UNIVERSIDAD MICHOACANA DE SAN NICOLÁS DE HIDALGO

FACULTAD DE BIOLOGÍA

PROGRAMA INSTITUCIONAL DE DOCTORADO EN CIENCIAS BIOLÓGICAS

OPCIÓN: CONSERVACIÓN Y MANEJO DE RECURSOS NATURALES

PATRONES FILOGEOGRÁFICOS EN PECES VIVÍPAROS DISTRIBUIDOS EN EL CENTRO DE MÉXICO

TESIS QUE PARA OBTENER EL GRADO DE DOCTORA EN CIENCIAS BIOLÓGICAS

Presenta

M. en C. Rosa Gabriela Beltrán López

Director de Tesis

Dr. Omar Domínguez Domínguez

Morelia, Michoacán. Septiembre de 2018.

AGRADECIMIENTOS

A mi director de tesis, el Dr. Omar Domínguez Domínguez por todo el apoyo recibido durante estos cuatro años, por darme total libertad, respaldo y confianza para llevar a cabo el proyecto ahora culmidado, y por impulsarme siempre a ser mejor y esforzarme más.

Al Dr. Ignacio Doadrio, por que aunque no tiene el título de Director de tesis, también lo fue, gracias por hacerme sentir como en mi casa el tiempo que estuve en el Museo Nacional de Ciencias Naturales, gracias también por darme total libertad de decisión y por abrirme las puertas de su laboratorio.

A mi comité tutoral: Dra. Sabina Irene Lara Cabrera, Dr. Antonio González Rodríguez y el Dr. Miguel Martínes Trujillo, por sus valiosos comentarios y correcciones durante este proceso, por que cuando proponíamos un nuevo capítulo, nunca dijeron que no, siempre tuvieron total confianza en mi, gracias por todo.

Al Dr. Rodolfo Pérez Rodríguez, por sus valiosos y acertados comentarios siempre que me sentía pérdida en algún capítulo, por la paciencia para escuchar y apoyarme, gracias.

A todos mis compañeros del laboratorio de Biología Acuática de la UMSNH, que siempre estuvieron presentes y que cuando algo me salía mal, siempre fueron los primeros en escucharme, además de todas las cosas divertidas y a veces nada divertidas que nos ha tocado vivir juntos, gracias por todo.

A la Universidad Autónoma del Estado de Morelos, el Centro de Investigaciones Biológicas y los integrantes del Laboratorio de Ictiología, por su total apoyo a este proyecto, por darme la oportunidad de llevarlo a cabo.

A las personas que integran el despacho del Dr. Doadrio y también algunos chicos de otros despachos, por que sin su ayuda y sin su amistad, me habría sentido pérdida en su país, muchas gracias por todos los momentos y por todo el apoyo, por mostrarme las cosas bellas de su país, espero que en algún momento de la vida podamos volver a coincidir.

A todos mis amigos que han seguido esta historia de cuatro años junto a mi, muchas gracias por existir en mi vida

A Conacyt por otorgarme la beca de Doctorado y beca mixta que hicieron posible estos estudios y mis estancias de investigación.

Finalmente, y no menos importante, a mi familia, a quien les agradezco profundamente que caminen conmigo de la mano, que cuando me caigo, siempre están para levantarme, y cuando soy féliz están ahí para compartir esa felicidad, sin ustedes, antes de empezar este proyecto habría renunciado, pero ustedes no lo permitieron, gracias por existir.

ÍNDICE

RESUMEN

La actividad volcánica y tectónica de placas desde el Mioceno y los ciclos glaciales e interglaciares durante el Plioceno y el Pleistoceno, son los principales eventos que han moldeado la distribución geográfica de la variación genética en varios taxa de peces de agua dulce del centro de México. En la presente tesis, se usaron marcadores mitocondriales (*cytb*, *coxI* y *dloop*) y nucleares (*S7*, *RAG1*, *RHO* y *B-ACT*) con enfoque filogenético, filogeográfico y de genética de poblaciones para inferir la historia evolutiva de las especies de peces vivíparos; *Xenotoca variata*, *Goodea* spp. (especies de importancia económico cultural purépecha), *Ilyodon* spp. (Goodeidae) y *Poeciliopsis infans* (Poeciliidae), a lo largo de su área de distribución en el centro de México. Los resultados muestran patrones filogeográficos contrastantes entre las especies analizadas. Las especies de *Xenotoca variata*, *Poeciliopsis infans* e *Ilyodon* spp., tienen historias evolutivas y biogeográficas complejas, y están fuertemente ligadas a la intensa actividad volcánica y tectónica durante el Plioceno y el Pleistoceno y a los ciclos glaciales-interglaciares durante el cuaternario. Para estas especies se encontró una fuerte estructura genética y alta diferenciación genética entre los linajes obtenidos, que puede estar ligada a la configuración antigua de las cuencas hidrológicas y no a la configuración actual. Debido a los linajes divergentes encontrados, se recomienda el re*Poeciliopsis infans* y *Xenotoca variata* es necesario realizar una revisión taxonómica por la probable presencia de especies taxonómicamente crípticas. En el presente trabajo se mostró que las características ecológicas y biológicas propias de cada especie, juegan un papel importante en la estructuración genética de las poblaciones, permitiendo en algunas la diferenciación genética, mientras que en otros casos, la homogeinización genética entre las poblaciones, como es el caso de *Goodea atripinnis*, la especie más ampliamente distribuida en las cuencas hidrológicas del centro de México. Se encontró baja diferenciación genética entre todas las poblaciones y haplotipos compartidos por varias regiones biogeográficas. Este patrón sugiere colonización reciente hacia la mayor parte de las cuencas hidrológicas donde la especie se distribuye, debido a su alta capacidad de dispersión y a que es una especie muy abundante y tolerante a los cambios ambientales. Se ha

reconocido que las especies sexualmente monomórficas generalmente presentan baja diferenciación genética entre las poblaciones debido a que no existe selección sexual fuerte como en las especies dimórficas. Además, para esta especie, la dispersión mediada por humanos es posible debido a que es de importancia para consumo humano desde el establecimiento de la cultura Purépecha (*ca*. 3000 años). En el presente trabajo, se mostró que además de la historia geológica del centro de México, las características biológicas y ecológicas de cada especie determinan el patrón filogeográfico encontrado, se encontraron resultados contrastantes entre las especies estudiadas; indicando que la historia geológica de la región en conjunto con las características biológicas y ecológicas propias de cada especie han moldeado la historia filogeográfica de cada una de estas de manera diferente.

PALABRAS CLAVE: Goodeinae, Poeciliinae, endémicos, Historia geológica, estructura genética

ABSTRACT

Volcanic and tectonic plates activities from the Miocene and glacial and interglacial cycles during the Pliocene and Pleistocene, are the main events that have shaped the geographic distribution of genetic variation in several taxa of freshwater fishes in central Mexico. In the present dissertation, we use mitochondrial (*cytb*, *coxI* y *dloop*) and nuclear markers (*S7*, *RAG1*, *RHO* y *B-ACT*) with phylogenetic, phylogeographic and population genetics frameworks to elucidate the evolutionary history of viviparous fish species *Xenotoca variata*, *Goodea* spp. (species of cultural and economic importance for the Purepecha culture), *Ilyodon* spp. (Goodeidae) and *Poeciliopsis infans* (Poeciliidae), along their distribution range in central Mexico. The results showed contrasting phylogeographic patterns among the analysed species. For the *Xenotoca variata*, *Poeciliopsis infans* and *Ilyodon* spp their evolutionary and biogeographical histories are complex, and strongly linked to the intense volcanic and tectonic activities during the Pliocene and Pleistocene and to Quaternary glacial-interglacial climate change. For these species we found strong genetic structure and high genetic differentiation between lineages that

could be linked to the ancient configuration of the hydrological Basins and not it's the current configuration. Due to the divergent lineages found, we encourage the rearrangement of the current taxonomic status for *Ilyodon* species, while *Poeciliopsis infans* and *Xenotoca variata* require a taxonomic review, given the probable existence of cryptic species. We showed that the ecological and biological characteristics of each species played an important role in the genetic structure of populations, allowing in some cases genetic differentiation, while in other cases, genetic homogenization between populations, as is the case of *Goodea atripinnis*, the most widely distributed species in the hydrological Basins of Central Mexico. Low genetic differentiation between all populations was found and shared haplotypes for several biogeographic regions. This pattern suggests a recent colonization towards most of the hydrological Basin where the species is distributed, due to the high dispersal capacity, its high abundance and its tolerance to environment changes. It has been recognized that sexually monomorphic species usually have low genetic differentiation among populations because sexual selection is not as important as in dimorphic species. Also, for this species, the dispersion mediated by humans is possible, due to this species its importance for human consumption since the establishment of the Purépecha culture (*ca*. 3000 years ago). We show that, in addition to the geological history of central México, biological and ecological characteristics of each species determine the phylogeographic pattern found, contrasting results between the studied species were found, indicating that the geological history of the region in conjunction with biological and ecological characteristics proper to each species have shaped the phylogeographical history of each in a different manner.

I. INTRODUCCIÓN GENERAL

I. I Historia evolutiva de los organismos

La biología evolutiva es una disciplina que se encarga de descubrir la diversidad biológica que existe en nuestro planeta. Una parte de la biología evolutiva trata de entender cómo la selección natural produce adaptaciones, mientras que otra parte infiere las relaciones y la historia de formación de las especies por medio de las genealogías y filogenias (Stearns y Hoekstra, 2000).

En 1859 Charles Darwin público su manuscrito titulado “El origen de las especies”, el cual es el cimiento en que descansa la teoría de la evolución, siendo esta teoría, la piedra angular de la biología moderna. Aunque Darwin no fue el primero en escribir sobre aspectos de la evolución, es a partir de él cuando se origina la teoría moderna de la evolución principalmente por dos razones: 1) examinó paciente y sistemáticamente todo tipo de evidencia relacionada con su tema, y 2) fue capaz de proporcionar un mecanismo plausible para explicar cómo las especies pueden cambiar: la selección natural, que de acuerdo a Darwin, puede influir en la evolución biológica (o lo que él llamó “descendencia con modificación”) (Darwin, 1859). A partir de esta fecha, la biología evolutiva ha tenido grandes avances metodológicos. En la actualidad, la historia evolutiva de las especies puede abordarse desde dos escalas diferentes, la microevolución y la macroevolución.

La microevolución describe los cambios genéticos que ocurren dentro de las poblaciones (periodos cortos de tiempo), a través de la selección natural: que puede ser selección natural positiva, selección natural negativa o selección natural balanceadora o disruptiva, dependiendo de los genotipos seleccionados en una población (Kelly, 2000; Templeton, 2006); la mutación: cambios en los genotipos de una generación que se transmiten a la descendencia y se van acumulando con el paso del tiempo generalmente de forma neutra (Templeton, 2006); flujo genético: intercambio de genotipos entre poblaciones distintas dentro de una misma especie como consecuencia del movimiento de individuos de unas poblaciones a otras, cambiando por tanto las frecuencias génicas de una generación a la siguiente (Allendorf, 1983) y la deriva genética: fijación de genotipos a lo largo del tiempo, como consecuencia del azar (Templeton,

2006). Para evaluar estas fuerzas, son importantes dos características de las poblaciones, la variación heredable de los rasgos y la variación en el éxito reproductivo entre individuos dentro de una población (Stearns y Hoekstra, 2000). Estos procesos provocan cambios en las frecuencias génicas o alélicas de una población.

Por su parte, la macroevolución describe e infiere los procesos evolutivos que ocurren a mayor escala evolutiva, es decir, aquellos que tienen lugar en grupos altos en la jerarquía taxonómica (especies, familias, órdenes, etc.) y que suponen un cambio evolutivo drástico dentro de los mismos, observable sólo después de que transcurra un largo periodo de tiempo (Dobzhansky, 1937; Rabosky y McCune, 2010). La macroevolución se encarga, por tanto, del estudio de las novedades evolutivas (Hallgrimsson et al., 2012; Shubin et al., 2009); los grandes patrones evolutivos como las tasas de cambio o las radiaciones evolutivas (Brawand et al., 2014; Harmon et al., 2003; Seehausen, 2004); la formación y extinción de linajes (Antonelli y Sanmartín. 2011; Burbrink y Pyron 2010; Mittelbach et al., 2007) y la disparidad morfológica entre grupos taxonómicos (Brusatte et al., 2011; Derek et al., 1992).

Una de las problemáticas en la biología evolutiva se da cuando el taxón se convierte en la unidad de estudio de entidades evolutivas, dificultando encontrar las fronteras y atributos de la unidad evolutiva, es decir el taxón, aún mas evidente, cuando la respuesta puede variar con el grupo de organismos de que se trate (Hey y Machado, 2003). Los conceptos de especie acuñados en los últimos 40 años son conflictivos y en ocasiones contradictorios, por lo que no existe un concepto único. En la biología evolutiva, el concepto filogenético de especie puede aportar resolución en este tema y se define como el agrupamiento mínimo de individuos, de poblaciones o grupos de poblaciones que son diagnosticables por un número dado de caracteres compartidos y dentro de los cuales hay un patrón de ancestría-descendencía (Cracraft, 1983; Davis y Nixon, 1992; McKitrick y Zink, 1988; Nixon y Wheler, 1990); es decir, la unidad taxonómica mínima que puede ser analizada desde un punto de vista filogenético.

Bajo este concepto, el uso de caracteres moleculares es de suma importancia en la delimitación de especies, en la actualidad, el uso de estos

caracteres se ha convertido en una herramienta importante, aunque no exclusiva, en la delimitación de especies (Avise y Ball, 1990).

Bajo este panorama los caracteres moleculares, sustentados en métodos filogenéticos (filogenias) o filogeográficos (tocogenias), se han convertido en una piedra angular para inferir la historia evolutiva de los organismos, siendo este un tema central en la biología (Godoy, 2009).

I. II Historia de la filogeografía y sus aplicaciones

El termino filogeografía fue acuñado por John C. Avise en el año de 1987 y la definió como el campo de estudio concerniente con los principios y procesos que gobiernan la distribución geográfica de los linajes genealógicos, especialmente aquellos dentro y entre especies cercanamente relacionadas (Avise et al., 1987; Avise, 2000). Esto quiere decir, que la filogeografía, considerada como una sub-disciplina de la biogeografía histórica, se ocupa de los componentes históricos y filogenéticos de la distribución espacial de los linajes genéticos, por lo que el tiempo y el espacio son los dos factores a considerar en los trabajos filogeográficos de las especies (Avise, 2000).

Los análisis filogeográficos y su interpretación usualmente integran conceptos y técnicas de genética molecular, genética de poblaciones, etología, demografía, biología filogenética, paleontología, geología y geografía histórica.

La filogeografía, por tanto, es una sub-disciplina integrativa, que engloba varias disciplinas microevolutivas (relaciones tocogenéticas) y macroevolutivas (relaciones jerárquicas). Dentro de las disciplinas microevolutivas se encuentran la etología, la demografía y la genética de poblaciones, entre otras, mientras que dentro de las disciplinas macroevolutivas se encuentran la paleontología y la biología filogenética (Avise, 2000).

Para llevar a cabo trabajos filogeográficos, la selección de los marcadores moleculares juega un papel primordial al inferir los patrones que rigen la distribución actual de los linajes genéticos. Desde la concepción de la filogeografía, el ADN mitocondrial en especies animales ha jugado un papel muy importante (Correa-Ramírez, 2010).

El uso del ADN mitocondrial se debe a sus características, entre ellas, es de herencia matrilineal, extensa variación intraespecífica y presenta altas

tasas de mutación (Lanteri y Confalonieri, 2003). Las variantes (secuencias) de ADNmt, conocidas como haplotipos, registran la historia matrilineal de eventos mutacionales, los cuales pueden conectarse de un modo filogenéticamente inteligible en un árbol de genes (Avise, 2000; 2008).

El uso exclusivo del ADNmt en algunos casos puede ser arriesgado, ya que implica un único locus, el cual puede estar ligado a selección, puede presentar introgresión o puede que no sea posible identificar su dispersión entre poblaciones como consecuencia de diferencias ecológicas o etológicas entre hembras y machos o entre especies (Avise, 2008; Domínguez-Domínguez et al., 2009). Lo mencionado anteriormente, puede implicar limitaciones al reconstruir las historias poblacionales (Avise, 2000), razones por las cuales cada vez es más frecuente el uso combinado y comparativo de genealogías obtenidas de datos de ADNmt y ADN nuclear (ADNn). Aunque el ADNn también puede presentar inconvenientes como son la recombinación, así como el riesgo de que cada locus, muestre historias evolutivas independientes (Zhang et al., 2005).

Recientemente, la secuenciación de nueva generación (NGS por sus siglas en inglés), ha facilitado el desarrollo de nuevos loci nucleares, es decir polimorfismos de nucleótidos simples (SNP por sus siglas en inglés) (Andres y Bogdanowicz, 2011). Los SNPs son considerados por muchos investigadores entre las herramientas más poderosas en estudios de genética de poblaciones (Guichoux et al., 2011). Los beneficios potenciales de otras aproximaciones con NGS en estudios filogeográficos han comenzado a incrementarse (Davey et al., 2011; Egan et al., 2012; Ekblom y Galindo, 2011; Emerson et al., 2010; Lexer et al., 2013; Zimmer y Wen, 2015), aunque actualmente la adopción de estos métodos han sido limitados como resultado de los desafíos metodológicos y computacionales (McCormack et al., 2013). Al menos por ahora, los estudios filogeográficos siguen dependiendo de la secuenciación tradicional de Sanger (Morris et al., 2018; Prince, 2015; Sanger y Coulson, 1975; Shaw et al., 2014).

