107*(954), 161–170. <https://doi.org/10.1086/282824>
- Lewis, H., & Roberts, M. R. (1956). The origin of *Clarkia Ungulata*. *Evolution*, *10*(2), 126–138.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., ... Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>

- Lombard, R. E., & Wake, D. B. (1986). Tongue evolution in the lungless salamanders, Family Plethodontidae:IV. Phylogeny of plethodontid salamanders and the evolution of feeding dynamics. *Systematic Zoology*, 35(4), 532–551.
- Luna Vega, I., Morrone, J. J., & Espinoza Organista, D. (2007). *Biodiversidad de la Faja Volcánica Transmexicana*. (I. Luna Vega, J. J. Morrone, & D. Espinoza Organista, Eds.). Ciudad de Mexico: Comisión Nacional para el Conocimiento y Uso de la Biodiversidad.
- Macher, J. N., Rozenberg, A., Pauls, S. U., Tollrian, R., Wagner, R., & Leese, F. (2015). Assessing the phylogeographic history of the montane caddisfly *Thremma gallicum* using mitochondrial and restriction-site-associated DNA (RAD) markers. *Ecology and Evolution*, 5(3), 648–662. <https://doi.org/10.1002/ece3.1366>
- Mastretta-Yanes, A., Moreno-Letelier, A., Piñero, D., Jorgensen, T. H., & Emerson, B. C. (2015). Biodiversity in the Mexican highlands and the interaction of geology, geography and climate within the Trans-Mexican Volcanic Belt. *Journal of Biogeography*, 42(9), 1586–1600. <https://doi.org/10.1111/jbi.12546>
- Mateos, M. (2005). Comparative phylogeography of livebearing fishes in the genera *Poeciliopsis* and *Poecilia* (Poeciliidae: Cyprinodontiformes) in central Mexico. *Journal of Biogeography*, 32(5), 775–780. <https://doi.org/10.1111/j.1365-2699.2005.01236.x>
- Mayr, E. (1942). Systematics and the Origin of Species. *Genome Biology*. <https://doi.org/10.1073/pnas.0502030102>
- Mayr, E. (1954). Change of genetic environment and evolution. *Evolution as a Process*. Retrieved from [http://www.sevin.ru/fundecology/speciation/mayr\(20\).pdf](http://www.sevin.ru/fundecology/speciation/mayr(20).pdf)
- Metcalf, S. E. (2006). Late quaternary environments of the northern deserts and central transvolcanic belt of Mexico. *Annals of the Missouri Botanical Garden*, 93(2), 258–273. [https://doi.org/10.3417/0026-6493\(2006\)93\[258:LQEOTN\]2.0.CO;2](https://doi.org/10.3417/0026-6493(2006)93[258:LQEOTN]2.0.CO;2)
- Miller, M. A., Pfeiffer, W., & Schwartz, T. (2010). Creating the CIPRES Science Gateway for inference of large phylogenetic trees. *2010 Gateway Computing Environments Workshop, GCE 2010*. <https://doi.org/10.1109/GCE.2010.5676129>
- Moritz, C., Schneider, C. J., & Wake, D. B. (1992). Evolutionary relationships within the *Desmognathus eschscholtzii* complex confirm the ring species interpretation. *Systematic Biology*, 41(3), 273–291. <https://doi.org/10.1093/sysbio/41.3.273>
- Mulcahy, D. G., Morrill, B. H., & Mendelson, J. R. (2006). Historical biogeography of lowland species of toads (*Bufo*) across the trans-Mexican neovolcanic belt and the isthmus of Tehuantepec. *Journal of Biogeography*, 33(11), 1889–1904. <https://doi.org/10.1111/j.1365-2699.2006.01546.x>
- Nadachowska-Brzyska, K., Burri, R., Smeds, L., & Ellegren, H. (2016). PSMC analysis of

- effective population sizes in molecular ecology and its application to black-and-white *Ficedula* flycatchers. *Molecular Ecology*, 25(5), 1058–1072. <https://doi.org/10.1111/mec.13540>
- Osete, M. L., Ruiz-Martínez, V. C., Caballero, C., Galindo, C., Urrutia-Fucugauchi, J., & Tarling, D. H. (2000). Southward migration of continental volcanic activity in the Sierra de Las Cruces, Mexico: Palaeomagnetic and radiometric evidence. *Tectonophysics*, 318(1–4), 201–215. [https://doi.org/10.1016/S0040-1951\(99\)00312-1](https://doi.org/10.1016/S0040-1951(99)00312-1)
- Osozawa, S., Su, Z. H., Oba, Y., Yagi, T., Watanabe, Y., & Wakabayashi, J. (2013). Vicariant speciation due to 1.55Ma isolation of the Ryukyu islands, Japan, based on geological and GenBank data. *Entomological Science*, 16(3), 267–277. <https://doi.org/10.1111/ens.12037>
- Otte, D., & Society, O. (1994). *The crickets of Hawaii : origin, systematics and evolution*. The Orthopterists' society.
- Parra-Olea, G. (2002). Molecular phylogenetic relationships of neotropical salamanders of the genus *Pseudoeurycea*. *Molecular Phylogenetics and Evolution*, 22(2), 234–246. <https://doi.org/10.1006/mpev.2001.1048>
- Parra-Olea, G., Rovito, S. M., García-París, M., Maisano, J. A., Wake, D. B., & Hanken, J. (2016). Biology of tiny animals: three new species of minute salamanders (Plethodontidae: Thorius) from Oaxaca, Mexico. *PeerJ*, 4, e2694. <https://doi.org/10.7717/peerj.2694>
- Parra-Olea, G., Windfield, J. C., Velo-Antón, G., & Zamudio, K. R. (2012). Isolation in habitat refugia promotes rapid diversification in a montane tropical salamander. *Journal of Biogeography*, 39(2), 353–370. <https://doi.org/10.1111/j.1365-2699.2011.02593.x>
- Patarnello, T., Bargelloni, L., Varotto, V., & Battaglia, B. (1996). Krill evolution and the Antarctic ocean currents: evidence of vicariant speciation as inferred by molecular data. *Marine Biology*, 126(4), 603–608. <https://doi.org/10.1007/BF00351327>
- Pavlidis, P., Laurent, S., & Stephan, W. (2010). MsABC: A modification of Hudson's ms to facilitate multi-locus ABC analysis. *Molecular Ecology Resources*, 10(4), 723–727. <https://doi.org/10.1111/j.1755-0998.2010.02832.x>
- Pereira, S. L., & Baker, A. J. (2014). Vicariant Speciation of Curassows (Aves , Cracidae): A Hypothesis Based on Mitochondrial DNA Phylogeny VICARIANT SPECIATION OF CURASSOWS (AVES , CRACIDAE): A HYPOTHESIS BASED ON MITOCHONDRIAL DNA PHYLOGENY, 121(3), 682–694.
- Rambaut, A. (2010). Tree figure drawing tool version 1.3. 1.
- Rittmeyer, E. N., & Austin, C. C. (2015). Combined next-generation sequencing and