Las aproximaciones filogeográficas son utilizadas principalmente en dos áreas de estudio, por un lado en la taxonomía, para el reconocimiento y establecimiento de límites entre especies y por otro lado, su uso específico en

biología de la conservación, sobre todo en la definición de unidades de conservación y manejo (Domínguez-Domínguez y Vázquez-Domínguez, 2009).

Los estudios filogeográficos han ayudado a reconocer y establecer límites entre las especies, de igual manera, la aplicación de métodos de análisis filogeográficos, ha sido una herramienta muy poderosa en estudios de biología de la conservación, permitiendo evaluar el potencial evolutivo de las especies. Además, los análisis filogeográficos han permitido evaluar hipótesis biogeográficas, describir procesos demográficos y evolutivos que resultan en unidades poblacionales diferenciables, así como inferir procesos que han determinado el origen, distribución y mantenimiento de la biodiversidad, información indispensable en taxonomía y conservación (Domínguez-Domínguez y Vázquez-Domínguez, 2009).

I. III El centro de México

El centro de México representa una zona de transición definida por Halffter (1987; 2003), como un área compleja en la cual los componentes bióticos Neotropicales y Neárticos se traslapan. Este traslape es particularmente evidente en la porción sur del Centro de México, en donde la Faja Volcánica Transmexicana (FVTM), caracteriza la zona orográficamente, resultado de extensos y complejos procesos geológicos (Halffter, 2003: Miller y Smith, 1986) y volcánicos desde el Mioceno temprano (hace aproximadamente 25 millones de años; Ferrari y Rosas-Elguera, 1999). La mayor actividad de la FVTM se desarrolló en los últimos 12 millones de años en la región central y oeste (Israde-Alcántara et al., 2010). Representa el arco del Neógeno más extenso en Norte América, abarcando 160,000 km^2 con una longitud de casi 1,000 km entre las coordenadas 18°30´ y 21°30´N (Ferrari et al., 2012).

Recientemente, esta zona ha recibido mucha atención por parte de los sistématas y biogeográfos (Beltrán-López et al., 2017; 2018; Domínguez-Domínguez et al., 2008a; 2016; Mateos et al., 2002; Mejía-Madrid et al., 2007; Pedraza-Lara et al., 2012; Pérez-Rodríguez et al., 2009; 2015) debido a la fuerte estructura genética que se ha encontrado relacionada a la historia geológica de la región.

Los ambientes dulceacuícolas de esta región del país han sido constantemente modificados debido a la dinámica e intensa actividad tectónica y volcánica de la FVTM; estos cambios, están reflejados en los procesos de génesis, destrucción, conexión, aislamiento y restructuración continua de las cuencas hidrológicas, incluidas las endorreicas, así como las cuencas asociadas con las vertientes del Pacífico y del Golfo de México, delimitadas por la Sierra Madre Occidental, la Sierra Madre Oriental y la Sierra Madre del Sur (Miller y Smith, 1986).

I. IV Subfamilia Goodeinae

I. IV. I Generalidades, biología y ecología de la subfamilia Goodeinae

Uno de los grupos mejor representados de peces endémicos para el Centro de México, son los miembros de la subfamilia Goodeinae, que es un grupo de pequeños peces vivíparos (de 40 a 180 mm de longitud estándar). Esta subfamilia está representada por aproximadamente 18 géneros y 41 especies, siendo muchas de sus especies endémicas a una cuenca específica o microendémica a un solo cuerpo de agua (Domínguez-Domínguez et al., 2010). La distribución de los goodeinos incluye drenajes a lo largo de las vertientes del Pacífico y del Golfo de México; en la vertiente del Pacífico se distribuyen desde la parte alta de la cuenca del Río Mezquital hasta las partes altas del sur de la cuenca del Río Balsas, mientras que en la vertiente del Golfo de México habitan en las partes altas de la cuenca del Río Pánuco. Además habitan drenajes interiores como el Río Aguanaval y los lagos de Cuitzeo, Pátzcuaro, Chapala y Zirahuén (Domínguez-Domínguez et al., 2010).

Se les conoce como mexclapiques, tiros y pintitos. Sus especies presentan diferencias importantes en su adaptacion a los hábitat más diversos: manantiales, humedales, lagos, arroyos, grandes ríos, canales y otros hábitat artificiales, principalmente en altitudes de 1000 a 2300 msnm (una especie habita hasta los 2800 msnm). La mayoría vive en aguas someras, por lo general a no más de 1 m de profundidad (Miller et al., 2005). Las especies de este grupo poseen rangos tróficos desde estrictamente carnívoros (*Alloophorus robustus* y *Allodonthichthys tamazulae*) a herbívoros (*Goodea atripinnis*), sin embargo, la mayoría de las especies poseen dietas similares, siendo omívoros,

y ocupan diferentes nichos dentro del espectro trófico del grupo (Soto-Galera et al., 1999).

La subfamilia tiene adaptaciones únicas para la reproducción, 1) estructura urogenital muscular (Nelson, 1975); 2) los primeros seis a ocho radios anales del macho, cortos y apiñados; 3) desarrollo de trofotenias complejas (placentas trofoténicas), estructuras relacionadas con la nutrición y la respiración (absorción de glucosa, aminoácidos y proteínas) en el embrión (Lombardi y Wourms, 1985); 5) espermatozeugmas (haces de espermatozoides; no espermatóforos), con las cabezas enterradas y las colas orientadas periféricamente, estructuras que no se encuentran en ningún otro grupo y 6) ausencia de órganos sensoriales, es decir la línea lateral, los cuales están presentes en todas las familias cercanas. El dimorfismo sexual es marcado en muchas de las especies; los machos suelen ser más pequeños y coloridos que las hembras, presentan un elaborado cortejo que involucra complejas exhibiciones de movimientos y danzas por parte de los machos (Domínguez-Domínguez y Pérez-Ponce de León, 2007), de esta manera, la selección sexual a través de la elección de pareja es una característica importante de este grupo.

Comparten caracteres osteológicos con dos géneros ovíparos actuales (*Empetrichthys* y *Crenichthys)* en el oeste de los Estados Unidos, estos dos géneros han sido incluidos dentro de la familia Goodeidae (Parenti, 1981), sin embargo, Miller et al. (2005), consideran que esos dos géneros constituyen la familia hermana Empetrichthyidae.

I. IV. II Relaciones filogenéticas y sistemática de la subfamilia Goodeinae

Se ha propuesto que el origen y evolución de la subfamilia Goodeinae, ha estado estrechamente vinculada a la compleja historia geológica y climática del Centro de México desde la época del Mioceno (16 millones de años atrás aproximadamente), factor que, junto con la peculiar forma de reproducción y selección sexual, así como la historia climática de la región, han sido ampliamente discutidos y considerados como factores determinantes de su extraordinaria radiación adaptativa (Domínguez-Domínguez et al., 2010; Pérez-Rodríguez et al., 2015). El fósil más antiguo de la subfamilia perteneciente a la época del Mioceno corresponde a la especie extinta *Tapatia occidentalis*

(Guzmán, 2015), especie que ya mostraba las especializaciones propias de la familia.

A pesar de considerarse un grupo ampliamente estudiado, investigaciones recientes usando técnicas moleculares han inferido valores altos de diversidad genética y divergencias entre poblaciones, proponiendo una subestimación en el número de especies dentro de la subfamilia (Beltrán-López et al., 2017; Doadrio y Domínguez 2004; Domínguez-Domínguez et al., 2008a; 2010; Piller et al. 2015), e incluso se han descrito especies nuevas (Domínguez-Domínguez et al., 2008b; 2016).

El trabajo filogenético y biogeográfico más completo hasta el momento para la subfamilia es el realizado por Domínguez-Domínguez et al. (2010), en el que incluyeron un total de 162 individuos de 138 poblaciones y 41 especies. En general, este trabajo apoya los resultados del estudio filogenético previamente reportado por Doadrio y Domínguez (2004), en el cual tres grupos de especies bien diferenciados fueron encontrados, las tribus Characodontini e Ilyodontini y un grupo que incluye a las especies restantes. Las hipótesis filogenéticas en este trabajo no soportaron la existencia de Godiini como un grupo monofilético en el que se incluía *Goodea* y *Ataenobius*.

Las tribus dentro de Goodeinae divergieron aproximadamente hace 14.05 millones de años (Foster y Piller, 2018). Un segundo evento cladogenético (hace diez millones de años aproximadamente) separó al ancestro del clado Ilyodontini en la parte oeste de las cuencas del Pacífico, a lo largo de las cuencas de los Ríos Ameca, Armería y Mascota-Purificación, del ancestro del resto de los godeinos, aislado en las partes altas del Centro de México. Aparentemente, como resultado de este evento cladogenético, el ancestro del resto de los goodeinos fue capaz de extender su rango de distribución, experimentando un proceso de diversificación a lo largo de la llamada Mesa Central de México (Domínguez-Domínguez et al., 2010). Estos eventos vicariantes, tienen correspondencia con el levantamiento de la Mesa Central de México durante el Mioceno.

Un trabajo realizado con la familia Goodeidae, mostró que dentro de estos, a comparación de la subfamilia Empetrichthyinae, la subfamilia Goodeinae tuvo una radiación adaptativa en el Centro de México en un periodo corto de tiempo, resultado de esto, son las altas tasas de evolución de la forma

corporal que presenta este grupo, ya que se ha mostrado que muchos de los miembros de Goodeinae colonizaron nuevas áreas debido a la alta variabilidad morfológica que presentan, mientras que los Empetrichthyinae no se expandieron hacia nuevos ambientes y no se diversificaron de la misma manera que los goodeinos, además, la evolución de la viviparidad en este grupo también pudo haber jugado un papel importante en su adaptación a nuevos ambientes (Foster y Piller, 2018).

Dentro de esta subfamilia, el género *Ilyodon*, cuya distribución geográfica está limitada a las cuencas de la vertiente del Pacífico del Centro de México (Ríos Balsas, Coahuayana, Armería, Ameca, Purificación y Marabasco), ha tenido una historia taxonómica controversial. Cinco especies han sido validamente descritas: *Ilyodon whitei* (Meek, 1904), *Ilyodon furcidens* (Jordan y Gilbert, 1882), *Ilyodon xantusi* (Hubbs y Turner, 1939), *Ilyodon lennoni* Meyer y Föerster, 1983, e *Ilyodon cortesae* Paulo-Maya y Trujillo-Jiménez, 2000. Sin embargo, diferentes investigaciones han diferido acerca del número de especies reconocidas. El catálogo de peces de Eschmeyer et al. (2016) incluye cuatro especies válidas (*Ilyodon whitei*, *Ilyodon furcidens*, *Ilyodon lennoni* e *Ilyodon cortesae*), mientras que en el trabajo de Doadrio y Domínguez (2004), basado en caracteres moleculares, fueron reconocidas solo dos especies (*Ilyodon whitei* e *Ilyodon furcidens*). Mientras que Domínguez-Domínguez et al. (2010), incluyendo todas las especies del género *Ilyodon*, con un solo gen mitocondrial, no encontraron grupos monofiléticos dentro del género. Se ha mostrado que este género exhibe niveles altos de plasticidad fenotípica, debido a diferentes morfotipos tróficos (Grudzien y Turner, 1984a,b; Turner y Grosse, 1980). Por lo anterior, un estudio que incluyera todas las especies identificadas a través de su rango de distribución y con secuencias de ADN mitoncondrial y nuclear era necesario para esclarecer la evolución y la taxonomía del género *Ilyodon* (Beltrán-López et al., 2017).

Por otra parte, dentro de la subfamilia Goodeinae, *Goodea* es uno de los géneros más ampliamente distribuidos (Miller et al., 2005) y ambientalmente tolerante (De la Vega-Salazar, 2006), con marcadas diferencias morfológicas entre poblaciones, y poco dimorfismo sexual. Al menos tres especies han sido descritas dentro de este género: *Goodea atripinnis* (Jordan, 1880) distribuido en la cuenca del Río Lerma, *Goodea luitpoldii* (Steindachner, 1894) distribuido

en el lago endorreico de Pátzcuaro, y *Goodea gracilis* (Hubbs y Turner, 1939), distribuido en la cuenca del Río Pánuco. A pesar del reconocimiento de estas tres especies, en el trabajo de Domínguez-Domínguez et al. (2010), con caracteres moleculares y un número reducido de muestras, concluyeron que *Goodea* está compuesto de una sola especie *Goodea atripinnis* (Domínguez-Domínguez et al., 2010; Webb et al., 2004), por lo que las variantes morfológicas parecen ser producto de una alta plasticidad fenotípica debida a una rápida respuesta al cambio ambiental (Foster et al., 2015). Por lo que un estudio que analizara el vínculo entre la historia evolutiva de *G. atripinnis* y la historia geológica y climática del centro de México, así como factores ecológicos y biológicos de la especie, incluyendo muestras de toda su área de distribución, era necesario para esclarecer la estructura y diversidad genética, así como la demografía histórica de sus poblaciones (Beltrán-López et al., *en preparación*).

Finalmente, la especie *Xenotoca variata*, es otra de las especies con amplia distribución dentro de la subfamilia. Esta capacidad de dispersión parece estar ligada a la amplia variedad de ambientes que habita, tolerante a diferentes características ecológicas y a diferentes condiciones ambientales (De la Vega Salazar, 2006; Miller et al., 2005). Sin embargo, se ha reconocido a *X. variata* como una especie altamente dimórfica con marcada selección sexual, en donde los machos realizan cortejos muy complejos en el que presentan a la hembra las características sexualmente dimórficas, particularmente el color y las aletas desplegadas (Fitzimons, 1976), lo que podría implicar una alta diferenciación genética entre sus poblaciones (Ritchie et al., 2005; 2007), a pesar de su alta capaciad de dispersion. Por lo que un trabajo filogeográfico usando genes mitocondriales y nucleares, incluyendo todas las poblaciones conocidas para esta especie, era necesario para evaluar el papel tanto de los eventos geológicos y climáticos en la diversificación de linajes, así como los aspectos ecológicos y etológicos de esta especie (Beltrán-López et al., *en preparación*).

I. V Subfamilia Poeciliinae

I. V. I Generalidades, biología y ecología de la subfamilia Poeciliinae

Los peces de esta subfamilia representan un grupo ampliamente distribuido y muy diverso. La subfamilia incluye de 22 a 29 géneros y más de 200 especies (Lucinda, 2003), tienen un tamaño entre 31 y 70 mm de longitud patrón (200 mm máximo; Miller et al., 2005). Los miembros de esta subfamilia son endémicos del nuevo mundo, con la mayoría de las especies distribuidas en México, América Central y las Antillas, aunque su distribución incluye desde el este de Estados Unidos hasta el noreste de Argentina (Berra, 2001). Representan uno de los cuatro grupos de peces del orden Cyprinodontiformes que han desarrolado la fecundación interna, siendo el gonopodio del macho un órgano intromitente representado por la modificación de los radios anales 3-5. Estos caracteres gonopodiales constituyen un caracter diagnóstico con fines taxonómicos, incluso la correcta identificación de ciertas especies requiere del examen minucioso de esta estructura gonopodial. Este grupo representa uno de los tres grupos que han evolucionado de manera independiente a la viviparidad, dando a luz a sus crías vivas (Parenti, 1981). A los peces de este grupo se les conoce como guayacones, topotes, espadas, gupis y molis y representan uno de los grupos de peces dominantes en aguas dulces y salobres, principalmente en zonas tropicales.

Las hembras grávidas pueden producir varias camadas sucesivas estando aisladas de los machos hasta por 10 meses o más, un proceso llamado retención de esperma. Algunas especies pueden presentar superfetación, de modo que dos o más camadas en diferentes etapas de desarrollo coexisten en una sola hembra. Esto ha evolucionado probablemente de manera independiente varias veces en esta subfamilia (Reznick y Miles, 1989). Muchas especies de esta subfamilia, entre ellas los platis y colas de espada (*Xiphophorus* spp.), el gupi (*Poecilia reticulata*) y los molis (*Poecilia*), son peces de acuario muy populares. El guayacón mosquito (*Gambusia affinis* y sus parientes cercanos) se ha introducido en muchas regiones del mundo como control de las larvas de mosquitos, pero cuando se le introduce fuera de su ámbito natural puede tener efectos negativos en la ictiofauna nativa (Miller et al., 2005).

I.V. II Relaciones filogenéticas y sistemática de la subfamilia Poeciliinae

Los Poeciliinos han jugado un papel importante en estudios biogeográficos debido a su amplia distribución y diversidad en Centro América y el Caribe (Rosen, 1975). Los poeciliinos se han agrupado en cinco grupos de acuerdo a su distribución geográfica y a las relaciones filogenéticas: Sur América, que incluye los géneros *Tomeurus*, *Phalloceros*, *Phallotorynus*, *Phalloptychus*, *Cnesterodon*, y *Poecilia*; Centro América, que incluye los géneros *Alfaro*, *Phallichthys*, *Carlhubbsia*, *Belonesox*, *Priapella*, *Xiphophorus*, *Pseudoxiphophorus*, *Poecilia*, *Gambusia*, *Xenodexia*, *Scolichthys* y *Poeciliopsis*; la parte sur de Centro América, que incluye los géneros *Brachyrhaphis*, *Priapichthys*, *Neoheterandria*, *Poeciliopsis*, *Poecilia* y *Gambusia*; Norte América, que incluye los géneros *Heterandria, Poecilia* y *Gambusia*; y las grandes Antillas, que incluye los géneros *Gambusia*, *Limia*, *Quintana* y *Girardinus*. La distribución geográfica de los poeciliinos también tiene correspondencia con linajes monofiléticos, aunque el linaje de Sur América es parafilético, mientras que *Gambusia* y *Poecilia* son los géneros más ampliamente distribuidos. Durante el Cretácico, parece haber existido una conexión entre Centro y Sur América, permitiendo dispersión bidireccional entre el Norte y el Sur por parte de los Poeciliinos, que representan un antiguo grupo que inicialmente estuvo restringido a Sur América, la cuál es su área de origen, experimentando después vicarianza que condujo a una radiación adaptativa en Centro América, las Grandes Antillas y la parte sur de Centro América, seguida de eventos de dispersión a las regiones de Sur América (Hrbek et al., 2007).