- morphology reveal fine-scale speciation in Crocodile Skinks (Squamata: Scincidae: Tribolonotus). *Molecular Ecology*, 24(2), 466–483. <https://doi.org/10.1111/mec.13030>
- Robinson, J. D., Bunnefeld, L., Hearn, J., Stone, G. N., & Hickerson, M. J. (2014). ABC inference of multi-population divergence with admixture from unphased population genomic data. *Molecular Ecology*, 23(18), 4458–4471. <https://doi.org/10.1111/mec.12881>
- Rosenberg, N. A., & Nordborg, M. (2002). Genealogical Trees, Coalescent Theory and the Analysis of Genetic Polymorphisms. *Nature Reviews Genetics*, 3(5), 380–390. <https://doi.org/10.1038/nrg795>
- Rovito, S. M. (2010). Lineage divergence and speciation in the Web-toed Salamanders (Plethodontidae: Hydromantes) of the Sierra Nevada, California. *Molecular Ecology*, 19(20), 4554–4571. <https://doi.org/10.1111/j.1365-294X.2010.04825.x>
- Rovito, S. M., Parra-Olea, G., Recuero, E., & Wake, D. B. (2015). Diversification and biogeographical history of Neotropical plethodontid salamanders. *Zoological Journal of the Linnean Society*, 175(1), 167–188. <https://doi.org/10.1111/zoj.12271>
- Ruiz-Sanchez, E., & Specht, C. D. (2013). Influence of the geological history of the trans-mexican volcanic belt on the diversification of *Nolina parviflora* (Asparagaceae: Nolinoideae). *Journal of Biogeography*, 40(7), 1336–1347. <https://doi.org/10.1111/jbi.12073>
- Shaw, K. L. (2002). Conflict between nuclear and mitochondrial DNA phylogenies of a recent species radiation: what mtDNA reveals and conceals about modes of speciation in Hawaiian crickets. *Proceedings of the National Academy of Sciences of the United States of America*, 99(25), 16122–16127. <https://doi.org/10.1073/pnas.242585899>
- Shen, X. X., Liang, D., Chen, M. Y., Mao, R. L., Wake, D. B., & Zhang, P. (2016). Enlarged Multilocus Data set Provides Surprisingly Younger Time of Origin for the Plethodontidae, the Largest Family of Salamanders. *Systematic Biology*, 65(1), 66–81. <https://doi.org/10.1093/sysbio/syv061>
- Singhal, S. (2013). De novo transcriptomic analyses for non-model organisms: An evaluation of methods across a multi-species data set. *Molecular Ecology Resources*, 13(3), 403–416. <https://doi.org/10.1111/1755-0998.12077>
- Soler, M. (2002). *Evolución la base de la Biología. Proyecto Sur de Ediciones, S.L* (Vol. 1). <https://doi.org/10.1017/CBO9781107415324.004>
- Stamatakis, A. (2006). RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, 22(21), 2688–2690. <https://doi.org/10.1093/bioinformatics/btl446>

- Sunnåker, M., Busetto, A. G., Numminen, E., Corander, J., Foll, M., & Dessimoz, C. (2013). Approximate Bayesian Computation. *PLoS Computational Biology*, 9(1). <https://doi.org/10.1371/journal.pcbi.1002803>
- Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, 123(3), 585–595. <https://doi.org/PMC1203831>
- Tavaré, S., Balding, D. J., Griffiths, R. C., & Donnelly, P. (1997). Inferring coalescence times from DNA sequence data. *Genetics*, 145(2), 505–518.
- Taylor, P. J., Maree, S., Cotterill, F. P. D., Missouf, A. D., Nicolas, V., & Denys, C. (2014). Molecular and morphological evidence for a Pleistocene radiation of laminate-toothed rats (*Otomys*: Rodentia) across a volcanic archipelago in equatorial Africa. *Biological Journal of the Linnean Society*, 113(1), 320–344. <https://doi.org/10.1111/bij.12317>
- Templeton, A. R. (2006). *Population Genetics and Microevolutionary Theory*.
- University of California, Berkeley, Ca, U. (2017). Amphibiaweb. Retrieved from <http://amphibiaweb.org>
- Velo-Antõn, G., Parra, J. L., Parra-Olea, G., & Zamudio, K. R. (2013). Tracking climate change in a dispersal-limited species: Reduced spatial and genetic connectivity in a montane salamander. *Molecular Ecology*, 22(12), 3261–3278. <https://doi.org/10.1111/mec.12310>
- Wake, D. B. (1966). Comparative Osteology and Evolution of the Lungless Salamanders, Family Plethodontidae. *Memoirs of the Southern California Academy of Sciences*, 4, 1–111.
- Wake, D. B. (2012). Taxonomy of Salamanders of the Family Plethodontidae (Amphibia: Caudata). *Zootaxa*, 3484, 75–82.
- Wegmann, D., & Excoffier, L. (2010). Bayesian inference of the demographic history of chimpanzees. *Molecular Biology and Evolution*, 27(6), 1425–1435. <https://doi.org/10.1093/molbev/msq028>
- Wegmann, D., Leuenberger, C., & Excoffier, L. (2009). Efficient approximate Bayesian computation coupled with Markov chain Monte Carlo without likelihood. *Genetics*, 182(4), 1207–1218. <https://doi.org/10.1534/genetics.109.102509>
- Wright, S. (1943). Isolation by Distance. *Genetics*, 28(2), 114–138. <https://doi.org/Article>