Se ha establecido que la mayor colonización de Centro y Norte América probablemente ocurrió a través de las Grandes Antillas durante la transición del Eoceno al Oligoceno (67.98 millones de años aproximadamente). La mayor radiación de este grupo tuvo lugar hace aproximadamente 44 millones de años, lo cual es compatible con una segunda invasión de Centro y Norte América a través del puente de las aves en las Grandes Antillas hace 35-33 millones de años (Iturralde-Vinent y MacPhee, 1999). Después de estas colonizaciones, varios eventos de colonización posteriores han sido documentados para diferentes géneros, mostrando que el escenario más probable para la historia evolutiva de los poeciliinos debe ser un origen en Sur América con recurrentes eventos de dispersión a Centro y Norte América, siendo los puentes terrestres

un factor crucial en la diversificación de este grupo (Hrbek et al., 2007). Los análisis filogenéticos han mostrado que los géneros *Poeciliopsis* y *Neoheterandria* son el grupo hermano del clado mas sureño de Centro América (Hrbek et al., 2007).

Actualmente, la taxonomía de este grupo se basa primordialmente en las características del gonopodio, reconociendo a las siguientes tribus: Poeciliini (gonopodio relativamente corto); Cnesterodontini (gonopodio desplazado hacia adelante y elongado); Gambusiini son los menos especializados (gonopodio de pequeña a mediana longitud); Girardinini (gonopodio largo, modificado y complejo). Heterandriini (modificaciones comparables con Girardinini, aunque con un gonopodio más simple). En diversos trabajos se ha reconocido que las tribus deben ser reevaluadas taxonómicamente, porque varias de ellas no son monofiléticas (Lucinda y Reis, 2005).

El género *Poeciliopsis* se distribuye mayormente en los drenajes del Pacífico, desde el sur de Arizona hasta el oeste de Colombia. Las especies de este género pueden distribuirse en un amplio rango de ambientes, desde cuencas tropicales y altitudes bajas, hasta lagos de grandes altitudes, así como manantiales en el desierto y ríos. Han sido reconocidas 20 especies para este género (Mateos et al., 2002).

Un trabajo filogenético realizado con las especies del género *Poeciliopsis* estableció que al menos dos eventos vicariantes en distintos periodos de tiempo en la FVTM fueron los responsables de la divergencia entre el linaje del norte de la FVTM y el linaje del sur de la FVTM. El primer evento ocurrió hace ocho y 16 millones de años, separando los miembros estrictos del norte del clado del sur. Mientras que el segundo evento vicariante tuvo lugar entre 2.8 y 6.4 millones de años (Mateos et al., 2002).

Dentro de este género, la especie *Poeciliopsis infans* es el único pez de origen Neotropical que ha colonizado las partes altas y templadas de la FVTM, incluyendo las Cuencas Lerma-Santiago, Ameca, Armería, Coahuayana, Balsas y Pánuco, así como los lagos endorreicos del Centro de México. Debido a estas características y a que se encuentra codistribuida con especies de origen Neártico, como los goodeinos y cipr ínidos, inferir la historia evolutiva de esta especie era necesario para poder explicar cómo la actividad volcánica y tectónica han moldeado la estructura filogeográfica de esta especie. Por otra

parte, al ser una especie que ha evolucionado en áreas marginales y templadas, en comparación con el resto de las especies de *Poeciliopsis*, *P. infans* podría mostrar patrones filogeográficos contrastantes a los mostrados para otras especies co-distribuidas en la región (Beltrán-López et al., 2018).

I. VI Estudios filogeográficos en el centro de México

Se ha reconocido que la actividad volcánica y tectónica desde la época del Mioceno ha tenido una importante influencia en la diversificación de muchos taxa del nuevo mundo (Castoe et al., 2009). En conjunto, los eventos geológicos y los eventos climáticos durante el cuaternario, son los principales factores que han moldeado la distribución geográfica de la variación genética en las especies, poblaciones y comunidades en diferentes taxa (Bermingham y Martin, 1998; Hewitt, 2000; 2004).

Bajo este contexto, la localización geográfica, la compleja topografía, el dinamismo geológico desde el Mioceno y la historia climática del Centro de México, han moldeado la compleja historia biogeográfica de la zona, que ha permitido la coexistencia de taxa de origen Neotropical y Neártico, así como de grupos endémicos a esta parte del país. Todos estos componentes físicos y biológicos, han influenciado la distribución de los linajes genéticos en espacio y tiempo, tanto de organismos terrestres como de organismos dulceacuícolas (Arellano et al., 2005; Beltrán-López et al., 2017; 2018; Betancourt-Resendes et al., 2018; Bryson et al., 2011a,b,c,d; Domínguez-Domínguez et al., 2008; 2010; Huidobro et al., 2006; Mejía-Madrid et al., 2007; Pedraza-Lara et al., 2012; Pérez-Rodríguez et al., 2009; 2015).

Para taxa dulceacuícolas distribuidos en el Centro de México como peces, crustáceos, plantas y parásitos, se ha demostrado que la estructura filogeográfica parece haber sido influenciada por los intensos eventos vicariantes asociados a la historia geológica y climática de la región (Beltrán-López et al., 2017; 2018; Domínguez-Domínguez et al., 2008a; 2010; Huidobro et al., 2006; Mejía-Madrid et al., 2007; Pedraza-Lara et al., 2012; Pérez-Rodríguez et al., 2009: 2015), principalmente por la configuración antigua de los ríos y no la configuración actual, como es el caso del goodeido *Zoogoneticus quitzeoensis* (Domínguez-Domínguez et al., 2008a), para

especies de cyprinidos del género *Algansea* (Pérez-Rodríguez et al., 2009), para poeciliidos (Beltrán-López et al., 2018; Mateos et al., 2002) y para atherinopsidos (Betancourt-Resendes et al., 2018).

Patrones filogeográficos contrarios, aunque menos comunes, se han reportado en el Centro de México, estudios recientes han demostrado que cuando existe conexión dentro y entre cuerpos de agua, la historia de vida y las características ecológicas de las especies pueden influenciar la capacidad de dispersión y por lo tanto, la conectividad genética de las poblaciones (Beltrán-López et al., *en preparación*; Betancourt-Resendes et al., 2018; Goto y Andoh, 1990). En este sentido, la diferenciación genética baja o nula en especies de peces dulceacuícolas en el Centro de México, ha sido explicada por la alta capacidad de dispersión de las especies, por la incertidumbre taxonómica de algunas especies o incluso debido a dispersión mediada por humanos, (Beltrán-López et al., 2018; Betancourt-Resendes et al., 2018; Corona-Santiago et al., 2015; Ornelas-García et al., 2012).

La presente tesis, incluye un trabajo filogeográfico con marcadores mitocondriales y nucleares de cuatro grupos de especies de peces dulceacuícolas vivíparos, *Ilyodon* spp., *Goodea* spp., *Xenotoca variata* (Goodeinae) y *Poeciliopsis infans* (Poeciliinae), todos con distribución en el Centro de México. Este trabajo aporta nuevos datos acerca de los patrones filogeográficos de peces víviparos del Centro de México, además de relacionar la historia geológica de la región y las características biológicas y ecológicas propias de cada especie con los patrones filogeográficos encontrados, sugiriendo nuevas hipótesis acerca de la historia evolutiva de especies de dos subfamilias distribuidas en el Centro de México.

II. OBJETIVOS.- OBJETIVO GENERAL

Describir la historia evolutiva de las especies de goodeinos *Ilyodon* spp., *Goodea* spp. y *Xenotoca variata,* así como del poecílido *Poeciliopsis infans* a lo largo de las cuencas del Centro de Mexico.

III.OBJETIVOS PARTICULARES

CAPÍTULO I

1. Entender las relaciones filogenéticas y filogeografícas de las especies que componen el género *Ilyodon* en toda su área de distribución.
2. Describir la historia biogegráfica en espacio y tiempo de las especies bajo estudio.
3. Entender la taxonomía de las especies que componen el género *Ilyodon*.
4. Evaluar a través de árboles de especies y pruebas de delimitación filogenética de especies, el número de especies que componen el género *Ilyodon*.
5. Relacionar las relaciones filogenéticas y la taxonomía del género con eventos antiguos de la actividad volcánica y tectónica de la región.

CAPÍTULO II

1. Describir las relaciones filogenéticas y filogeográficas de las especies que han sido descritas para el género *Goodea* en todo su rango de distribución.
2. Describir la estructura genética de *Goodea*.
3. Describir la diversidad genética de *Goodea*.
4. Examinar la historia demográfica de *Goodea*.
5. Examinar la relación entre la historia evolutiva de *Goodea atripinnis* (Goodeinae) y la historia geológica y climática del Centro de México.
6. Comparar los patrones encontrados para *G. atripinnis* con los patrones encontrados en estudios previos de goodeinos y otras especies de peces co-distribuidas.

CAPÍTULO III

1. Describir los patrones filogeográficos de *Xenotoca variata* (Goodeinae).

2. Inferir la estructura genética de las poblaciones en toda su área de distribución.
3. Describir la diversidad genética de las poblaciones de esta especie.
4. Elucidar el tiempo de divergencia de los linajes genéticos recuperados.
5. Describir la historia biogeográfica de los linajes genéticos recuperados.
6. Evaluar el papel de los eventos geológicos y climáticos en la diversificación de los linajes.
7. Evaluar como los aspectos ecológicos y biológicos afectan la historia evolutiva de la especie.

CAPÍTULO IV

1. Inferir la variación filogeográfica de todas las poblaciones de la especie de poecilido *Poeciliopsis infans* en toda su área de distribución.
2. Inferir la historia biogeográfica en espacio y tiempo de los linajes recuperados para la especie.
3. Describir la historia demográfica de las poblaciones de *P. infans*.
4. Examinar la influencia que han tenido los ciclos glaciales e interglaciares en la historia demográfica de los linajes genéticos así como en la distribución potencial en el pasado (último interglaciar y último máximo glacial) para la especie.
5. Examinar la influencia de la historia geológica de la región en espacio y tiempo en la historia evolutiva de la especie Neotropical (*P. infans*) evolucionando en un área predominantemente Neártica.

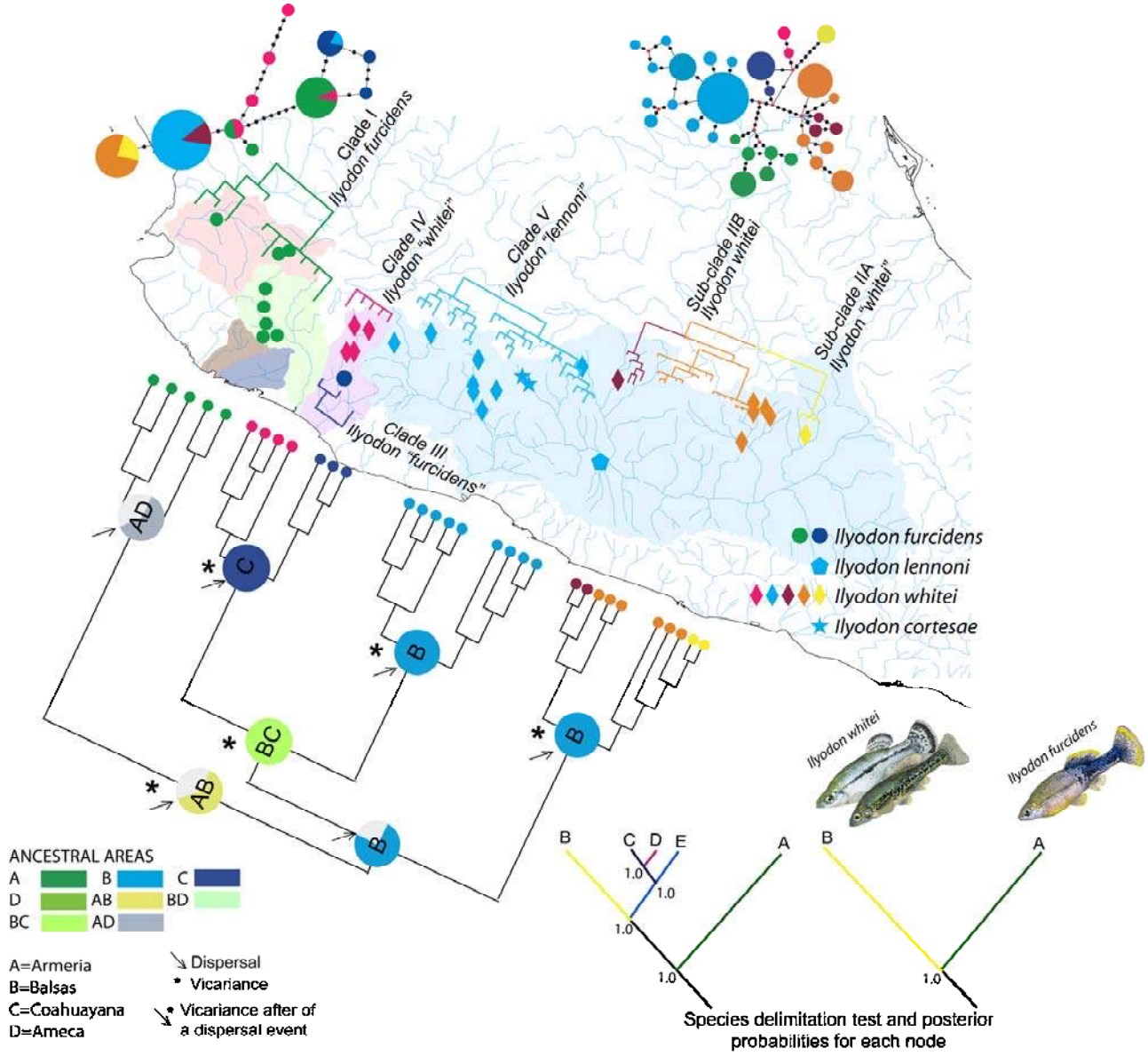
IV. RESULTADOS

CAPÍTULO I

Phylogeny and taxonomy of the genus *Ilyodon* Eigenmann, 1907 (Teleostei: Goodeidae), base on mitochondrial and nuclear DNA sequences

Beltrán-López, R. G., Domínguez-Domínguez, O., Guerrero, J. A., Corona-Santiago, D. K., Mejía-Mojica, H., & Doadrio, I. (2017). Phylogeny and taxonomy of the genus *Ilyodon* Eigenmann, 1907 (Teleostei: Goodeidae), based on mitochondrial and nuclear DNA sequences. *Journal of Zoological Systematics and Evolutionary Research*, 55(4), 340-355. DOI: 10.1111/jzs.12175

Graphical abstract



Abstract

Taxonomy of the live-bearing fish of the genus *Ilyodon* Eigenmann, 1907 (Goodeidae), in Mexico, is controversial, with morphology and mitochondrial genetic analyses in disagreement about the number of valid species. The present study accumulated a comprehensive DNA sequences dataset of 98 individuals of all *Ilyodon* species and mitochondrial and nuclear loci to reconstruct the evolutionary history of the genus. Phylogenetic inference produced five clades, one with two sub-clades, and one clade including three recognized species. Genetic distances in mitochondrial genes (*cytb*: 0.5-2.1%; *coxI*: 0.5-1.1% and *d-loop*: 2.3-10.2%) were relatively high among main clades; while, as expected, nuclear genes showed low variation (0.0-0.2%), with geographic concordance and few shared haplotypes among river basins. High genetic structure was observed among clades and within basins. Our genetic analyses, applying the priority principle, suggest the recognition only of *Ilyodon whitei* and *Ilyodon furcidens,* with *I. cortesae* relegated to an invalid species, the populations of which belong to *I. whitei*.

KEYWORDS

Diversification, endemic fish, mitochondrial DNA, nuclear DNA, taxonomy

INTRODUCTION

The geological history of Central Mexico is characterized by high tectonic and volcanic activity since the Miocene, at least 16 Mya that continues to the present, generating an ongoing process of hydrological reconfiguration (Ferrari Conticelli, Vaggelli, Potrone, & Manetti, 2000). This dynamic geomorphology has been postulated as the primary cause of the complex evolutionary history of the freshwater fish fauna of Central Mexico, exceeding the effects of biological characteristics and the evolution of climate conditions (Barbour 1973; Domínguez-Domínguez, Doadrio, Martínez-Meyer, Zambrano, & Pérez-Ponce de León, 2006; Domínguez-Domínguez et al., 2010; Pérez-Rodríguez, Domínguez-Domínguez, Doadrio, Cuevas-García, & Pérez-Ponce de León, 2015; Smith, 1980). Of the nearly 100 described species of freshwater fish in Central Mexico, ~70% are endemic (Miller Minckley, & Norris, 2005) as a result of paleogeological isolation processes, especially volcanism and tectonic events.