Anexos

Anexo 1

Relación de ejemplares colectados en el Eje Neovolcánico

Clado Edomex						
ID	Especie	Número de catálogo	Index-Illumina	Trascriptoma (*-id.fasta)	Estado	Localidad
1	<i>P. longicauda</i>	SMR2460	2	SMR2460-1b2	Estado de México	Villa de Allende
2	<i>P. robertsi</i>	SMR2471	4	SMR2471-1d4	Estado de México	Zinacantepec
3	<i>P. robertsi</i>	SMR2468	7	SMR2468-1g7	Estado de México	Zinacantepec
4	<i>P. altamontana</i>	SMR2488	13	SMR2488-1m13	Estado de México	Ocuilan
5	<i>P. longicauda</i>	SMR2461	1	SMR2461-1a1	Estado de México	Villa de Allende
6	<i>P. altamontana</i>	SMR2489	9	SMR2489-1i9	Estado de México	Ocuilan
7	<i>P. altamontana</i>	SMR2487	8	SMR2487-1h8	Estado de México	Ocuilan
8	<i>P. altamontana</i>	SMR2490	10	SMR2490-1j10	Estado de México	Ocuilan
9	<i>P. longicauda</i>	SMR2459	15	SMR2459-1o15	Estado de México	Villa de Allende
10	<i>P. longicauda</i>	SMR2462	16	SMR2462-1p16	Estado de México	Villa de Allende
11	<i>P. longicauda</i>	SMR2463	18	SMR2463-1q18	Estado de México	Villa de Allende
12	<i>P. robertsi</i>	SMR2469	19	SMR2469-1r19	Estado de México	Zinacantepec
13	<i>P. robertsi</i>	SMR2470	20	SMR2470-1s20	Estado de México	Zinacantepec
14	<i>P. robertsi</i>	SMR2472	21	SMR2472-1t21	Estado de México	Zinacantepec
15	<i>P. robertsi</i>	SMR2473	22	SMR2473-1u22	Estado de México	Temascaltepec
16	<i>P. altamontana</i>	SMR2500	23	SMR2500-1v23	Estado de México	Ocuilan
17	<i>P. longicauda</i>	SMR2465	1	SMR2465-2a1	Estado de México	Villa de Allende

Clado Veracruz						
ID	Especie	Número de catálogo	Index-Illumina	Trascriptoma (*-id.fasta)	Estado	Localidad
1	<i>P. melanomolga</i>	SMR2532	5	SMR2532-2e5	Veracruz	Las Vigas de Ramírez
2	<i>P. melanomolga</i>	SMR2542	10	SMR2542-2j10	Veracruz	Perote
3	<i>P. melanomolga</i>	SMR2544	11	SMR2544-2k11	Veracruz	Perote
4	<i>P. melanomolga</i>	SMR2545	12	SMR2545-2l12	Veracruz	Perote
5	<i>P. melanomolga</i>	SMR2546	13	SMR2546-2m13	Veracruz	Perote
6	<i>P. gadovii</i>	SMR2549	15	SMR2549-2o15	Veracruz	La Perla
7	<i>P. gadovii</i>	SMR2550	16	SMR2550-2p16	Veracruz	La Perla
8	<i>P. gadovii</i>	SMR2554	19	SMR2554-2r19	Puebla	Atzitzintla
9	<i>P. gadovii</i>	SMR2555	20	SMR2555-2s20	Puebla	Atzitzintla
10	<i>P. gadovii</i>	SMR2560	23	SMR2560-2v23	Puebla	Atzitzintla

Anexo 2

Preparación de soluciones

1. Clorotone

Disolver algunos cristales de clorobutanol hidratado (aproximadamente 1g) en 200 ml de etanol al 96 %. Agitar vigorosamente hasta disolver.

2. Formalina al 10 %

Solución	Volumen/Peso
Formaldehido (37-40 %)	100 ml
Fosfato de Sodio monobásico	4 g
Fosfato de Sodio dibásico	6.5 g
Agua destilada	900 ml

Se deben mezclar las sales con el agua destilada y finalmente homogenizar con los 100 ml de formaldehido. Reposar la solución por 24 hrs y almacenar a temperatura ambiente.

3. RNA Later

Preparar las siguientes soluciones:

- EDTA al 0.5M a pH 8 (18.61 g de EDTA disódico dihidratado/100 ml de agua, ajustar pH con NaOH).
- Citrato de Sodio (24.9 g de sal trisódica dihidratada/100 ml de agua desionizada).

Después se debe mezclar en un vaso de precipitado:

Solución	Volumen/Peso
40 ml de EDTA 0.5M, pH 8	40 ml
25 ml de Citrato de Sodio 1M	25 ml
Sulfato de Amonio pulverizado	700 g
Agua desionizada esteril	935 ml

Calentar y agitar hasta disolver el Sulfato de Amonio. Enfriar y ajustar pH a 5.2 con Ácido Sulfúrico al 1M. Esterilizar por calor húmedo y almacenar a temperatura ambiente o en refrigeración.

4. Buffer de extracción de ADN con sales

Solución	Volumen
Tris-Base-HCl pH 8, 1M	2 mL
NaCl 5M	4 mL
EDTA pH8, 0.5M	4 mL
Agua MiliQ	190 mL
Volumen final	200 mL

Anexo 3

Protocolo para la extracción de ADN con sales

Consumibles:

- Microtubos de 1.5 mL
- Puntas para micropipeta de 200 μ l y 1 mL
- Guantes de nitrilo
- Buffer de extracción de ADN con sales
- SDS al 10 %
- Proteinasa K
- NaCl 5M
- Isopropanol absoluto frío
- Etanol fresco al 80 %
- Agua miliQ

Nota: Algunas cantidades y tiempos pueden variar de acuerdo al organismo con el que se esté trabajando.