Studies of endemic and native freshwater fishes in Central Mexico have chiefly focussed on phylogeny based on DNA sequences or biogeographic aspects of complete groups of fishes in Mexico (Corona-Santiago, Doadrio, & Domínguez-Domínguez, 2015; Doadrio & Domínguez, 2004; Domínguez-Domínguez, Pérez-Rodríguez, Escalera-Vázquez, & Doadrio, 2009; Domínguez-Domínguez et al., 2010; Pérez-Rodríguez, Domínguez-Domínguez, Pérez-Ponce de León, & Doadrio, 2009; Pérez-Rodríguez et al., 2015; Schönhuth & Doadrio, 2003; Schönhuth, Doadrio, Domínguez-Domínguez, Hillis, & Mayden, 2008), while within-species phylogeographic studies are scarce (Domínguez-Domínguez, Alda, Pérez-Ponce de León, García-Garitagoitia, & Doadrio, 2008; Mateos, Sanjur, & Vrijenhoek, 2002), especially of species distributed in Central Mexico Pacific drainages (CMPD) (Domínguez-Domínguez et al., 2006; Mateos 2005; Piller, Kenway-Lynch, Camak, & Domínguez-Domínguez, 2015). Phylogeographic and population studies are essential tools in understanding evolutionary patterns and provide useful information on genetic isolation of populations on a geographic and temporal scale. Phylogeographic data are especially relevant when the populations studied are under threat, as is the case for the Central Mexico endemic

subfamily Goodeinae (Domínguez-Domínguez & Pérez-Ponce de León, 2007). Phylogeographic studies can identify divergent populations and evolutionarily isolated lineages undetected by traditional taxonomy (Domínguez-Domínguez et al., 2008; Mateos, 2005; Piller et al., 2015).

The goodeids in Central Mexico include the endemic subfamily Goodeinae represented by approximately 19 genera, including *Ilyodon*, and 40 species of viviparous fishes with internal fertilization and matrotrophy (Doadrio & Domínguez, 2004; Dominguez-Domínguez & Pérez-Ponce de León, 2007). Geographic distribution of *Ilyodon* is limited in the CMPD to the main basin of the Balsas River and the adjacent Coahuayana, Armería, Ameca, Purificacion, and Marabasco River basins (Figure1).

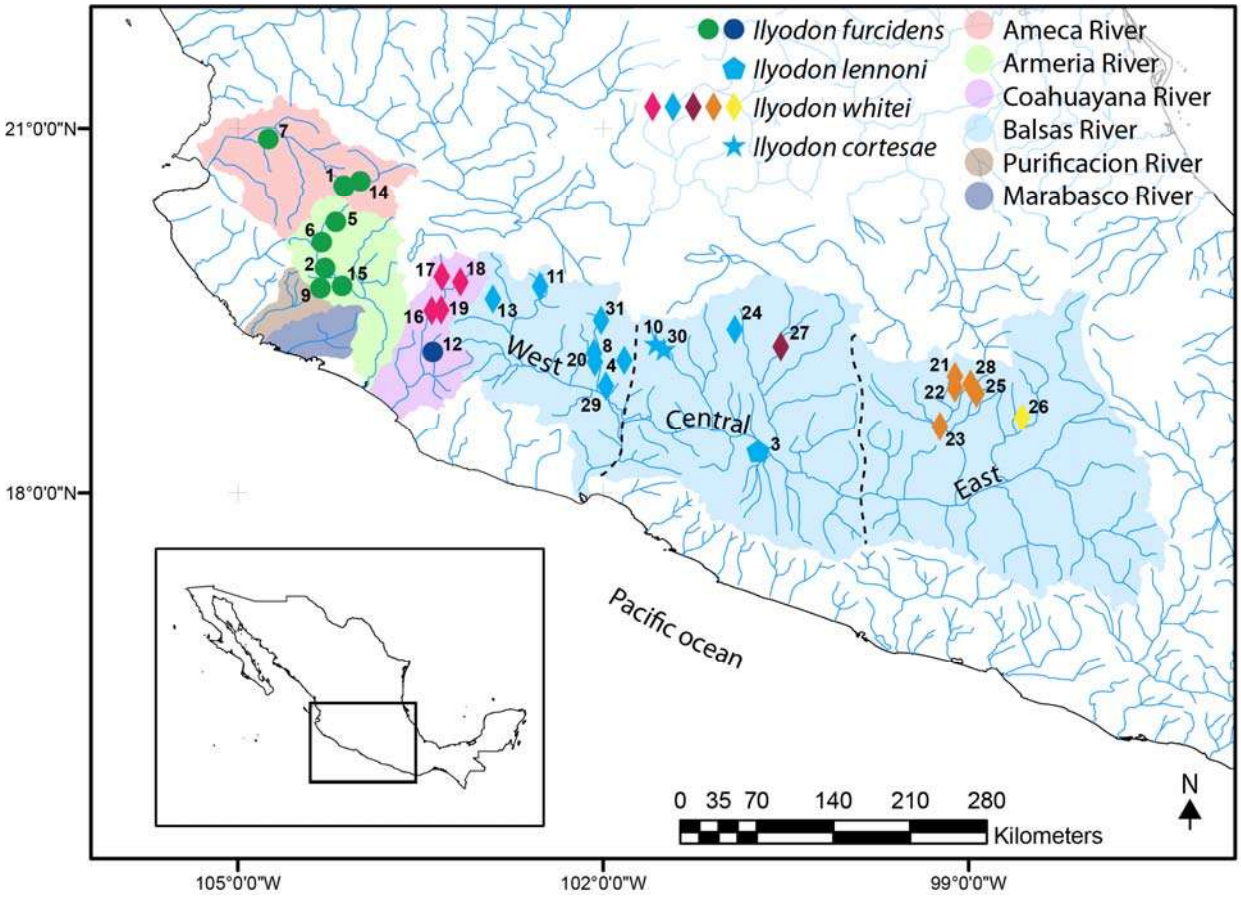


FIGURE 1 Sampling locations (symbols) and the hydrological basins (filled areas) where *Ilyodon* is distributed, numbers corresponded with Table 2. The lines of the Balsas basin show the west, central, and east divisions considered in the present study. The colors of symbols correspond to the colors of the clades.

Species of *Ilyodon* have long been taxonomically controversial. Five species have been described: *Ilyodon whitei* (Meek, 1904), *Ilyodon furcidens* (Jordan & Gilbert, 1882), *Ilyodon xantusi* (Hubbs & Turner, 1939), *Ilyodon lennoni* Meyer & Föerster, 1983, and *Ilyodon cortesae* Paulo-Maya & Trujillo-Jimenéz, 2000. However, opinions have differed with regard to the number of valid species. The catalogue of fishes (Eschmeyer, Fricke & Van der Laan, 2016) includes four valid species (*I. whitei*, *I. furcidens*, *I. lennoni*, and *I. cortesae*), whereas taxonomy based on molecular studies has identified only *I. whitei* and *I. furcidens* (Doadrio & Domínguez, 2004). The results of a comprehensive study of Goodeidae, including all species of *Ilyodon,* using a single mitochondrial DNA gene, did not find monophyletic groups (Domínguez-Domínguez et al., 2010).

At the population level, morphological and genetic data also generate wide discussion of *Ilyodon* speciation *vs.* phenotypic plasticity in previously separated species or subspecies (Kingston, 1979; Turner & Grosse, 1980). Two trophic morphs found in sympatry in the Coahuayana and Armería River basins, described as *I. furcidens* and *I. xantusi* have been considered the same species with an incipient signature of reproductive isolation, resulting from trophic differentiation (Grudzien & Turner, 1984a,b; Turner & Grosse, 1980).

The present study comprises a comprehensive report of *Ilyodon,* including all identified species throughout their distribution range and DNA sequences analyses of three mitochondrial and two nuclear markers. The use of combined mitochondrial and nuclear genes allows a better understanding of the evolution and taxonomy of *Ilyodon.* The aims of the study were to infer the evolutionary history of *Ilyodon* and elucidate the relationships among described species.

MATERIALS AND METHODS

Fish sampling

Ninety-eight specimens of *I. cortesae, I. furcidens, I. lennoni*, and *I. whitei* from throughout their distribution range were collected from the east, central, and west sub-basins of the Balsas River basin, as well as the Coahuayana, Armería, and Ameca River basins (Figure 1). *Ilyodon xantusi* has been described from a tributary of the Armería River, near Colima city. In this work,

following Turner and Grosse (1980), we considered *I. xantusi* a junior synonym of *I. furcidens*.

Fish were captured by electrofishing and trawl nets and anesthetized with tricaine mesylate (MS-222). A fin fragment of each fish was preserved in 95% ethanol for DNA extraction. A maximum of five specimens from each locality were preserved in 5% formalin and deposited in the fish collection of several institutions, approved by the Ministry of Environment and Natural Resources for Mexico (SEMARNAT), with the permission number: SGPA/DGVS/08473/15.

The remaining fish were released at the capture site. The fin clips were deposited in the fish collection at the Universidad Michoacana de San Nicolás de Hidalgo, México (SEMARNAT registration number MICH-PEC-227-07-09), the fish collection of the Universidad Autónoma del Estado de Morelos, México (SEMARNAT registration number MOR-CC-243-201), and the collection of the Museo Nacional de Ciencias Naturales, Spain (Table 1). Based on published reports and available samples, we used *Allodontichthys* as outgroup (Doadrio & Domínguez 2004; Domínguez-Domínguez et al., 2010). Information on sampling is provided in Table 2.

TABLE 1 Geographical information, collection where the tissue are deposited and the voucher number of the sample

Locality	GPS Coordinates	Fish Collection	Tissue voucher number
Río Ameca, puente la muerta	20° 31´ 44´´ N, 104° 7´ 47.3´´ W	MNCN	31943, 31944, 31945, 31946
Río Armeria	19° 51´ 1.3´´ N, 104° 17´ 0.0´´ W	MNCN	31971
Arroyo Chacambero	18° 20´ 44.53´´ N, 100° 43´ 43.97´´ W	MNCN	32155, 32156, 32157, 32752
Río Las Trojes	19° 05´ 38.8´´ N, 101° 49´ 33.5´´ W	MNCN	32158, 32159
Presa Copales	20° 13´ 49.2´´ N, 104° 11´ 42.9´´ W	MNCN	33008
Presa Tacotán	20° 3´ 42.78´´ N, 104° 18´ 43.67´´ W	MNCN	64259
Río Las Rosas	20° 54´ 51.27´´ N, 104° 45´ 1.75´´ W	MNCN	4266, 64367, 64368
Río Cajones	19° 9´ 13.59´´ N, 102° 4´ 23.7´´ W	UMSNH	5179, 5180
Río Ahuacapan	19° 39´ 54.21´´ N, 104° 19´ 19.79´´ W	UMSNH	8842, 8843
Manantial Cutzaróndiro	19° 11´ 0.6´´ N, 101° 30´ 8.3´´ W	UMSNH	9162, 9164, 9166
		MNCN	64229, 64230
Manantial Tocumbo	19° 42´ 7´´ N, 102° 30´ 60´´ W	UMSNH	9260, 9984, 9986
Arroyo El Tule, Tuxpan	19° 19´ 32.3´´ N, 103° 22´ 18.8´´ W	UMSNH	9267, 9268, 9269
Los Horcones	19° 35´ 48.1´´ N, 102° 54´ 15.2´´ W	UMSNH	9393
Potrero Grande	20° 31´ 15´´ N, 104° 7´ 36´´ W	UMSNH	9938, 9940, 9942, 9946
Achacales	19° 42´ 14.1´´ N, 104° 8´ 37.9´´ W	UMSNH	11989, 11990, 11991
Atenquique, Tuxpan	19° 31´ 46.35´´ N, 103° 25´ 56.39´´ W	UMSNH	12018, 12019
San Jerónimo, Tuxpan	19° 41´ 42.3´´ N, 103° 21´ 8.2´´ W	UMSNH	12033
Los Pitayos, Tuxpan	19° 45´ 35.3´´ N, 103° 11´ 8.4´´ W	UMSNH	12056
Arroyo La Purisima, Tuxpan	19° 31´ 19.8´´ N, 103° 20´ 32.9´´ W	UMSNH	13042
Arroyo coróndiro, Nueva Italia.	19° 4´ 46.7´´ N, 102° 4´ 2.5´´ W	UMSNH	36420, 36421, 36422, 36423, 36424
Barranca de Cuernavaca, Morelos	18° 52´ 0´´ N, 99° 6´ 38.6´´ W	UMSNH	36434, 36435, 36436, 36437, 36438
Barranca San Andrés de la Cal	18° 57´ 41.1´´ N, 99° 7´ 46´´ W	UMSNH	36444, 36445, 36446, 36447, 36448, 36449

Río Apatlaco, Jojutla	18° 33´ 0´´ N, 99° 14´ 0´´ W	UAEM	51
Río Chinapa, Tzitzio	19° 20´ 50.8´´ N, 100° 55´ 6.7´´ W	UMSNH	36415,36416, 36417, 36418, 36419
Río Cuautla, el ojito	18° 49´ 18.3´´ N, 98° 56´ 0.1´´ W	UMSNH	36429, 36430, 36431, 36432, 36433
Río Rijo, Izúcar de Matamoros	18° 37´ 23.8´´ N, 98° 33´ 39´´ W	UAEM	52, 53, 54, 55
Río Zitacuaro, Tuzantla	19° 12´ 18.3´´ N, 100° 32´ 24.2´´ W	UMSNH	36406, 36407, 36408, 36409, 36450
Río Yautepec, Oaxtepec	18° 53´ 55´´ N, 98° 58´ 59´´ W	UMSNH	36439, 36440, 36441, 36442, 36443
Río Zicuirán, Zicuirán	18° 53´ 1.4´´ N, 101° 58´ 35.7´´ W	UMSNH	36410, 36411, 36412, 36413, 36414
Río Tacámbaro, Puruarán	19° 11´ 0.6´´ N, 101° 30´ 8.3´´ W	UMSNH	36425, 36426, 36427, 36428, 36451
Río Cupatitizio, Parque Uruapan	19° 23´ 19´´ N, 102° 0´ 51´´ W	MNCN	33649
		UMSNH	9584, 9585

UMSNH, Universidad Michoacana de San Nicolás de Hidalgo; UAEM, Universidad Autónoma del Estado de Morelos, MNCN, Museo Nacional de Ciencias Naturales

TABLE 2 Sampling locations

Site	Locality	Sub-basin	Basin	Species
1	Río Ameca, puente la muerta	Ameca	Ameca	*Ilyodon furcidens*
2	Río Armeria	Ayuquila	Armería	*Ilyodon furcidens*
3	Arroyo Chacambero	Medio Balsas	central Balsas	*Ilyodon lennoni*
4	Río Las Trojes	Cupatitzio	west Balsas	*Ilyodon whitei*
5	Presa Copales	Ayuquila	Armería	*Ilyodon furcidens*
6	Presa Tacotán	Ayuquila	Armería	*Ilyodon furcidens*
7	Río Las Rosas	Mascota	Ameca	*Ilyodon furcidens*
8	Río Cajones	Cupatitzio	west Balsas	*Ilyodon whitei*
9	Río Ahuacapan	Ahuacapan	Ameca	*Ilyodon furcidens*
10	Manantial Cutzaróndiro	Tacámbaro	central Balsas	*Ilyodon cortesae*
11	Manantial Tocumbo	Tepalcatepec	west Balsas	*Ilyodon whitei*
12	Arroyo El Tule, Río Tuxpan	Tamazula	Coahuayana	*Ilyodon furcidens*
13	Los Horcones	Tepalcatepec	west Balsas	*Ilyodon whitei*
14	Potrero Grande	Ameca	Ameca	*Ilyodon furcidens*
15	Achacales	Ayuquila	Armería	*Ilyodon furcidens*
16	Atenquique, Río Tuxpan	Tamazula	Coahuayana	*Ilyodon whitei*
17	San Jerónimo, Río Tuxpan	Tamazula	Coahuayana	*Ilyodon whitei*
18	Los Pitayos, Río Tuxpan	Tamazula	Coahuayana	*Ilyodon whitei*
19	Arroyo La Purisima, Río Tuxpan	Tamazula	Coahuayana	*Ilyodon whitei*
20	Arroyo coróndiro, Nueva Italia.	Cupatitzio	west Balsas	*Ilyodon whitei*
21	Barranca de Cuernavaca, Morelos	Amacuzac	east Balsas	*Ilyodon whitei*
22	Barranca San Andrés de la Cal	Amacuzac	east Balsas	*Ilyodon whitei*
23	Río Apatlaco, Jojutla	Amacuzac	east Balsas	*Ilyodon whitei*
24	Río Chinapa, Tzitzio	Cutzamala	central Balsas	*Ilyodon whitei*
25	Río Cuautla, el ojito	Amacuzac	east Balsas	*Ilyodon whitei*
26	Río Rijo, Izúcar de Matamoros	Atoyac	east Balsas	*Ilyodon whitei*
27	Río Zitacuaro, Tuzantla	Cutzamala	central Balsas	*Ilyodon whitei*
28	Río Yautepec, Oaxtepec	Amacuzac	east Balsas	*Ilyodon whitei*
29	Río Zicuirán, Zicuirán	Cupatitzio	west Balsas	*Ilyodon whitei*
30	Río Tacámbaro, Puruarán	Tacámbaro	central Balsas	*Ilyodon cortesae*
31	Río Cupatitizio, Parque Uruapan	Cupatitzio	west Balsas	*Ilyodon whitei*

DNA Extraction, amplification, *and sequencing*

Total genomic DNA was isolated with the Qiagen Dneasy Tissue and Blood Kit (Qiagen, Valencia, CA, USA) following the manufacturer's protocol. Fragments of three mitochondrial genes and two nuclear genes were amplified: cytochrome b (*cytb*: 533 bp), cytochrome oxidase subunit I (*coxI*: 626 bp), and control region (*d-loop*: 441 bp), for a total of 1600 bp from 98 individuals, and a fragment of the nuclear β-actin gene (*ACTB*: 979) and the exon 3 of the

recombinant activating gene 1 (*RAG1*: 1453), for a total of 2432 bp from a subset of 51 individuals, representing the variation found in mtDNA haplotypes. Polymerase chain reactions (PCRs) were conducted in a reaction volume of 12.5 µl containing 4.25 µl ultrapure water, 0.5 µl of each 0.2 µM primer, 6.25 µl Dream Taq Green PCR Master Mix 2x (Thermo Scientific, Waltham, MA, USA), and 1 µl (ca 10-100 ng) of DNA template. The protocols for amplification are presented in Table S1. The PCR products were purified using ExoSAP-IT (USB Corp. Cleveland, OH, USA) and submitted to Macrogen Inc. (Netherlands) for sequencing. Nucleotide sequences were edited and aligned in Mega v. 6.06 (Tamura et al. 2013). The sequences of *ACTB* showed heterozygous positions defined by indels, a manual reconstruction of the two allele phases was performed following the procedure described by Sousa-Santos, Robalo, Collares-Pereira & Almada (2005). The *d-loop* and *ACTB* genes showed ambiguously aligned positions that are showed in Table S2. For the *RAG1* gene, the phase of heterozygous genotypes was resolved using DNAsp v. 5.10 (Librado & Rozas, 2009) and conducted with the algorithm provided by PHASE v. 2.0 (Stephens & Donelly, 2003). Recombination of nuclear genes was assessed with the phi test in Splitstree v. 4.13 (Huson & Bryant, 2006) and did not find significant evidence for recombination in either gene (p= 1 for both). Codification of amino acids was used to verify the alignment and the absence of stop codons. The obtained sequences were deposited in GenBank under accession numbers for *cytb*: KY204452-KY204540, for *coxI*: KY118827-KY118914, for *d-loop*: KY204628-KY204716, for *ACTB*: KY204717-KY204778, and for *RAG1*: KY204541-KY204627 (Table S3). All raw data: alignments of each one of the genes, are show in Dataset S1, S2, S3, S4 and S5.

Phylogeny based on DNA sequences and haplotype networks

The incongruence length difference test (partition homogeneity test; Farris, Källersjö, Kluge, & Bult, 1995) was conducted in Phylogenetic Analysis Using Parsimony* and other methods (PAUP*) v. 4.0b10 (Swofford, 2003) to evaluate the significance of conflict among datasets, using 1000 resampling of characters. Phylogenetic analyses were conducted for each gene, for the concatenated dataset for mitochondrial genes, and for the five genes combined. Model selection based on the Akaike information criterion and optimal partition-

setting analyses, conducted using PartitionFinder v. 1.1.0 (Lanfear, Calcott, Ho, & Guindon, 2012), suggested that optimal partition-setting was obtained by assigning a substitution model for each gene (Table S4). Genetic trees were constructed using maximum likelihood and Bayesian inference. Maximum likelihood analyses were carried out using RAxMLGUI v.1.3.1 (Silvestro & Michalak, 2012; Stamatakis, 2014), with the substitution model GTR+gamma and 10,000 bootstrap replicates.

The relative stability of clades was evaluated by 1,000 nonparametric bootstrap replicates (Alfaro, Zooler, & Lutzoni, 2003). Bayesian analyses were implemented using MrBayes v. 3.2.1 (Ronquist et al., 2012). The analysis was run for 10 million of generations, with two independent runs implementing four Markov chain Monte Carlo (MCMC) processes, sampling every 100 generations. We evaluated the chain convergence with the log-likelihood (-lnL) values of the runs on Tracer v. 1.5 (Rambaut & Drummond, 2007), discarding 10% of generations as burn-in to construct the consensus tree (σ= 0.0002).

To determine the geographic correspondence with the genetic structure, a haplotype network for each gene was constructed using the median-joining algorithm as implemented in Network v. 4.6.1.3 (Bandelt, Forster, & Röhl, 1999).

Genetic distances and structure

To analyze the genetic structure of populations of *Ilyodon* spp., we conducted analyses of molecular variance (AMOVA) and calculated the components of the fixation index Φ_{CT} (variation between groups), Φ_{ST} (variation within populations), and Φ_{SC} (variation among populations within groups) using Arlequin v. 3.5.1.3 (Excoffier & Lischer, 2010). The analyses were implemented for the five genes separately, as well as grouped, according to various criteria to estimate partitioning of the genetic variance at different hierarchical levels (Excoffier, Smouse, & Quattro, 1992). The first analysis considered each hydrological basin as a group. In the second analysis, each group comprised species that have been described and recognized as valid. Finally, in a third analysis all recovered groups within the main clades in phylogenetic inference were considered as groups. Analyses were performed using 10,000 permutations to significance values estimated in Arlequin v. 3.5.1.3.

The uncorrected genetic distances were calculated between the recovered groups in phylogenetic trees for each mitochondrial gene (*cytb*, *coxI* and *d-loop*), and between all individuals for *ACTB* and *RAG1* in Mega v.6.06 (Tamura et al., 2013), a bootstrapping process was performed with 1,000 repetitions.

Species delimitation test

Species tree analysis was conducted to obtain a guide tree and speciation model, using a multispecies coalescent model (Heled & Drummond, 2010) in BEAST v. 1.8.1 (Drummond, Suchard, Xie, & Rambaut, 2012), for implementation in the Bayesian species delimitation test using Bayesian phylogenetics and phylogeography (BPP v. 3.1; Yang & Rannala, 2010; Yang, 2015). For estimating the species tree model, the analysis was performed using the assumption that each clade recovered in the phylogenetic analyses represented a different species.

For the species tree ancestral reconstruction (StarBEAST) implemented in BEAST, the model parameters were unlinked across *cytb*, *coxI*, *d-loop*, *ACTB*, and *RAG1* genes. The dataset consisted of one sequences for each gene (*cytb*, *coxI*, *d-loop*, *ACTB* and *RAG1*), of each one of the clades and sub-clades. Substitution models were set according to the selected model for each gene by PartitionFinder v. 1.1.0 (Lanfear et al., 2012). We applied a lognormal relaxed clock (Uncorrelated) model on branch length and calibrated the *cytb* partition using the mutation rate of *cytb* in teleosts of 0.76–2.2%/million years (Machordom & Doadrio, 2001; Near & Benard, 2004; Zardoya & Doadrio, 1999). We estimated the evolutionary rate of the *coxI*, *d-loop*, *RAG1* and *ACTB* genes relative to the *cytb* gene. We selected the tree prior-species Tree: Yule process model. Markov chain Monte Carlo analysis was run for 70 million generations, sampled every 1000 generations. We evaluated the chain convergence with the -InL values in Tracer v. 1.5 (Rambaut & Drummond, 2007), and summarized the results using TreeAnnotator v. 1.8.1 (Drummond et al., 2012).

For the BPP analyses of the five concatenated genes, we used the reversible-jump Markov chain Monte Carlo (rjMCMC) (Yang & Rannala, 2010) algorithm to determine whether to collapse or retain nodes throughout the phylogeny. Using the entire dataset coded by each gene, we tested with two

algorithms: Analysis A10, in which the rjMCMC algorithm was used to move between species-delimitation models that were compatible with a fixed guide tree (Rannala & Yang, 2013; Yang & Rannala, 2010), and Analysis A11 that explored species delimitation models and species phylogenies with the nearest neighbour interchange or sub-tree pruning and re-grafting used to change the species-tree topology and test all species-tree models from a fixed tree (Yang & Rannala, 2014).

To determine whether lineages could be considered distinct species under a general lineage species concept, the program assessed the probability of the node separating the species (de Queiroz, 2007). We used algorithm 0 with values of 5, 10, 15, 20 for the fine-tuning parameter to ensure that the rjMCMC mixed effectively in species-delimitation models. We conducted analyses with priors θ and τ_0 (Leaché & Fujita, 2010) to discern how the effective ancestral population size and time of divergence influenced results. We initially set the gamma prior at θ and τ to the values α = 1 and 2 and β = 10, 100, and 2000 and ran four analyses of each with different starting seeds for two independent chains of 500,000 generations with a burn-in of 50,000 and thinning every five generations. Finally, to test the robustness of the results, the analysis was repeated, randomizing individuals to either group to minimize the over-splitting effect and changing the speciation model according to the genetic results obtained (two to six species).

Ancestral area reconstruction

The ancestral area reconstruction for the species of genus *Ilyodon* was estimated using the dispersal-extinction-cladogenesis (DEC) model of LAGRANGE (Ree, Moore, Webb & Donoghue, 2005; Ree & Smith, 2008), implemented in RASP v. 3.2 software (Yu, Harris, Blair & He, 2015). The ultrametric and dichotomous tree obtained for the five concatenated genes in BEAST software was used as the tree topology on which mapping ancestral areas. The number of maximum areas was kept as 2. For this analysis, we divided the distribution area of *Ilyodon* in four according with hydrological regions of distribution: Ameca River, Armeria River, Balsas River and Coahuayana River.

RESULTS

Phylogenetic relationships

The incongruence length difference test did not show significant differences, indicating that all genes presented the same phylogenetic signal. The phylogenetic analyses for the mitochondrial (*cytb*, *coxI*, d-*loop*: 1,600 bp; Figure S1) and the concatenated gene dataset (*cytb*, *coxI*, d-*loop*, *ACTB*, *RAG1*: 4032 bp), based on maximum likelihood and Bayesian methods, recovered the same topology. Five well-differentiated clades were geographically segregated but did not correspond to actual basin configuration, corresponding to Ameca-Armería (Clade A), central and east Balsas (Clade B), Coahuayana lower (Clade C), Coahuayana upper (Clade D), and west Balsas (Clade E) watersheds. The phylogenetic relationships among the five clades were not resolved, appearing as a large basal polytomy (Figure 2). Clade A clustered individuals identified as *I. furcidens* from Ameca and Armería basins. Clade B clustered individuals of the central and east Balsas basin identified as *I. whitei*. For clade B two well supported sub-clades were identified: B1 included the Atoyac sub-basin (east Balsas) specimens, and B2 included clustering samples from the Zitacuaro River of Cutzamala sub-basin (central Balsas) and the Amacuzac sub-basin (east Balsas). Clade C consisted of samples from the lower Coahuayana basin identified as *I. furcidens*. Clade D clustered samples of the upper Coahuayana basin identified as *I. whitei*. These two species were identified on the basis of the morphology of their mouth, teeth, head, and coloration patterns. Finally, Clade E grouped individuals of the central and western Balsas, including the sub-basins Cutzamala, Tacámbaro, middle Balsas, Cupatitzio, and Tepalcatepec, comprising specimens identified as *I. whitei*, *I. lennoni*, and *I. cortesae*, the last two collected in the type locality. Phylogeny based on nuclear genes was unresolved and high polytomy was recovered, as expected, for genes with low variation in closely related species (Figure S2).

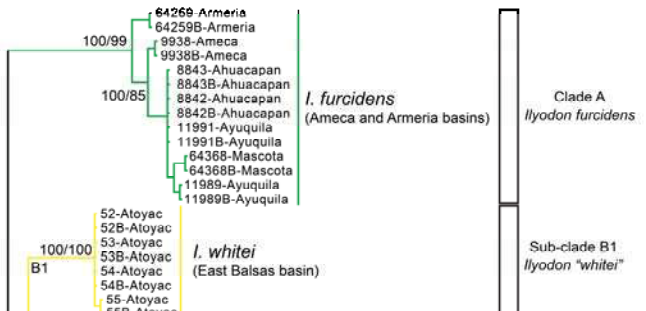


FIGURE 2 The Bayesian inference tree of *Ilyodon* species inferred from concatenated sequences of three mitochondrial genes (*cytb*, *coxI*, and *d-loop*; 1,600 bp) and two nuclear genes (*ACTB* and *RAG1*; 2,432 bp) concatenated. Bayesian posterior probability (>90%) and maximum likelihood bootstrap values (>80%) are indicated. Under the name of each clade, the taxonomic proposal of the present work is found.

The haplotype networks for the mitochondrial genes essentially showed the

general pattern of the phylogenetic analyses, with no shared haplotypes among groups. Five corresponded to the main clades A, B, C, D, and E and two to the sub-clades B1 and B2 (Figure S3). The number of mutation steps between haplogroups differed depending on the marker, with 14–31 for the *d-loop*, 5–12 for *cytb*, and 3–9 for the *coxI*. The six haplogroups found in mitochondrial networks were not recovered in the haplotype networks of the nuclear genes (Figure S4), and shared haplotypes were observed in nuclear genes of the defined mitochondrial groups. For the nuclear *ACTB* gene, structure was found in Ameca, Armería, and lower Coahuayana specimens with shared haplotypes. They were separated from the Balsas specimens, with the exception of the samples from the Tepalcatepec sub-basin of the west Balsas basin, which were closely related to the upper Coahuayana samples. Samples from upper and lower Coahuayana showed no shared haplotypes. A single mutation step separated most of the samples from east-central Balsas from those to the west-central Balsas. The samples from the Zitacuaro River shared haplotypes with those of west Balsas, whereas the mtDNA was consistent with samples from the Amacuzac River, in the east Balsas. For *RAG1*, the same general pattern was observed, but haplotypes from Ameca, Armería, and the lower and upper Coahuayana showed more shared haplotypes than they did for the *ACTB* gene.

Genetic distances and structure

The uncorrected mean genetic distances calculated between the main clades ranged from 3.7–10.2% for the *d-loop*, 1–2.1% for *cytb*, and 0.6–1.1% for *coxI* (Tables S5, S6), and for nuclear genes ranged from 0–0.2% (Table S7). Highest genetic distance for the *d-loop* and *cytb* genes was found between clade A and sub-clade B1, at 10.2% and 2.1% respectively. Based on *coxI*, the maximum genetic distances were observed in clade E and C with respect to sub-clade B2, 1.1% in both cases. Within the Balsas basin (clade B, east-central and clade E, west-central), the mean genetic distances were 4.7% for the *d-loop*, 1.6% for *cytb*, and 1.1% for *coxI*. Sub-clades B1 and B2 showed mean genetic distances of 4.3% for the *d-loop*, 1.6% for the *cytb*, and 0.8% for the *coxI* genes. The genetic distances between clade C and D (lower and upper Coahuayana) were 0.5% for *cytb* and *coxI*, 2.3% for *d-loop*, and 0.1% for the nuclear genes.

In all analyzed genes, significant genetic structure was observed among *a priori* groups, among populations within groups, and within populations.

For the three mitochondrial genes, the highest ($p < 0.0001$) percent of variation among groups was when populations were grouped according to phylogenetic analyses (*cytb*, 72.9%; *coxI*, 73.8%; *d-loop*, 78.6%; Table 3) and not according to hydrological basin (*cytb*, 35.5%; *coxI*, 29.9%; *d*-loop, 54.3%), or recognized species (*cytb*, 16.4%; *coxI*, 24.2%; *d-loop*, 31.1%). For nuclear genes, the highest ($p < 0.0001$) percent of variation was also among groups according the phylogenetic analyses (*ACTB*, 80.09%; *RAG1*, 45.91%) (Table 4), but differences for *RAG1* were lower than found in mitochondrial genes and *ACTB* (Table 3).

TABLE 3 Analyses of molecular variance for groups according to hydrological basin, recognized species, and those recovered in phylogenetic analyses [A = Ameca and Armeria; B1 = east Balsas; B2 = east-central Balsas; C = lower Tuxpan River (Tamazula); D = upper Tuxpan River (Tamazula); E = west-central Balsas] for the mitochondrial genes.

	Cytb			
Testing assumptions	Source of variation	% of variance	Fixation index	*P*- value
Grouped according to hydrological basin [Ameca, Armeria, Coahuayana and Balsas]	Among groups	35.46	Φ_{CT}: 0.35	ns
	Among populations within groups	51.85	Φ_{SC}: 0.80	<0.0001
	Within populations	12.69	Φ_{ST}: 0.87	<0.0001
	Total	100		
Grouped according to recognized species	Among groups	16.41	Φ_{CT}: 0.16	ns
	Among populations within groups	69.45	Φ_{SC}: 0.83	<0.0001
	Within populations	14.14	Φ_{ST}: 0.85	<0.0001
	Total	100		
Grouped according to recovered clades and sub-clades	Among groups	72.94	**Φ_{CT}: 0.72**	<0.0001
	Among populations within groups	16.62	Φ_{SC}: 0.61	<0.0001
	Within populations	10.44	Φ_{ST}: 0.89	<0.0001
	Total	100		
	cox1			

Testing assumptions	Source of variation	% of variance	Fixation index	*P*- value
Grouped according to hydrological basin [Ameca, Armeria, Coahuayana and Balsas]	Among groups	29.95	Φ_{CT}: 0.29	<0.0001
	Among populations within groups	41.85	Φ_{SC}: 0.59	<0.0001
	Within populations	28.20	Φ_{ST}: 0.71	<0.0001
	Total	100		
Grouped according to recognized species	Among groups	24.24	Φ_{CT}: 0.24	ns
	Among populations within groups	66.78	Φ_{SC}: 0.88	<0.0001
	Within populations	8.97	Φ_{ST}: 0.91	<0.0001
	Total	100		
Grouped according to recovered clades and sub-clades	Among groups	73.87	**Φ_{CT}: 0.73**	<0.0001
	Among populations within groups	21.52	Φ_{SC}: 0.82	<0.0001
	Within populations	4.61	Φ_{ST}: 0.95	<0.0001
	Total	100		

d-loop

Testing assumptions	Source of variation	% of variance	Fixation index	*P*- value
Grouped according to hydrological basin [Ameca, Armeria, Coahuayana and Balsas]	Among groups	54.32	Φ_{CT}: 0.54	<0.0001
	Among populations within groups	40.41	Φ_{SC}: 0.88	<0.0001
	Within populations	5.27	Φ_{ST}: 0.94	<0.0001
	Total	100		
Grouped according to recognized species	Among groups	31.10	Φ_{CT}: 0.31	Ns
	Among populations within groups	62.36	Φ_{SC}: 0.90	<0.0001
	Within populations	6.54	Φ_{ST}: 0.93	<0.0001
	Total	100		
Grouped according to recovered clades and sub-clades	Among groups	78.64	**Φ_{CT}: 0.78**	<0.0001
	Among populations within groups	15.34	Φ_{SC}: 0.71	<0.0001
	Within populations	6.01	Φ_{ST}: 0.93	<0.0001
	Total	100		

TABLE 4 Analyses of molecular variance of groups according to hydrological basin, recognized species, and groups recovered in phylogenetic analyses [A =

Ameca and Armeria; B1 = east Balsas; B2 = east-central Balsas; C = Lower Tuxpan River (Tamazula); D = Upper Tuxpan river (Tamazula); E = west-central Balsas] for the nuclear genes.