1. Corta finamente 2 a 4 mg de tejido dentro de un tubo de 1.5 mL.
2. Si la muestra se encontraba en etanol dejar evaporar por 1 a 3 minutos.
3. Agregar en el siguiente orden:
410 μ l de Buffer de extracción
80 μ l de SDS al 10 %
15 μ l de proteinasa K
4. Incuba la(s) muestra(s) a 55°C por 12 a 16 horas (mínimo 3hrs)
5. Centrifugar a 14000 rpm durante 5 minutos.
6. Transferir por decantación el sobrenadante a un tubo nuevo y agregar 180 μ l de NaCl

- 5M, mezclar por vórtex durante 20 segundos y centrifugar a 14000 rpm durante 20 minutos.
7. Transferir el sobrenadante a un tubo nuevo y agregar 500 μ l de isopropanol y mezclar por inmersión.
 8. Precipitar durante 8 hrs (3 hrs) mínimo a -20°C .
 9. Centrifugar a 14000 rpm por 10 minutos y eliminar sobrenadante por decantación (debe ser visible un pellet en el fondo del tubo).
 10. Agregar 500 μ l de etanol y precipitar por 1 hr (mínimo 30 minutos), centrifugar a 14000 rpm por 10 minutos.
 11. Descartar sobrenadante por pipeteo con cuidado de no sustraer el pellet de ADN y dejar secar por 15 a 20 minutos a 45°C .
 12. Resuspender en 50 μ l de agua miliQ y mezclar por pipeteo.
 13. Almacenar a -20°C .

Anexo 4

Protocolo de extracción de RNA con Trizol

Consideraciones previas

- Se debe tener si es necesario el peso del tejido u órgano del cual se realizará la extracción.
- Compuestos como Trizol, Cloroformo e Isopropanol se deben manejar en campana de extracción y con equipo de seguridad apropiado (Guantes, cubrebocas, y lentes de seguridad).
- Se deben mantener en refrigeración hasta el momento de su uso
- Se debe limpiar el área de trabajo y todo el equipo y material a utilizar con Rnase zap o ELIMINase o solución descontaminante de RNasas.
- Se debe mantener el tejido congelado en nitrógeno líquido hasta el momento de su uso.
- El etanol al 75 % debe prepararse utilizando agua ultrapura tratada con DEPC.

Consumibles

Microtubos de 1.5 mL debidamente rotulados (3 tubos por extracción)
Puntas para micropipeta de 10 μ l, 200 μ l y 1mL
Guantes de nitrilo
Trizol
Cloroformo
Isopropanol
RnaseZap
Etanol fresco al 75 %
Agua ultrapura tratada con DEPC

1. Descongelar el tejido a utilizar en hielo.
2. Añadir 500 μ l de trizol para 3-49 mg de tejido y 1 ml por cada 50 a 100 mg de tejido.
3. Homogenizar la muestra con trizol inmediatamente durante 1 minuto (tiempo máximo). Después de la homogenización con trizol, la muestra puede almacenarse por unas

- horas a temperatura ambiente o un mes a -70°C .
4. Una vez homogenizado con trizol, se considera centrifugar a $12000\times g$ durante 10 min a 4°C . *Esto se realiza cuando la muestra contenga un alto contenido de polisacáridos, proteínas o lípidos.
 5. Transferir el sobrenadante a un tubo nuevo. Nota: la pastilla formada contiene DNA de alto peso molecular y la capa superficial contiene lípidos.
 6. Incubar la muestra durante 5 minutos a temperatura ambiente.
 7. Agregar un 20 % de volumen de cloroformo del volumen de trizol agregado (ejemplo: 100 μl de cloroformo por cada 500 μl de trizol añadido).
 8. Agitar por inversión durante 15 seg. Aproximadamente.
 9. Incubar durante 3 minutos a temperatura ambiente.
 10. Centrifugar a $12000\times g$ durante 15 minutos a 4°C .
 11. Recuperar la fase acuosa (fase superior) por pipeteo con precaución de no acarrear la interfase o fase orgánica.
 12. Añadir a la fase acuosa el 50 % de volumen de isopropanol absoluto del volumen de trizol agregado (ejemplo: 250 μl de isopropanol por cada 500 μl de trizol añadido).
 13. Incubar 20 minutos a 4°C o en su defecto se puede dejar durante toda la noche.
 14. Centrifugar a $12000\times g$ durante 15 minutos a 4°C . Se debe observar un pequeño pellet o algo similar a un gel transparente.
 14. Remover sobrenadante y conservar únicamente el pellet.
 15. Lavar el pellet o pastilla con 1 ml de etanol al 75 %. Nota: El RNA puede almacenarse en etanol durante 1 año a -20°C o una semana a 4°C .
 16. Vortexear suavemente la pastilla con el etanol y centrifugar a $7500\times g$ durante 5 minutos a 4°C y descartar el sobrenadante.
 17. Secar parcialmente la pastilla por aireación.
 18. Resuspender la pastilla por pipeteo en 50 μl agua libre de RNasas o agua tratada con DEPC.
 19. Incubar en termoblock a $55-60^{\circ}\text{C}$ por 10 a 15 minutos. Y transferir 10 μl a un tubo de 200 μl (este ARN será el que se utilice para cuantificaciones y electroforesis, mientras que los 40 μl restantes permanecerán en congelación hasta el momento de ser utilizados para evitar su degradación)
 20. Almacenar a -70°C .

Anexo 5

Preparación y prueba de viabilidad de solución de perlas magnéticas SeraPure

- Preparación de solución de perlas magnéticas SeraPure como sustituto de SeraMag

Nota: La preparación de la solución de perlas SeraPure se realiza en campana de flujo laminar.

Consumibles:

Perlas magnéticas SeraMag
Polietilen-Glicol (PEG) 8000

Agua desionizada estéril
Tris 1 M, pH. 8 estéril
EDTA 0.5M estéril
Hcl 1N (para ajustar pH)
Tween 20
NaCl 5M estéril
Tubos falcon de 50 mL estériles.
Puntas para micropipéta (200 μ l, 1000 μ l).
Microtubos de 1.5 mL

1. En un tubo o probeta esteril de 50 ml preparar TE (Tris-Hcl 10 mM-EDTA 1mM= 500 μ l Tris pH 8, 1 M + 100 μ l EDTA 0.5M, llenar hasta 50 mL con agua desionizada) .
2. Homogenizar las perlas SeraMag usando un vórtex y transferir 1 mL a un microtubo de 1.5 mL y colocarlo en el stand magnético por 5 minutos o hasta que las perlas se hayan separado por completo y remover el sobrenadante sin perturbar o tocar las perlas.
3. Retirar las perlas del stand magnético, agregar 1 mL de TE, homogenizar por pipeteo y regresar el microtubo con la solución de perlas al stand magnético.
4. Esperar a que las perlas se hayan separado por completo y remover con cuidado el sobrenadante sin perturba rni tocar las perlas.
5. Repetir los pasos 3 y 4 para un total de 2 lavados de perlas con TE.
6. Retirar el microtubo del stand magnético y adicionar 1 mL de TE, homogenizar completamente por pipeteo y colocar el microtubo en una gradilla.
7. En un nuevo tubo falcon de 50 mL colcoar 9 g de PEG-8000, 10 mL de NaCl 5M 500 μ l de Tris-HCl 1M y 100 μ l de EDTA 0.5M.
8. Llena rle tubo falcon hasta 49 mL de agua desionizada y agitar usando vórtex hasta que el PEG-8000 se haya solubilizado por completo.
9. Adicionar 27.5 μ l de Tween 20 al tubo falcon y mezclar suavemente para evitar la formación de burbujas excesivas.
10. Agitar el mililitro de solución de perlas magnéticas-TE hasta homogenizar y tranferirlo al tubo falcon con la solución de PEG-8000, EDTA, NACL, Tris-HCl y Tween 20. Llenar hasta los 50 mL con agua desionizada si aún no esta en la marca y mezclar suavemente hasta homogenizar por completo.
11. Envolver el tubo con la solución de perlas con estaño o aluminio para protegerlo de la expocisión a la luz y almacenar a 4°C.

Nota. Se recomienda proceder inmediatamente a probar la viabilidad de la solución de perlas magnéticas y después realizar el test mensualmente.

- Prueba de viabilidad de solución de perlas magnéticas

Nota: Se recomienda utilizar un marcador de 50 pb para encontrar la relación volumen-volumen más precisa entre la muestra y la solución de perlas de acuerdo al fragmento de ADN/ARN que se desee purificar.

Consumibles

Solución de perlas SeraPure
Etanol fresco al 70 %
Marcador de peso molecular de 50 pb
Tubo falcon de 50 ml
Microtubos de 2 mL
Agua desionizada

1. Adicionar 1 μl de marcador de peso molecular (PM) por cada 9 μl de agua desionizada para obtener un marcador de PM 1:10 utilizar de acuerdo a la siguiente tabla

Marcador de PM 1:10	Solución de Perlas SeraPure	Relación volumen-volumen
10 μl	9 μl	0.7X
10 μl	12 μl	1.2X
10 μl	15 μl	1.5X
10 μl	18 μl	1.8X
10 μl	21 μl	2.1X
10 μl	24 μl	2.4X
4 μl	-----	-----

2. Preparar cada una de las soluciones de la tabla anterior en microtubos de 2 mL e incubar por 5 minutos a $^{\circ}\text{T}$ ambiente en una gradilla.
3. Pasados los 5 minutos colocar los tubos en un stand magnético y esperar hasta que las perlas se hayan seprado por completo y se hayan adherido a la pared de los tubos.
4. Remover y descartar sobrenadante por pipeteo sin perturbar las perlas.
5. Adicionar 500 μl de etanol al 70 % a cada uno de los tubos aún en el stand magnético sin tocar las perlas con la punta de la micropipeta e incubar durante 1 minuto.
6. Terminada la incubación, remove y descartar el etanol de cada uno de los tubos.
7. Repetir los pasos 5 y 6 para realizar un total de dos lavados con etanol al 70 %.
8. Retirar los tubos del stand magnético para colcoarlos en un termomixer a 37 $^{\circ}\text{C}$ durante 3 a 4 minutos o hasta que se hayan secado por completo.
9. Remover los tubos del termomixer y rehidratar con 20 μl de agua desionizada, homogenizar completamente por pipeteo y trasnferirlos al stand magnético por 5 minutos o hasta que las perlas se hayan separado por completo.

10. Remover el sobrenadante y colectarlo en un nuevo tubo (este sobrenadante contiene el ADN del marcador de peso molecular).
11. Realizar una electroforesis en gel de agarosa al 1.5 % con TBE utilizando 2 µl de cada sobrenadante para colcoar en los pozos. Correr la electroforesis durante 60 minutos a 100v. Utilizar el marcador de PM 1:10 como control en la visualización.

Anexo 7

Preparación de librerías de RNA con el Kit TruSeq RNA Library Preparation Kit v2 set A

Programar el termociclador con los siguientes programas.