ACTB				
Testing assumptions	Source of variation	% of variance	Fixation index	*P*- value
Grouped according to hydrological basin [Ameca, Armeria, Coahuayana and Balsas]	Among groups Among populations within groups Within populations Total	75.81 9.97 14.22 100	Φ_{CT}: 0.75 Φ_{SC}: 0.41 Φ_{ST}: 0.85	<0.0001 <0.0001 <0.0001
Grouped according to recognized species	Among groups Among populations within groups Within populations Total	71.18 12.78 16.05 100	Φ_{CT}: 0.71 Φ_{SC}: 0.44 Φ_{ST}: 0.83	<0.0001 <0.0001 <0.0001
Grouped according to clades and sub-clades	Among groups Among populations within groups Within populations Total	80.09 1.64 18.27 100	**Φ_{CT}: 0.80** Φ_{SC}: 0.08 Φ_{ST}: 0.81	<0.0001 Ns <0.0001
RAG1				
Testing assumptions	Source of variation	% of variance	Fixation index	*P*- value
Grouped according to hydrological basin [Ameca, Armeria, Coahuayana and Balsas]	Among groups Among populations within groups Within populations Total	46.90 15.13 37.97 100	**Φ_{CT}: 0.46** Φ_{SC}: 0.28 Φ_{ST}: 0.62	<0.0001 <0.0001 <0.0001
Grouped according to recognized species	Among groups Among populations within groups Within populations Total	41.22 18.33 40.46 100	Φ_{CT}: 0.41 Φ_{SC}: 0.31 Φ_{ST}: 0.59	<0.0001 <0.0001 <0.0001
Grouped according to recovered clades and sub-	Among groups Among populations within	45.91 10.52	**Φ_{CT}: 0.46** Φ_{SC}: 0.19	<0.0001 <0.0001

clades	groups	43.57	Φ_{ST}: 0.56	<0.0001
	Within populations	100		
	Total			

Species delimitation test

The speciation model based on the species tree estimate strongly supported the assumption of six species. In the tests of species delimitation implemented in BPP, we obtained strong support (posterior probability of 1) for the tested speciation model of six *a priori* defined species within *Ilyodon* (clade A: Ameca and Armería Rivers; sub-clade B1: Atoyac sub-basin of the east Balsas River basin; sub-clade B2: Zitacuaro River of the Central Balsas and Amacuzac sub-basins of the east Balsas basin; clade C: lower Coahuayana Basin; clade D: upper Coahuayana Basin; and clade E: west Balsas River basin). However, in the posterior analyses conducted to minimize the over-splitting effect, reducing the number of species in the model and randomizing individuals or splitting populations to construct new clades, the posterior probability was 1 in all speciation models applied (Figure 3). The BPP was not sensitive for species delimitation, and no alteration of posterior probabilities of the speciation model was seen when we applied different values of root age (τ_0) and population size (θ), showing high posterior probabilities for models tested with both the A10 and A11 algorithms.

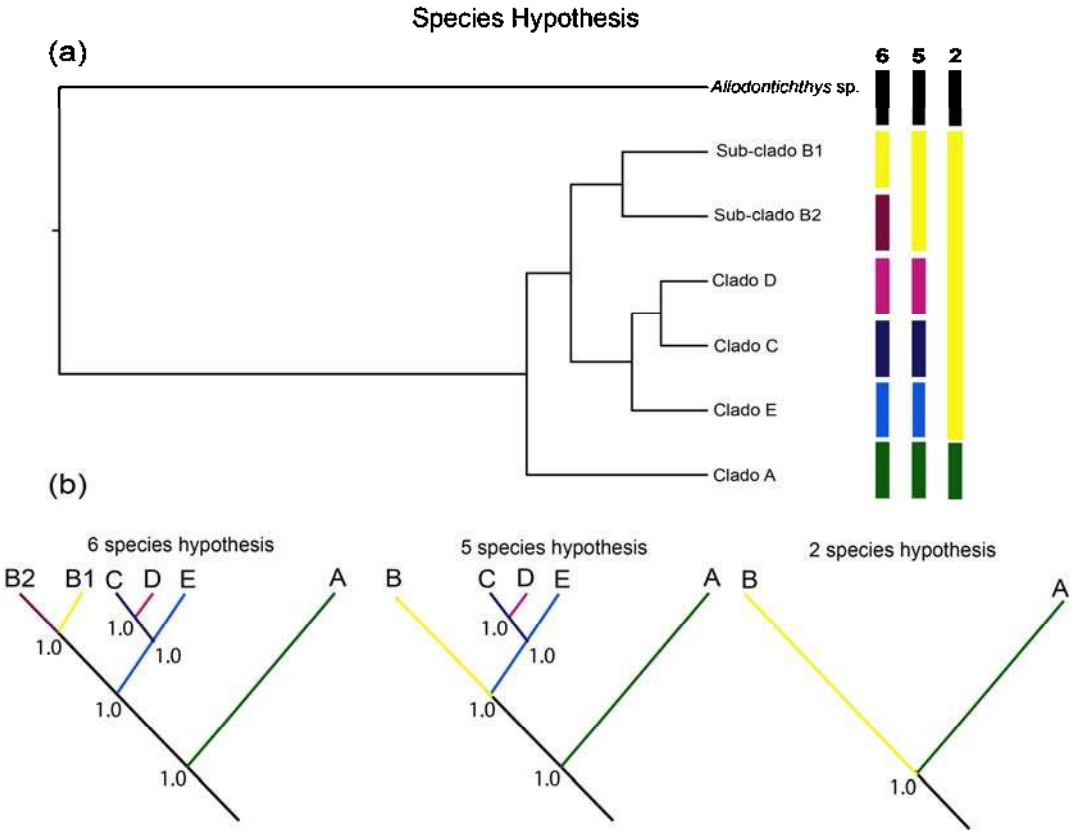


FIGURE 3 Alternative species hypothesis. (a) Guide topology based on the StarBEAST analysis. Bayesian species delimitation results for *Ilyodon* assuming six species, five species and two species (b) guide trees. The speciation probabilities are provided for each node. We consider speciation probability values >0.95 as strong support for speciation event.

Ancestral area reconstruction

Ancestral area reconstruction revealed a complex biogeographical history for *Ilyodon* species, with different events of dispersion and vicariance. The ancestral areas estimated for *Ilyodon* spp., were Armería and Balsas Rivers with a marginal probability of 0.315, followed by dispersion events to Ameca River, and one vicariance event in which Armería and Ameca Rivers were isolated from Balsas River. A second dispersal event were estimated from Balsas River toward Coahuayana River with a marginal probability of 0.769; inside of Balsas basin, several dispersal and vicariance events were estimated (Figure 4).

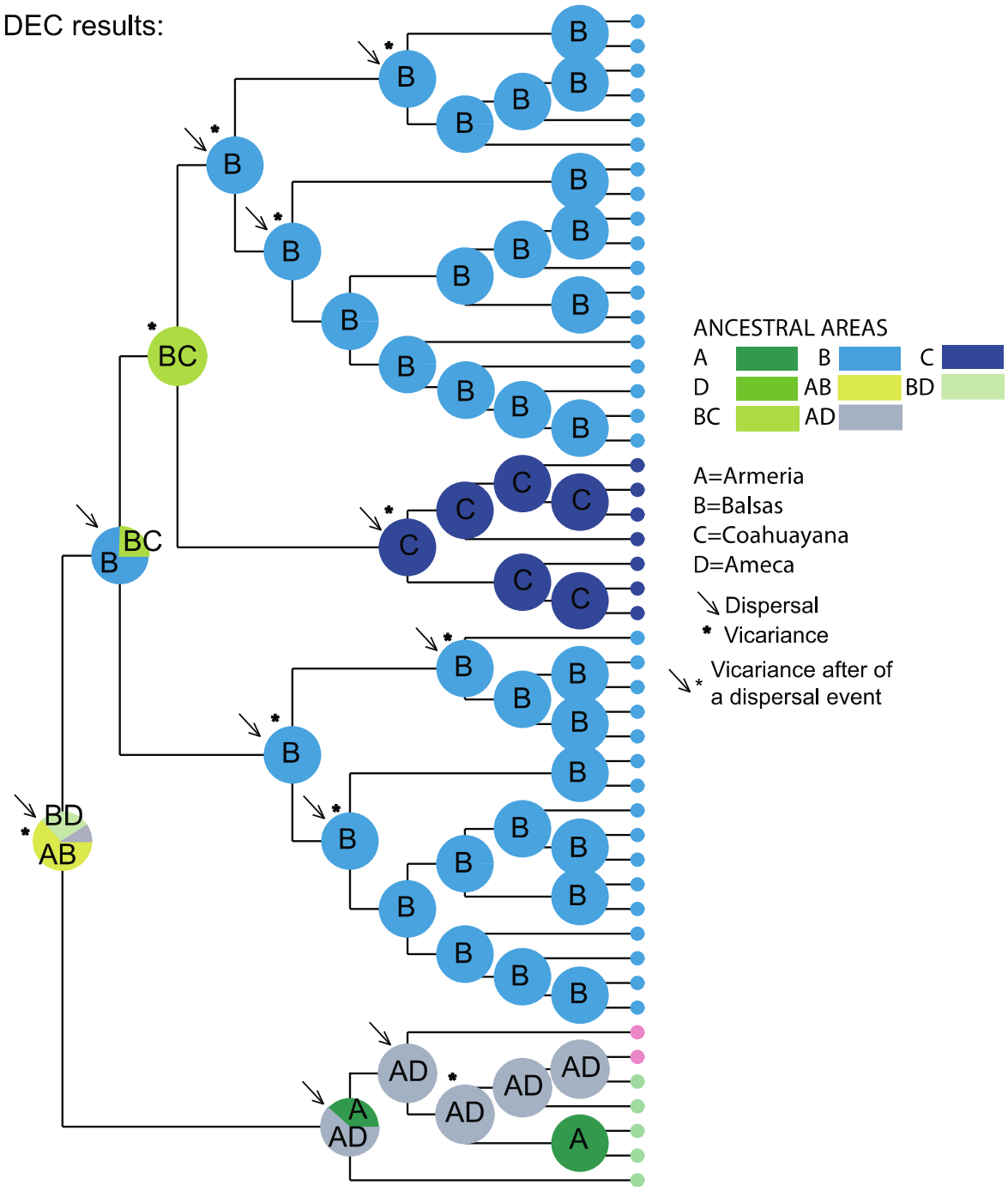


FIGURE 4 Ancestral area reconstruction with DEC for all species of genus *Ilyodon*, with the dichotomous tree obtained in BEAST and with the biogeographical regions Armería, Ameca, Coahuayana and Balsas Basins.

DISCUSSION

In Central Mexico, the Pacific coast river drainages show a configuration in which the upper areas of the basins drain parts of the Mexico Plateau, while the low sections are in the Pacific Plain (Domínguez-Domínguez et al., 2006). This area is located in an active geological zone, with high volcanic activity during the late Pliocene and early Pleistocene (1.5-3.5 Mya) (Rosas-Elguera & Urrutia-Fucugauchi, 1998). Specifically, the activity in the triple junction (boundaries of the Tepic-Zacoalco, Chapala and Colima rifts), the Tenochtitlán fault system, and the Chapala-Oaxaca fault system (García-Palomo et al., 2002; Garduño-Monroy et al., 1998; Rosas-Elguera, Ferrari, Lopez-Martinez, & Urrutia-Fucugauchi, 1997), along with climate change at the pluvial-interpluvial period beginning *ca*. 0.9 Mya, had a strong influence on the configuration of river basins in the area and on distribution of freshwater fish populations (Hewitt 2000; Smith et al., 2002; Webb & Bartlein, 1992).

The presence of the hard polytomy in the main clades recovered in the phylogenetic analyses makes the evolutionary history of *Ilyodon* difficult to interpret.

The phylogenetic, phylogeographic, and AMOVA results showed six well-differentiated groups, including main clades and sub-clades, that lack taxonomic and river basin configuration congruence: one distributed in the Armería and Ameca basins (clade A), two in the Coahuayana basin (clades C and D), and three belonging to the Balsas basin (B1, B2, and E).

At least four scenarios can be proposed for the genetic formation of the main groups: i) the ancestor of *Ilyodon* evolved in the Armería or Balsas basins and later dispersed into the other basins, as is hypothesised by the DEC analysis and is partial supported by the results showed in Domínguez-Domínguez et al. (2010). This hypothesis was also supported by a study of helminth parasites of *Ilyodon* (Martínez-Aquino, Ceccarelli, Eguiarte, Vázquez-Domínguez, & Pérez- Ponce de León, 2014). 2). (ii) isolation of *Ilyodon* populations occurred, resulting in significant genetic structure in all analyzed genes, followed by secondary contact, supported by the shared haplotypes between drainages in the nuclear genes. This would involve a significant but low number of migrants, with genetic drift purifying the mitochondrial haplotype

of migrants and acting on nuclear genes to a lesser extent than expected in large populations (Qu et al. 2012; Sefc, Payne & Sorenson, 2005). Possibly only males, which likely made up the bulk of migrants, reproduced, or selective pressures promoted the reproductive isolation of migrant females resulting in no shared haplotypes in mitochondrial genes (Qu et al., 2012). (iii) The genetic structure found in mitochondrial genes and the lack of resolution in nuclear genes could be due to relatively recent divergence of the main clades that shape *Ilyodon*, which resulted in the nuclear genes of the two most divergent groups (Ameca/Armería *vs*. central and east Balsas) showing no shared haplotypes, while divergent groups of the west Balsas basin and Coahuayana Basin do exhibit shared haplotypes. This pattern in which nuclear genes resolve some structure but not to the extent of mitochondrial genes, due to recent diversification, has been reported for other freshwater fishes of Central Mexico (Pérez-Rodríguez et al., 2009). The last and more likely scenario for *Ilyodon* genetic groups evolution is (iv) a recent and simultaneous differentiation of the six genetic groups, that is supported by the similar values of genetic distance between them, the hard basal polytomy, and the lack of shared haplotypes in mitochondrial genes accompanied by incomplete lineage sorting in nuclear genes (Ballard, Chernoff, & James, 2002; Qu et al., 2012), and this also explain the low marginal probability found in the DEC analyses for the most plausible ancestral area. We suggest that our data are not enough for determinate the ancestral area of *Ilyodon.* In any case, is evident that the biogeographical history of *Ilyodon* is more complex than previously reported (Doadrio & Domínguez, 2004; Domínguez-Domínguez et al., 2010), mainly for lineage evolved in the Balsas Basin, which several events of dispersal and vicariance were estimated in DEC (Figure 4).

This complex history seems to be closely related with the complex hydrological system in the area. Genetic patterns related to connection and disconnection in Pacific slope drainages are partially supported by the goodeine species *Allodontichthys* spp., *Xenotoca eiseni*, and *Xenotoca melanosoma* (Domínguez-Domínguez et al., 2010; Piller et al. 2015; Webb, 2002) and other freshwater fish, such as *Algansea aphanea, Moxostoma* sp*,* and *Astyanax aeneus* (Ornelas-García, Domínguez-Domínguez, & Doadrio, 2008; Pérez-Rodríguez et al., 2009, 2015). Incomplete genetic data for most of these groups

prevent accurate comparisons with *Ilyodon*. Certain geological events provide an independent line of evidence, such as the uplift of the Sierra de Manantlan and Cacoma, the volcanic activity of the Talpa-Mascota graben, dated 3.6 Mya (Carmichael, Lange, & Luhr, 1996), and the reactivation of the Colima and Tamazula graben in the Pliocene. These geologic events are related to the configuration of the river basin beds (Allan 1986; Garduño-Monroy et al., 1998).

Coahuayana groups

Two groups (clades C and D) were recovered in the specimens from Coahuayana River basin, one distributed in the lower and other in the upper Cohuayana basin. A genetic split between the upper and lower Coahuayana populations has been suggested, based on two cytotypes, one distributed in the upper and other in the lower Coahuayana (Turner, Grudzien, Adkisson, & Worrell, 1985). This is also supported for *Allodontichthys*, with two related species showing the same pattern: *Allodontichthys hubbsi* mainly distributed in the lower Coahuayana, and *Allodontichthys tamazulae* in the upper Coahuayana, but with higher divergence than in *Ilyodon* (Doadrio & Domínguez, 2004; Domínguez-Domínguez et al., 2010; Webb 2002).