- a) Tapa a 100°C
 - 65°C por 5 minutos, 4°C hold, volumen: 50 µl (guardar como mRNADENA)
 - 80°C por 2 minutos, 25°C hold, volumen: 25 µl (guardar como MRNAELU1)
 - 94°C por 6 minutos (4 minutos si el RNA está degradado, 8 si la claidad es muy alta), 4°C hold, volumen: 9.75 µl (Guardar como ELU2FRPR).
- b) Tapa a 100°C
 - 25°C por 10 minutos, 42°C por 50 minutos, 70°C por 15 minutos, 4°C hold, volumen: 12.5 µl (Guardar como 1STRAND).
- c) Temperatura de la tapa en OFF, 16°C por 1 hr, volumen final: 25 µl (Guardar como 2STRAND)
- d) Tapa 100°C
 - 30°C por 30 minutos, volumen: 50 µl (guardar como ENDR)
- e) Tapa a 100°C
 - 37°C por 30 minutos, 70°C por 5 minutos, 4°C hold, volumen: 15 µl (Guardar como ATAIL70).
- f) Tapa a 100°C
 - 30°C por 10 minutos, volumen: 18.75 µl (guardar como 3010MIN)
- g) Tapa a 100°C
 - 98°C por 30 segundos,
 - 98°C por 10 segundos,
 - 60°C por 30 segundos,
 - 72°C por 30 segundos,
 - 72°C por 5 minutos,
 - 10°C hold (Guardar como PCRLIB1)

} **13X**

Notas: Todas las centrifugaciones son a temperatura ambiente.

Se deben probar las perlas SeraPure antes de comenzar las librerías.

Se recomienda el uso de RnaseZap o semejante para descontaminar hasta antes de la síntesis de la 1ra cadena de ADN.

Se recomienda alicuotar los reactivos y soluciones del kit para evitar descogelarlos más de 4 veces.

El etanol al 80 % para los lavados debe ser frío y recién preparado.

Todas las homogenizaciones en microtubo de 200 μ l son por pipeteo suave hasta homogenizar completamente.

Mientras se realizan los lavados con etanol al 80 % los tubos permanecen en el stand magnético.

En cada paso al remover y descartar el sobrenadante de los tubos que contienen perlas se debe ser muy cuidadoso de no perturbar las perlas magnéticas y de no removerlas durante el pipeteo.

Es necesario contar con un stand magnético para tubos de 200 μ l.

Todas las incubaciones en el stand magnético se realizan a temperatura ambiente.

Protocolo:

A) Purificación y Fragmentación de ARN

Consumibles:

- Bead binding Buffer (BBB)
- Bead Washing Buffer (BWB)
- Elute, Prime Fragment Mix (EPF)
- Elution Buffer (ELB)
- Resuspension Buffer (RSB)
- RNA Purification Beads (RPB)
- Tubos de 200 μ l etiquetados RBT (RNA Bead Tubes).

Remover: BBB, BWB, ELB, EPF, RSB de congelación y/o refrigeración y colocarlos en hielo 5 minutos antes de utilizarlos.

Selección de Poly A

Nota: Pasos 1 y 2 se realizan en hielo.

1. Diluir 2 a 4 μ g de ARN total en 25 μ l de agua libre de RNasas en un tubo RBT etiquetado debidamente.
2. Vortexear vigorosamente las RNA purification Beads (RPB) hasta que se resuspendan por completo y adicionar 25 μ l de RPB a cada tubo (librería) que contiene el ARN, homogenizar completamente. Tapa los tubos y colócalos en el termociclador y selecciona el programa MRNADENA.
3. Alcanzado el hold, retira los tubos del termociclador y colócalos en la mesa a temperatura ambiente durante 5 minutos.
4. Terminada la incubación, abre los tubos y colócalos en el stand magnético por 5 minutos hasta que se observe que las perlas se han separado por completo.
5. Remueve y descarta el sobrenadante de cada tubo y remueve los tubos del stand magnético.
6. Lava las perlas de cada tubo agregando 100 μ l de Bead Washing Buffer (BWB) en cada tubo y homogenizar por completo por pipeteo. Coloca los tubos en el stand magnético por 5 minutos.
7. Mientras los tubos se encuentran en el stand magnético, alicuota el Elution Buffer (ELB) a utilizar y centrifúgalo por 5 segundos a 600xg y colócalo en hielo.
8. Terminada la incubación en el stand magnético, remueve y descarta el sobrenadante de

- cada tubo y retíralos del stand magnético.
9. Adiciona 25 μ l de ELB a cada tubo y homogeniza por pipeteo, tapa los tubos y colócalos en el termociclador y selecciona e inicia el programa MRNAELU1. (Almacenar el ELB restante a 4°C).
 10. Mientras los tubos están en el termociclador alicuota y centrifuga durante 5 segundos a 600xg el Bead Binding Buffer (BBB) a utilizar y almacena el resto a 4°C.
 11. Cuando el termociclador haya alcanzado el hold retira los tubos y colócalos en la mesa a temperatura ambiente y abre la tapa.
 12. Adiciona 25 μ l de BBB a cada tubo y homogeniza por pipeteo, incuba los tubos 5 minutos a temperatura ambiente.
 13. Al concluir la incubación, coloca los tubos en el stand magnético por 5 minutos o hasta que las perlas se hayan separado completamente y remueve y descarta el sobrenadante y retira los tubos del stand magnético.
 14. Lava las perlas con 100 μ l de BWB y homogeniza por pipeteo cuidando de no formar burbujas y coloca los tubos durante 5 minutos en el stand magnético. Almacena el BWB a 4°C.
 15. Después de la incubación, remueve y descarta el sobrenadante de cada tubo y remueve los tubos del stand magnético.
 16. Adiciona 9.75 μ l de EPF a cada tubo y homogeniza por pipeteo, tapa los tubos y colócalos en el termociclador e inicia el programa ELU2FRPR. Almacena el EPF a -20°C.
 17. Una vez alcanzado el hold en termociclador, retira los tubos, centrifúgalos brevemente (spin) y procede inmediatamente a la síntesis de la primera cadena de cADN.

B) Síntesis de la primera cadena de cADN.

Consumibles:

- First strand Master Mix (FSM)
- Super Script II-transcriptasa reversa (SSII)
- Tubos de 200 μ l etiquetados CDT (cDNA Tubes)

Remover: FSM de -20°C y llevarlo a temperatura ambiente en hielo.