The relationships within these groups were not resolved in the phylogenetic trees, and the haplotype networks indicate different relationships depending on the marker analyzed. Also, these two clades showed the lowest genetic distances among all the pairwise comparisons (Tables S5, S6 and S7). Mitochondrial and nuclear genes showed no shared haplotypes between clades C and D, with the exception of a single haplotype in *RAG1*. The nuclear *ACTB* showed the lower Coahuayana basin samples to share haplotypes with the Ameca-Armería clade, and the upper Coahuayana basin shared haplotypes with the west Balsas population. For the *RAG1* gene, the lower Coahuayana specimens possessed unique haplotypes, with the exception of one shared with the upper Coahuayana, Ameca, and Armería basins. The upper Coahuayana specimens shared haplotypes with lower Coahuayana, Ameca-Armería, and west Balsas. The most plausible scenario is a recent isolation event of the ancestor of these two Coahuayana groups, one in the upper and other in the lower Coahuayana basin, as could be indicated by the low genetic distances between them (2.3% for *d-loop* and 0.5% for *cytb* and *coxI*), and, by the DEC

analysis, which showed several dispersal and vicariance events that separated these two groups inside Coahuayana River (marginal probability=1.0). In this scenario, the relationships of the nuclear genes are a product of incomplete lineage sorting or of secondary contact between the lower Coahuayana and Armería-Ameca populations. Secondary contact is also supported by the occurrence of *Allodontichthys zonistius*, a species previously considered endemic to Armeria drainages, in the lower reaches of the Coahuayana River, probably related to a river piracy event of nearby (15 km) tributaries (Webb, 2002). Cytogenetic data in *Ilyodon* show the cytotypes from the lower Coahuayana to be more closely related to the Armería population than those of the upper Coahuayana (Turner et al., 1985). Evidence of a founded population in the Ameca and Coahuayana drainages, the source of which was an Armería population, has also been suggested for *Allodontichthys* (Webb 2002). In most of the genes analyzed, the upper Coahuayana population seems to be close related with west Balsas specimens and even the genetic distances are similar between lineage within Coahuayana than between upper Coahuayana and West Balsas. This may indicate that west Balsas specimens had secondary contact or that the upper Coahuayana population was founded by specimens from Balsas River as indicate the DEC.

Balsas basin

Three well-differentiated groups that show significant genetic structure and high divergence were identified within the Balsas River basin (Figure 2 and Figure S3), the clades and subclades B1, B2, and E. This scenario of the formation of differentiated groups within the Balsas river basin is supported by the results of DEC, in which several dispersal and vicariance events have been promoted the actual distribution of these three groups (Figure 4). Isolation of other fish species in the Balsas watershed has been documented, including restriction of *Notropis boucardi* (Schönhunth & Doadrio, 2003) to the east Balsas and evidence for two divergent groups of *Astyanax,* one distributed in west and other in the east Balsas basin (Ornelas-García et al., 2008). The significant genetic structure and the relatively high genetic distances observed in the Balsas populations are at the same rank with that seen between populations of isolated drainages. This could be explained by ancient ecological or geological

barriers within the basin, which is inferred in the mitochondrial genes. The shared haplotypes in *RAG1*, along with close relationships without shared haplotypes in *ACTB*, in the Balsas basin populations could be indicative of secondary contact and gene flow between previously isolated groups, as was previously suggested for the lower Coahuayana and Armería samples. Evidence of a gene flow process within Balsas populations was shown by the central Balsas specimens from Zitacuaro River (Figures S1, S3 and S4), which showed a close relationship to an east Balsas population in mitochondrial genes, but a closer relationship to west Balsas populations in the *ACTB* gene. Gene flow previous to isolation and gene flow among *Ilyodon* populations have been suggested previously (Webb, 2002). These processes are as complex as the geological and climatic history of the Balsas depression. Geological activity has been suggested to have similarly effect on other endemic species, including spiders, butterflies, birds, amphibians, and reptiles, especially in a sector of the Amacuzac sub-basin (Escalante-Pliego, Navarro, & Peterson, 1993; Luna-Reyes, Llorente-Bousquets, & Luis-Martínez, 2008; Nieto-Castañeda, Pérez-Miguel, & García-Cano, 2014). Also, the Balsas depression is located between the Trans-Mexican Volcanic Belt and the Sierra Madre del Sur (Castañeda-Rico, León-Paniagua, Vázquez-Domínguez, & Navarro-Sigüenza, 2014; Ferrusquía-Villafranca, 1993) which has been active from the Eocene and Oligocene to the present (Yarza de De la Torre, 1992), specifically the Guerrero and Morelos platforms, the Tierra Caliente metamorphic complex and Guerrero terrane, the Taxco fault and Arcelia graben, the Tenochtitlan fault system, and the Chapala-Oaxaca fault system (García-Palomo et al., 2002; González-Torres et al., 2013).

Taxonomic implications

The species delimitation tests did not resolve the species-level taxonomy of *Ilyodon*. We suggest that a speciation model based on different criteria (phylogenetic relationships, genetic structure, haplotype networks, genetic distances, and geography) is not informative when phylogenetic relationships are unresolved, low genetic divergences in nuclear genes are observed, or shared nuclear haplotypes are present, violating the algorithm assumptions in the species tree and species delimitation analyses. The BEAST analysis

assumes a model in which the separation of species is complete, if this separation is nor complete, can result in an incorrectly specified guide tree or speciation model (Leaché & Fujita, 2010) that detects species before they are fully separated (incipient species) (Heled, Bryant, & Drummond, 2013). Eberle, Warnock & Ahrens (2016) have shown that uncertainties in analyses implemented in BPP, such as guide tree inference, individual species assignment, and prior parameter choice, may impact the accuracy of results. However, in our tests, use of different prior parameters (τ_0 and θ) did not affect the results, which showed high values of posterior probability in all analyses (>0.95). We consider that basal hard polytomy obtained in the phylogenetic tree, and possible over-splitting, could explain the results obtained with the species delimitation tests (similar high posterior probabilities for two to six species of *Ilyodon*). Hence, we consider that our data do not meet the conditions and parameters necessary for the species delimitation test. It has been shown that both BEAST and BPP may be impacted by putative incomplete lineage sorting and are inadequate for delimiting very young species, which are difficult to distinguish on the basis of molecular or morphological data alone (Eberle et al., 2016).

Despite the failure of the species delimitation test, the finding of six well-differentiated lineages, together with the significant differentiation between them revealed by AMOVA, seemed to indicate a separate genetic identity of each group. The genetic distances calculated with mitochondrial genes are similar to those previously reported between *Ilyodon* species (Doadrio & Domínguez 2004; Domínguez-Domínguez et al., 2010; Webb et al. 2004). The highest genetic distance was found between the *d-loop* and *cytb* gene of the Ameca-Armería population (clade A) with respect to the other clades (B1, B2, C, D, and E). For nuclear genes, the genetic distance between these clades ranged from 0.1-0.2% (Table S7). Previous studies of other goodeine species showed similar genetic distances and lack of resolution in phylogenetic analyses, as did some species of *Allotoca* and *Goodea,* associated with recent isolations (< 1 Mya) or secondary contact events promoted by river piracy or founder effect (Corona-Santiago et al., 2015; Doadrio & Domínguez, 2004; Domínguez-Domínguez et al., 2010;).

Not all species within *Ilyodon* were identified as monophyletic in the phylogenetic results. The pattern of species or genera mixed in the phylogenetic tree has also been reported for other freshwater fishes of Mexico (Corona-Santiago et al., 2015; McMahan, Geheber & Piller, 2010; Ornelas-García et al., 2008; Pérez-Rodríguez et al., 2009). The variation among *Ilyodon* with morphological recognized groups has been associated with trophic structure, with variation in the shape and arrangement of the head, mouth, and teeth and in fish size influenced by habitat and feeding (Grudzien & Turner 1984a,b; Kingston, 1979; Turner & Grosse, 1980).

All currently recognized *Ilyodon* species were described morphologically (Paulo-Maya & Trujillo-Jiménez, 2000), with no descriptions based on molecular analyses. In the present study, results of molecular analyses of five genes disagree with the currently recognized taxonomy of *Ilyodon*. Our results showed lower genetic distance between most clades than the average found for all the recognized species of goodeines (1.7% in *cytb*), as well as shared haplotypes among most main clades in nuclear genes. We also found higher genetic divergence between the *I. whitei* sampled in west and east Balsas than between samples from the Balsas and Coahuayana basins, with evidence of interbreeding between highly divergent lineages. We found significant geographic structure in *Ilyodon*, but not concordance with the five previously recognized species *I. furcidens*, *I. xantusi*, *I. whitei*, *I. lennoni*, and *I. cortesae* showing polyphyletic relationships (Tables 3 and 4). Hence, taxonomic interpretations are difficult to assess. Further genetic and morphological analyses need to be conducted to provide a clearer picture of the taxonomy and evolution of genetically divergent populations of *Ilyodon*, but some suggestions can be made according to the priority principle and type locality.

All specimens belonging to clade A were identified as *Ilyodon furcidens.* Although the type locality was given as Cape San Lucas by Eigenmann (1907), later researchers state the type locality to be Río Colima, a tributary of Armería River basin (Hubbs & Turner, 1939). Although we did not include samples from the Colima River in the lower Armería basin, and because of high genetic differentiation found within other drainages, we provisionally designate the specimens of clade A as *I. furcidens.* Since the type locality for *I. xantusi* is the

Colima River, in the Armería River basin, more samples from the lower Armería need to be examined to draw a robust conclusion.

The specimens of sub-clade B1 were identified as *Ilyodon whitei,* but this group showed high genetic divergence and significant structure with respect to other samples, so we considered this group as a differentiated group of *Ilyodon "whitei".*

Specimens belonging to sub-clade B2 were also identified as *I. whitei*. This group included specimens collected at its type locality (upper tributaries of Balsas, at Cuautla and Yautepec, in Morelos state), and we considered this clade as *I. whitei*. Within this clade we found genetic flow between east Balsas (B2) and west Balsas (E) in nuclear genes.

Specimens belonging to clade C were identified as *I. furcidens*, but, because of the genetic divergence from other groups and the possibility of interbreeding with Armería populations, we considered this group a differentiated group of *Ilyodon "furcidens."*

Specimens belonging to clade D were identified as *I. whitei*, but, based on the divergence from other genetically identified *I. whitei*, and the high structure found, we considered this group a differentiated group of *Ilyodon "whitei."*

The specimens belong to clade E were identified as *I. whitei*, *I. cortesae*, and *I. lennoni*, with the latter two species collected from the type locality. Since specimens identified as *I. whitei* in the type locality (upper tributaries of Balsas, at Cuautla and Yautepec) belong to the sub-clade B2, the specimens of clade E must be considered *Ilyodon "lennoni,"* while *I. cortesae* was not considered a valid species.

Our results show more complex evolutionary and taxonomic history of *Ilyodon* than was previously revealed by molecular studies (Doadrio & Domínguez, 2004; Domínguez-Domínguez et al., 2010). In this scenario, and due to the high level of morphological differentiation, a broader taxonomic and systematic work for *Ilyodon* species is necessary to confirm the taxonomic status of each described species.

Implications for conservation

We identified at least six genetic groups, with geographic correspondence, in analyses performed with mtDNA and with concatenated mtDNA + nDNA. Each

of these groups should be considered an operational conservation unit (OCU), and effective protection of the OCUs could guarantee the conservation and preservation of the genetic pool (Doadrio, Perdices & Machordom, 1996) found within genus *Ilyodon*. *Ilyodon whitei* is catalogued since 1996 in the red list of endangered species as critically endangered (Contreras-Balderas & Almada-Villela, 1996), and *I. furcidens* is catalogued as threatened in the NOM-059 for the Ministry of Environmental and Natural Resources (SEMARNAT, 2010). Based on our genetic groups recovered for mitochondrial genes, we suggest a re-evaluation of the conservation status of the *Ilyodon* species or populations.

ACKNOWLEDGMENTS

This work was partially conducted while M. Sc. Studies R. G. Beltrán-López.

We thank Topiltizin Contreras MacBeath and Manuel Rivas González for their help with sampling, to Silvia Perea and Berenice García for their comments to improve the study. Finally, we thank the Ministerio de Economía y Competitividad for funding the project (CGL-2013 41375-P).

REFERENCES

Alfaro ME, Zooler S, Lutzoni F (2003) Bayes or bootstrap? A simulation study comparing the performance of Bayesian Markov chain Monte Carlo sampling and bootstrapping in assessing phylogenetic confidence. Mol Biol Evol **20:** 255–266. PMID: 12598693

Allan J (1986) Geology of the Northern Colima and Zacoalco grabens, southwest Mexico: late Cenozoic rifting in the Mexican Volcanic Belt. Geol Soc Am Bull **97:** 473–485.

Ballard JW, Chernoff B, James AC (2002) Divergence of mitochondrial DNA is not corroborated by nuclear DNA, morphology, or behavior in Drosophila simulans. Evolution **56:** 527–545. PMID: 11989683

Bandelt HJ, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol **16:** 37–48. PMID: 10331250

Barbour CD (1973) The systematics and evolution of the genus *Chirostoma* Swainson (Pisces: Atherinidae). Tulane Stud Zool Bot **18:** 97–141.

Castañeda-Rico S, León-Paniagua L, Vázquez-Domínguez E, Navarro-Sigüenza G (2014) Evolutionary diversification and speciation in rodents of the Mexican lowlands: The *Peromyscus melanophrys* species group. Mol Phylogenet Evol **70:** 454–463. PMID: 24140979

Carmichael ISE, Lange RA, Luhr J (1996) Quaternary minettes and associated volcanic rocks of the Mascota, western Mexico: a consequence of plate extension above a subduction modified mantle wedge. Contrib Mineral Petr **124:** 302–333.

Contreras-Balderas S, Almada-Villela P (1996**)** *Ilyodon whitei*. *The IUCN Red List of Threatened Species*. Version 1996: e. Accessed online 05 March 2016 at: www.iucnredlist.org

Corona-Santiago DK, Doadrio I, Domínguez-Domínguez O (2015) Evolutionary History of the Live-bearing endemic *Allotoca diazi* species complex (Actinopterygii, Goodeinae): evidence of founder effect events in the Mexican Prehispanic period. Plos One **10:** e0124138. D oi: 10.1371/journal.pone.0124138. PMID: 25946217

De Queiroz K (2007) Species concepts and species delimitation. Syst Biol **56:** 879–886. PMID: 18027281

Doadrio I, Perdices A, Machordom A (1996) Allozymic variation of the endangered killifish *Aphanius iberus* and its application to conservation. Environ Biol Fish **45:** 259–271.

Doadrio I, Domínguez O (2004**)** Phylogenetic relationships whitin the fish family Goodeidae based on cytochrome b sequence data. Mol Phylogenet Evol **31:** 416–430. PMID: 15062784

Domínguez-Domínguez O, Doadrio I, Martínez-Meyer E, Zambrano L, Pérez-Ponce de León G (2006) Using ecological niche modeling as a conservation tool for freshwater species: the live-bearing fishes (Goodeidae) in Central Mexico. Conserv Biol **20:** 1730–1739. PMID: 17181808

Domínguez-Domínguez O, Pérez Ponce de León G (2007) Los goodeidos, peces endémicos del centro de México. CONABIO. *Biodiversitas* **75:** 12–15.

Domínguez-Domínguez O, Alda F, Pérez-Ponce de León G, García-Garitagoitia JL, Doadrio I (2008) Evolutionary history of the endangered fish

Zoogoneticus quitzeoensis (Bean, 1898) (Cyprinodontiformes: Goodeidae) using a sequential approach to phylogeography base on mitochondrial and nuclear DNA data. BMC Evol Biol **8:** 161. doi: 10.1186/1471-2148-8-161 PMID: 18503717

Domínguez-Domínguez O, Pérez-Rodríguez R, Escalera-Vázquez LH, Doadrio I (2009) Two new species of the genus Notropis Rafinesque, 1817 (Actinopterygii, Cyprinidae) from the Lerma River Basin in Central Mexico. Hidrobiologica **19:** 159–172.

Domínguez-Domínguez O, Pedraza-Lara C, Gurrola-Sánchez N, Perea S, Pérez-Rodríguez R, Israde-Alcántara I, Garduño-Monroy VU, Doadrio I, Pérez-Ponce de León G, Brooks DR (2010) Historical Biogeography of the Goodeinae (Cyprinodontiformes). In: Uribe-Aranzabal MC, Grier HJ (eds), Viviparous Fishes II. New Life Publications, Florida, pp 13–30.

Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol Biol Evol **29:** 1969–1973. PMID: 22367748

Eberle J, Warnock RCM, Ahrens D (2016) Bayesian species delimitation in *Pleophylla* chafers (Coleoptera) – the importance of prior choice and morphology. BMC Evol Biol **16:** 94. doi: 10.1186/s12862-016-0659-3

Eigenmann CH (1907) The poeciliid fishes of Rio Grande do Sul and the La Plata Basin. Proc U S Nat Mus **32:** 425–433.

Escalante-Pliego P, Navarro A, Peterson T (1993) A geographic, ecological and historical analysis of land bird diversity in Mexico. In: Ramamoorthy TP, Bye R, Lot A, Fa J (eds), Biological Diversity of Mexico: Origins and Distribution. Oxford, Nueva York, pp 281–307.

Eschmeyer WN, Fricke R Dan der Laan R (eds). **CATALOG OF FISHES: GENERA, SPECIES, REFERENCES.** (http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp). Accessed online 1 March 2016. [This version was edited by Bill Eschmeyer.]

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes application to human mitochondrial DNA restriction data. Genetics **131:** 479–491. PMID: 1644282

Excoffier L, Lischer HE (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour **10:** 564–567. PMID: 21565059

Farris JD, Källersjö M, Kluge AG, Bult C (1995) Constructing a significance test for incongruence. Syst Biol **44:** 570–572.

Ferrari L, Conticelli S, Vaggelli G, Potrone CM, Manetti P (2000) Late Miocene volcanism and intra-arc tectonics during the early development on the Trans-Mexican volcanic belt. Tectonophysics **318:** 161–185.

Ferrusquía-Villafranca I (1993) Geología de México: una sinopsis. In: Ramamoorthy TP, Bye R, Lot A, Fa J (eds), Biological Diversity of Mexico: Origins and Distribution. Oxford, Nueva York, pp 3–108

García-Palomo A, Macías JL, Arce JL, Capra L, Garduño VH, Espíndola JM (2002) *Geology of Nevado de Toluca Volcano and surrounding areas, central Mexico.* Boulder, Colorado, Geological Society of America Map and Chart Series MCH089, 26 p.