1. Destapa los tubos RBT y colócalos en el stand magnético por 5 minutos.
2. Con los tubos en el stand magnético y transfiere 8.5 μ l de sobrenadante de cada tubo RBT a su tubo correspondiente CDP. Descartar los tubos RBT.
3. Centrifuga el FSM descongelado por 5 segundos a 600xg y adiciona 50 μ l de Super Script II. Si no se usa todo el FSM, adicionar 1 μ l de la enzima Super Script II por cada 9 μ l de FSM. homogenizar por pipeteo.
4. Adiciona 4 μ l de FSM+ SSII a cada tubo CDP y homogeniza por pipeteo, tapa los tubos y centrifuga rápidamente (spin). Almacena el FSM+SSII a -20°C.
5. Coloca los tubos en el termociclador e inicia el programa 1STRAND, cuando el programa llegue a la temperatura del hold, retira los tubos y procede inmediatamente a la síntesis de la segunda cadena de ADN.

C) Síntesis de la segunda cadena de ADNc

Consumibles:

Resuspension Buffer (RSB)

Second Strand Master Mix (SSM)

Etanol al 80%

Perlas magnéticas SeraPure

Tubos de 200 µl etiquetados IMT (Insert Modification Tubes)

1. Centrifuga el SSM descongelado por 5 minutos a 600xg, destapa los tubos CDP y adiciona 12.5 µl de SSM a cada tubo y homogeniza por pipeteo.
2. Coloca los tubos en el termociclador, cierra la tapa e inicia el programa 2STRAND. Almacena el SSM a -20°C.
3. Remueve los tubos del termociclador y colócalos en la mesa para que lleguen a temperatura ambiente y abre la tapa.
4. Vortexea las perlas SeraPure hasta que se homogenizen por completo y agrega 45 µl de perlas a cada tubo CDP que contiene 25 µl de cADN, homogeniza por pipeteo e incubalos por 5 minutos a temperatura ambiente.
5. Coloca los tubos CDP en el stand magnético y una vez que las perlas se han separado por completo desacrta 67.5 µl de sobrenadante de cada tubo.
6. Con los tubos en el stand magnético adiciona 100 µl de etanol al 80 % a cada tubo con cuidado, despues de 30 segundos, retira y descarta el etanol de los tubos. Repite este paso una vez más para realizar un total de 2 lavados con etanol.
7. Deja los tubos CDP en el stand magnético durante 10 a 15 minutos para que el etanol restante se evapore (mientras los tubos están en el stand magnético centrifuga el RSB a utilizar a 600xg durante 5 minutos), transcurrido el tiempo remueve los tubos del stand.
8. Adiciona 26.25 µl de RSB a cada tubo CDP y homogeniza por completo, incuba los tubos durante 2 a 3 minutos a temperatura ambiente.
9. Transfiere los tubos al stand magnético y espera a que las perlas magnéticas se separen por completo (aproximadamente 5 minutos) y transfiere 25 µl de sobrenadante a un tubo correspondiente IMT.

Punto de paro seguro:

Se pueden almacenar los tubos IMT a -20 °C hasta por 7 días si no se desea proceder inmediatamente a Reparación de final de cadenas.

D) Reparación de finales de cadenas

Consumibles:

End Repair Mix (ERP)

Resuspension Buffer (RSB)

Etanol al 80%

Perlas magnéticas SeraPure

Tubos de 200 µl etiquetados ALT (Adapter Ligation Tubes)

Remover: ERP de -20°C y llevarlo a temperatura ambiente, colocar en hielo.

Tubos IMT de -20°C y llevarlos a temperatur ambiente, centrifugar a 280xg durante 1 minuto.

Realizar IMT

1. Si no se utiliza el reactivo de control, añade 5 µl de RSB a cada tubo IMT que contiene 25 µl de ADNcds, adiciona también 20 µl de ERP y homogeniza.
2. Coloca los tubos en el termociclador y selecciona el programa ENDR. Remuevelos cuando el termociclador haya terminado la incubación y colcoalos en la mesa.

Limpieza IMT

3. Vortexea las perlas SeraPure hasta homogenizar por completo y adiciona 80 µl de perlas a cada tubo IMT y homogeniza completamente. Después, incuba los tubos a temperatura ambiente durante 5 minutos.
4. Coloca los tubos en el stand magnético y espera 5 minutos o hasta que las perlas se separen por completo y descarta 63.75 µl de sobrenadante de cada tubo dos veces (En total se descarta 127.5µl de sobrenadante de cada tubo).
5. Con los tubos aún en el stand magnético añade 100 µl de etanol al 80 % a cada tubo procurando que resbale sobre las perlas sin perturbarlas, después de 30 segundos retira el etanol por pipeteo evitando no susturar las perlas de los tubos. Realiza otro lavado con etanol por un total de dos lavados para cada tubo.
6. Deja secar los tubos durante 10 minutos aún en el stand magnético y después retira los tubos del stand y colcoalos en una gradilla.
7. Adiciona 8.75 µl de RSB a cada tubo y homogeniza completamente, incubalos durante 2 a 3 minutos en la gradilla y después colócalos en el stand magnético por 5 minutos o hasta que las perlas se hayan separado por completo.
8. Con los tubos en el stand magnético transfiere 7.5 µl de sobrenadante a cada tubo correspondiente ALT.

Punto de paro seguro

Si no se desea seguir inmediatamente con la Adenilación de extremos 3' se pueden almacenar los tubos ALT a -20°C hasta por 7 días.

E) Adenilación de extremos 3'

Consumibles:

A-Tailing Mix (ATL)

Resuspension Buffer (RSB)

Remover: ATL de -20°C y llevarlo a temperatura ambiente, mantener en hielo.

Tubos ALT de -20°C y llevarlos a temperatura ambiente; centrifuga los tubos a 280xg durante 1 minuto.

Adición de ATL e incubación 1 de tubos ALT

1. Si no se usa A-Tailing control, adicionar 1.25 µl de RSB y 6.25 µl de ATL a cada tubo y homogenizar completamente. Tapa los tubos, colocalos en el termociclador y selecciona el programa ATAIL70.

2. Cuando el programa llegue al hold retira los tubos y procede inmediatamente a la ligación de adaptadores.

F) Ligación de Adaptadores

Consumibles:

- Ligation Mix (LIG)
- Resuspension Buffer (RSB)
- RNA Adapter Index (1 diferente para cada librería)
- Stop Ligation Buffer (STL)
- Tubos de 200 μ l etiquetados: CAT (Clean Up ALT Tubes) y PCRT (Polymerase Chain Reaction Tubes)
- Etanol al 80 %
- Perlas SeraPure

Remover: RNA adapter index necesarios (de acuerdo a las librerías que se realizarán) de -20°C y llevarlos a temperatura ambiente. Mantener en hielo.

Adición de adaptadores por ligación

1. Una vez descongelados centrifuga los tubos de RNA index adapters y el STL a 600xg durante 5 segundos.
2. Si no se usa el Ligation Control adiciona 1.25 μ l de RSB a cada tubo ALT e inmediatamente remueve el LIG de -20°C y añade 1.25 μ l de ligation mix a cada tubo (regresa el LIG a -20°C), también adiciona 1.25 μ l de RNA adapter index a cada tubo (se debe tener un control de que número de índice lleva cada librería) y homogeniza completamente (regresa los adaptadores a -20°C).

Incubación 2 ALT

1. Coloca los tubos en el termociclador e inicia el programa 3010MIN, cuando finalice retira los tubos y colócalos en la mesa.

Adición de Stop Ligation Buffer y Limpieza de ATL

1. Destapa los tubos y añade 2.5 μ l de STL a cada tubo y homogeniza por completo.
2. Vortexea las perlas SeraPure hasta que se homogenizen por completo y adiciona 21 μ l a cada tubo ALT y homogeniza.
3. Incuba los tubos a temperatura ambiente por 15 minutos y después colócalos en el stand magnético por 5 minutos o hasta que las perlas se hayan separado por completo y remueve por pipeteo cuidadosamente y descarta 39.75 μ l de sobrenadante de cada tubo ATL (deja los tubos en el stand magnético para los lavados posteriores).
4. Realiza dos lavados con etanol al 80 % en cada tubo cuidando de no perturbar las perlas al momento de verterlo ni de removerlas al momento de retirar el etanol.
5. Deja secar los tubos aun en el stand magnético por 10 minutos y después colocalos en una gradilla.
6. Adiciona 26.25 μ l de RSB a cada tubo y homogeniza por completo. Incuba los tubos

- durante 2 a 3 minutos y después colócalos en el stand magnético.
7. Una vez que las perlas han separado por completo (~ 5 minutos) transfiere 25 μ l de sobrenadante de cada tubo ALT a su tubo correspondiente CAT con cuidado de no perturbar ni tocar las perlas (desecha los tubos ALT).
 8. Vortexea las perlas SeraPure hasta que se dispersen completamente y añade 25 μ l a cada tubo CAT y homogeniza por completo.
 9. Incuba los tubos durante 15 minutos y después transfírelos al stand magnético por 5 minutos para que las perlas se separen, después de esto, remueve y descarta 47.5 μ l de sobrenadante de cada tubo.
 10. Realiza dos lavados a cada tubo con 100 μ l de etanol al 80 % por cada lavado con cuidado para no perturbar ni tocar las perlas.
 11. Deja secar los tubos aún en el stand magnético por 10 minutos y después retíralos para adicionar 11.25 μ l de RSB a cada tubo CAT, homogeniza por pipeteo e incuba durante 2 a 3 minutos en una gradilla.
 12. Coloca los tubos en el stand magnético y espera por 5 minutos o hasta que las perlas se hayan separado completamente, después remueve y transfiere 5 μ l de sobrenadante a cada tubo PCRT (tubo A y tubo B para cada librería). Si se desea cuantificar se debe transferir a cualquiera de los tubos 1.2 μ l más de sobrenadante para que alguno de los tubos (A o B) tenga un volumen total de 6.2 μ l.

Punto de paro seguro

Se pueden almacenar los tubos PCRT a -20°C hasta por 7 días si no se desea continuar de inmediato al enriquecimiento de fragmentos por PCR.

G) Enriquecimiento de fragmentos de ADN

Consumibles:

PCR Master Mix (PMM)

PCR Primer Cocktail (PPC)

Resuspension Buffer (RSB)

Tubos etiquetados TST (Target Sample Tubes + número de index)

Etanol 80 %

Perlas SeraPure

Remueve: los tubos PMM y PPC de -20°C y llévalos a temperatura ambiente, centrifúgalos a 600xg durante 5 segundos, mantener en hielo.

Remueve los tubos PCRT de -20°C y llévalos a temperatura ambiente, centrifuga a 280xg durante 1 minuto y colócalos en hielo.

PCR de enriquecimiento

1. Añade 1.25 μ l de PPC a cada tubo y 6.25 μ l de PMM, homogeniza por completo y da un pequeño spin, coloca los tubos en el termociclador e inicia el programa PCRLIB1 (almacena inmediatamente el PMM y el PPC a -20°C).

Limpieza de productos de PCR

1. Cuando el termociclador haya alcanzado el hold retira los tubos PCRT y colocalos en la mesa.
2. Vortexea las perlas SeraPure y añade 12.5 μl a cada tubo homogeniznaod por completo. Incuba los tubos por 15 minutos y después transfierelos al stand magnético.
3. Una vez que las perlas se hayan separado completamente remueve y descrata 23.75 μl de sobrenadante de cada tubo.
4. Realiza dos lavados con 100 μl de etanol al 80 % a cad atubo como en las ocaisiones anteriores cuidando no perturbar d eninguna manera las perlas.
5. Deja secar los tubos en el stand magnético por 10 minutos y depsués remuevelos y colócalos en una gradilla para adicional 8.125 μl de RSB a cada tubo e incubalos por 2 a 3 minutos.
6. Tranfiere los tubos nuevamente al stand magnético por 5 minutos o hasta que se hayan separado completamente la sperlas y después remueve 7.5 μl de sobrenadante de cada tubo y colócalos en su tubo correspondiente TST.
7. Almacena los tubos TST a -20°C hasta que se proceda a realiza rla validación de las librerías.