Garduño-Monroy V, Saucedo-Girón R, Jiménez Z, Gavilanes-Ruiz J, Cortés A, Uribe-Cifuentes R (1998) La falla Tamazula-límite suroriental del bloque Jalisco y sus relaciones con el complejo volcánico de Colima, México. Rev Mex Cienc Geol **15:** 132–144.

González-Torres EA, Morán-Zenteno DJ, Mori L, Díaz-Bravo B, Martiny BM, Solé J (2013) Geochronology and magmatic evolution of the Huautla volcanic field: last stages of extinct Sierra Madre del Sur igneous province of southern Mexico. Int Geol Rev **55:** 1145–1161.

Grudzien TA, Turner BJ (1984a) Genic identity and Geographic Differentiation of Trophically Dichotomous *Ilyodon* (Teleostei: Goodeidae). Copeia **1984:** 102–107.

Grudzien TA, Turner BJ (1984b) Direct evidence that the *Ilyodon* morphs are a single biological species. Evolution **38:** 402–407.

Heled J, Drummond AJ (2010) Bayesian inference of species trees from multilocus data. Mol Biol Evol **27:** 570–580. doi: 10.1093/molbev/msp274.

Heled J, Bryant D, Drummond AJ (2013) Simulating genes trees under the multispecies coalescent and time-dependent migration. BMC Evol Biol **13:** 44. doi: 10.1186/1471-2148-13-44

Hewitt G (2000) The genetic legacy of the Quaternary ice ages. Nature **405:** 907–913, doi:10.1038/35016000

Hubbs LC, Turner LC (1939) Studies of the fishes of the order Cyprinodontes. XVI. A revision of the Goodeidae. Misc Publ, Museum of Zoology. University of Michigan. 42.

Huson DH, Bryant D (2006) Application of Phylogenetic Networks in Evolutionary Studies. Mol Biol Evol **23:** 254–267. PMID 16221896

Jordan DS, Gilbert CH (1882) Catalogue of the fishes collected by Mr. John Xantus at Cape San Lucas, which are now in the U. S. National Museum, with descriptions of eight new species. Proc U S Nat Mus **5:** 353–371.

Kingston DIL (1979) *Behavioral and morphological studies of the goodeid genus Ilyodon, and comparative behavior of the fishes of the family Goodeidae*. PhD Thesis, University of Michigan.

Lanfear R, Calcott B, Ho SY, Guindon S (2012) PartitionFinder: Combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol Biol Evol **29:** 1695–1701. PMID: 22319168

Leaché AD, Fujita MK (2010) Bayesian species delimitation in West African forest geckos (*Hemidactylus fasciatus*). Proc R Soc B **277:** 3071–3077. doi: 10.1098/rsBp.2010.0662

Lee WJ, Conroy J, Howell WH, Kocher TD (1995) Structure and evolution of teleost mitochondrial control regions. J Mol Evol **41:** 54–66. PMID: 7608989

Librado P, Rozas J (2009) DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics **25:** 1451–1452. PMID: 19346325

Luna-Reyes M, Llorente-Bousquets J, Luis-Martínez A (2008) Papilionoidea de la Sierra de Huautla, Morelos y Puebla, México (Insecta: Lepidoptera). Rev Biol Trop **56:** 1677–1716. PMID: 19419076

Machordom A, Doadrio I (2001) Evidence of a cenozoic Betic-Kabilian connection based on freshwater fish phylogeography (*Luciobarbus*, Cyprinidae). Mol Phylogenet Evol **18:** 252–263. PMID:11161760

Martínez-Aquino A, Ceccarelli FS, Eguiarte LE, Vázquez-Domínguez E, Pérez-Ponce de León G (2014) Do the historical biogeography and evolutionary history of the digenean Margotrema spp. across central Mexico mirror

those of their freshwater fish hosts (Goodeinae)?. Plos one **9:** e101700. doi:10.1371/journal.pone.0101700. PMID: 24999998

Mateos M, Sanjur OI, Vrijenhoek RC (2002) Historical biogeography of the livebearing fish genus *Poeciliopsis* (Poeciliidae: Cyprinodontiformes). Evolution **56:** 972–984. PMID: 12093032

Mateos M (2005) Comparative phylogeography of livebearing fishes in the genera *Poeciliopsis* and *Poecilia* (Poeciliidae: Cyprinodontiformes) in central Mexico. J Biogeogr **32:** 775–780. doi: 10.1111/j.1365-2699.2005.01236.x

McMahan CD, Geheber AD, Piller KR (2010) Molecular systematics of the enigmatic Middle American genus *Vieja* (Teleostei: Cichlidae). Mol Phylogenet Evol **57:** 1293–1300. PMID: 20849962

Meek SE (1904) The fresh-water fishes of Mexico of the isthmus of Tehuantepec. Publ Field Columbian Mus Zool Ser **5:** 1-252.

Meyer MK, Föerster W (1983) Eine neue *Ilyodon*-Art aus Guerrero, Mexiko (Osteichthyes: Goodeidae). Zool Abh (Dres) **38:** 257-263.

Miller RR, Minckley WL, Norris SM (2005) *Freshwater Fishes of Méxic*o. Chicago: The University of Chicago Press.

Near TJ, Benard MF (2004) Rapid allopatric speciation in logperch darters (Percidae: Percina). Evolution **58:** 2798–2808. PMID:15696757

Nieto-Castañeda G, Pérez-Miguel L, García-Cano AA (2014) New records of spiders (Arachnida: Araneae) from the Balsas Basin in central Mexico. Southwest Nat **59:** 426–430. doi: http://dx.doi.org/10.1894/JC-77.1

Ornelas-García CP, Domínguez-Domínguez O, Doadrio I (2008) Evolutionary history of the fish genus Astyanax Baird and Girard (1854) (Actinopterygii, Characidae) in Mesoamerica reveals multiple morphological homoplasies. BMC Evol Biol **8:** 340. PMID: 19102731

Paulo-Maya J, Trujillo-Jiménez P (2000) Nueva especie de *Ilyodon* (Cyprinodontiformes: Goodeidae) de la cuenca del río Balsas, México. Rev Biol Trop **48:** 465–472. PMID:11354953

Pérez-Rodríguez R, Domínguez-Domínguez O, Pérez-Ponce de León G, Doadrio I (2009) Phylogenetic relationships and biogeography of the genus Algansea Girard (Cypriniformes: Cyprinidae) of central Mexico inferred from molecular data. BMC Evol Biol **9:** 223. PMID: 19735558

Pérez-Rodríguez R, Domínguez-Domínguez O, Doadrio I, Cuevas-García E, Pérez-Ponce de León G (2015) Comparative historical biogeography of three groups of Neartic freshwater fishes across central Mexico. J Fish Biol **86:** 993–1015. PMID: 25644334

Piller KR, Kenway-Lynch CS, Camak DT, Domínguez-Domínguez O (2015) Phylogeography and population structure of the imperilled redtail splitfin (Goodeidae: *Xenotoca eiseni*): Implications for Conservation. Copeia **103:** 440–454. doi: http://dx.doi.org/10.1643/CI-14-067

Quenouille B, Bermingham E, Planes S (2004) Molecular systematics of the damselfishes (Teleostei: Pomacentridae): Bayesian phylogenetic analyses of mitochondrial and nuclear DNA sequences. Mol Phylogenet Evol **31:** 66–88. PMID: 15019609

Qu Y, Zhang R, Quan Q, Song G, Hsien S, Lei F (2012**)** Incomplete lineage sorting or secondary admixture: disentangling historical divergence from recent gene flow in the Vinous-throated parrotbill (*Paradoxornis webbianus*). Mol Ecol **21:** 6117–6133. doi: 10.1111/mec.12080.

Rambaut A, Drummond A (2007) Tracer v1.5. Available: http://tree.bio.ed.ac.uk/software/tracer.

Rannala B, Yang Z (2013) Improved Reversible Jump Algorithms for Bayesian Species Delimitation. Genetics **194:** 245–253. doi:10.1534/genetics.112.149039

Ree RH, Moore BR, Webb CO, Donoghue MJ (2005) A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. Evolution **59:** 2299-2311. PMID: 16396171

Ree RH, Smith SA (2008) Maximum-Likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. Syst Biol **57:** 4-14. PMID: 18253896

Robalo JI, Sousa-Santos C, Levy A, Almada VC (2006) Molecular insights on the taxonomic position of the paternal ancestor of the *Squalius alburnoides* hybridogenetic complex. Mol Phylogenet Evol **39:** 276–281. PMID: 16213165

Ronquist F, Teslenko M, Van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP (2012) MrBayes 3.2: Efficient

Bayesian Phylogenetic Inference and Model Choice across a Large Model Space. Syst Biol **61:** 539–542. PMID: 22357727

Rosas-Elguera J, Ferrari L, Lopez-Martinez M, Urrutia-Fucugauchi J (1997) Stratigraphy and tectonics of the Guadalajara region and the triple junction area, western Mexico: Int Geol Rev **39:** 125–140.

Rosas-Elguera J, Urrutia-Fucugauchi J (1998**)** Tectonic control on the volcano-sedimentary sequence of the Chapala graben, western Mexico. Int Geol Rev **40:** 350–362. doi: 10.1080/00206819809465214

Schönhuth S, Doadrio I (2003) Phylogenetic relationships of Mexican minnows of the genus *Notropis* (Actinopterygii, Cyprinidae). Biol J Linnean Soc **80:** 323–337. doi:10.1046/j.1095-8312.2003.00241.x

Schönhuth S, Doadrio I, Domínguez-Domínguez O, Hillis DM, Mayden RL (2008) Molecular evolution of Southern North American Cyprinidae (Actinopterygii), with the description of the new genus *Tampichthys* from central Mexico. Mol Phylogenet Evol **47:** 729–756. PMID: 18353691

Sefc KM, Payne RB, Sorenson MD (2005) Genetic continuity of brood-parasitic indigobird species. Mol Ecol **14:** 1407–1419. PMID: 15813780

SEMARNAT. [Cited 13 February 2016]. Secretaría de Medio Ambiente y Recursos Naturales. NORMA Oficial Mexicana NOM–059–ECOL–2010, Protección ambiental–Especies nativas de México de flora y fauna silvestres-Categorías de riesgo y especificaciones para su inclusión, exclusión o cambio-Lista de especies en riesgo. Diario Oficial de la Federación, segunda sección. 30 de diciembre de 2010. Available: http://www.semarnat.gob.mx/leyesynormas/SEMARNAT%20DOF/Norma%20Oficial%20Mexicana%20NOM-059-SEMARNAT-2010.pdf

Silvestro D, Michalak I (2012). raxmlGUI: a graphical front-end for RAxML. Org Divers Evol **12:** 335–337. doi: 10.1007/s13127-011-0056-0

Smith ML (1980) *The evolutionary and ecological history of the fish fauna of the Rio Lerma basin, Mexico*. Unpublished. PhD. Thesis, Ann Arbor, University, Michigan.

Smith GR, Dowling T, Gobalet KW, Lugaski T, Shiazawa D Evans RP (2002) Biogeography and timing of evolutionary events among Great Basin fishes. *In*: Hershler R, Madsen DB Currey DR (eds), Great Basin Aquatic

Systems History. Smithsonian Contributions to Earth Sciences, Number 33, pp 175–234.

Sousa-Santos C, Robalo JI, Collares-Pereira MJ, Almada VC (2005) Heterozygous indels as useful tools in the reconstruction of DNA sequences and in the assessment of ploidy level and genomic constitution of hybrid organisms. DNA Seq **16:** 462–467. PMID: 16287626

Stamatakis A (2014) RAxML Version 8: A tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies. Bioinformatics **30:** 1312–1313. PMID: 24451623

Stephens M, Donelly P (2003) A comparison of Bayesian methods for haplotype reconstruction from population genotype data. Am J Hum Genet **73:** 1162–1169. PMID: 14574645

Swofford DL (2003) PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5. Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Mol Biol Evol **28:** 2731–2739. PMID: 21546353

Turner BJ, Grosse DJ (1980) Trophic differentiation in *Ilyodon,* a genus of stream dwelling Goodeid fishes: speciation versus ecological polymorphism. Evolution **34:** 259–270.

Turner BJ, Grudzien TA, Adkisson KP, Worrell RA (1985) Extensive chromosomal divergence within a single river basin in the goodeid fish, *Ilyodon furcidens*. Evolution **39:** 122–134.

Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PDN (2005) DNA barcoding Australia´s fish species. Philos Trans R Soc Lond B: Biol Sci **360:** 1847–1857. PMID: 16214743

Webb SA (2002) Molecular systematic of the genus *Allodontichthys* (Cyprinodontiformes: Goodeidae). Rev Fish Biol Fish **12:** 193–205

Webb T, Bartlein PJ (1992) Global changes during the last 3 million years: climatic controls and biotic responses. Annu Rev Ecol Syst **23:** 141–173.

Webb SA, Graves JA, Macías-García C, Magurran AE, Foighil DÓ, Ritchie MG. (2004) Molecular phylogeny of the livebearing Goodeidae (Cyprinodontiformes). Mol Phylogenet Evol **30:** 527–544. PMID: 15012937

Yang ZH, Rannala B (2010) Bayesian species delimitation using multilocus sequence data. Proc Natl Acad Sci U S A **107:** 9264–9269. doi: 10.1073/pnas.0913022107

Yang ZH, Rannala B (2014) Unguided species delimitation using DNA sequence data from multiple loci. Mol Biol Evol **31:** 3125–3135. doi: 10.1093/molbev/msu279

Yang ZH (2015) The BPP program for species tree estimation and species delimitation. Current Zoology **61:** 854–865.

Yarza de De la Torre E (1992) Los volcanes del Sistema volcánico transversal. Volcanes de México. Instituto de Geografía, UNAM, México. doi: http://dx.doi.org/10.14350/rig.30441

Yu Y, Harris AJ, Blair C, He J (2015) RASP (Reconstruct Ancestral State in Phylogenies): A tool for historical biogeography. Mol Phylogenet Evol **87:** 46-49. PMID: 25819445

Zardoya R, Doadrio I (1999) Molecular evidence on the evolutionary and biogeographical patterns of european Cyprinids. J Mol Evol **49:** 227–237. PMID:10441674

Supporting information

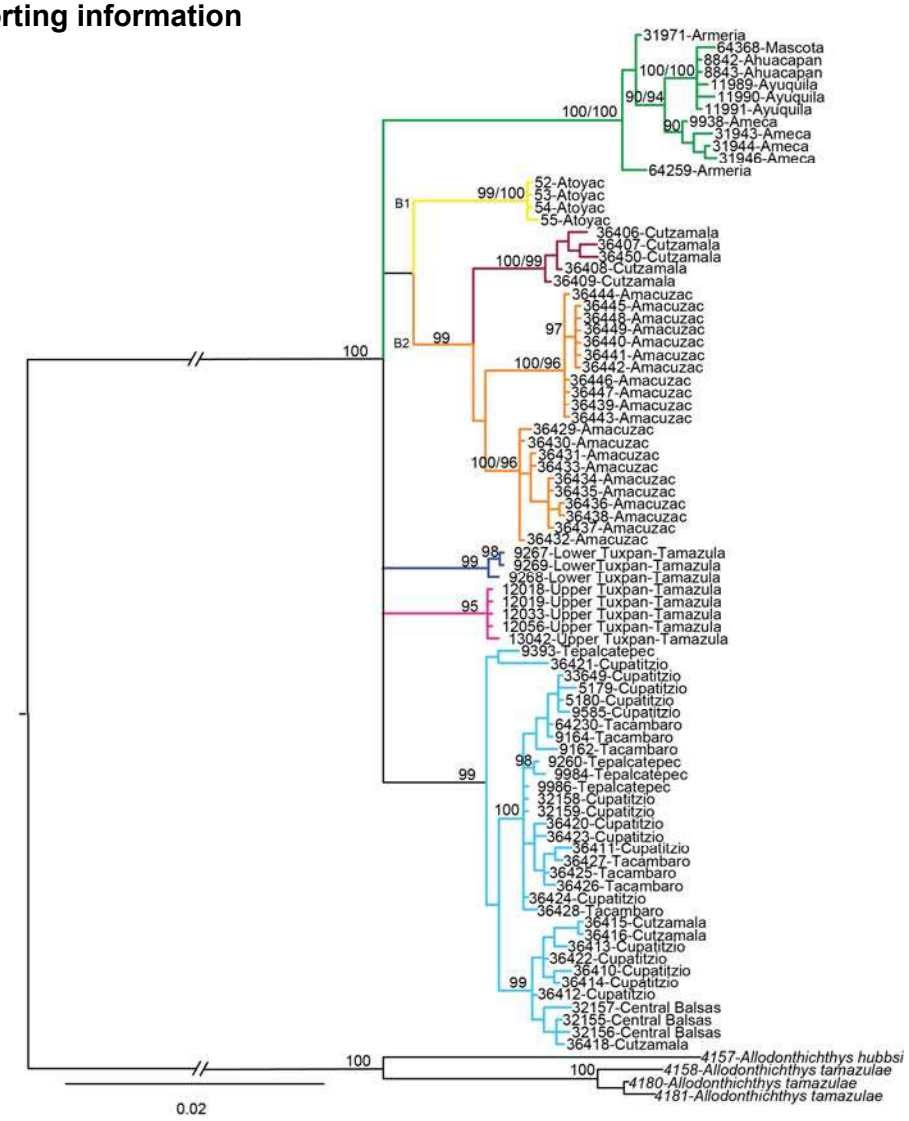


Figure S1. The Bayesian inference tree of *Ilyodon* species inferred from concatenated sequences of three mitochondrial genes (*cytb*, *coxI*, and *d-loop*; 1600 bp) concatenated. Bayesian posterior probability (>90%) and maximum likelihood bootstrap values (>80%) are indicated.

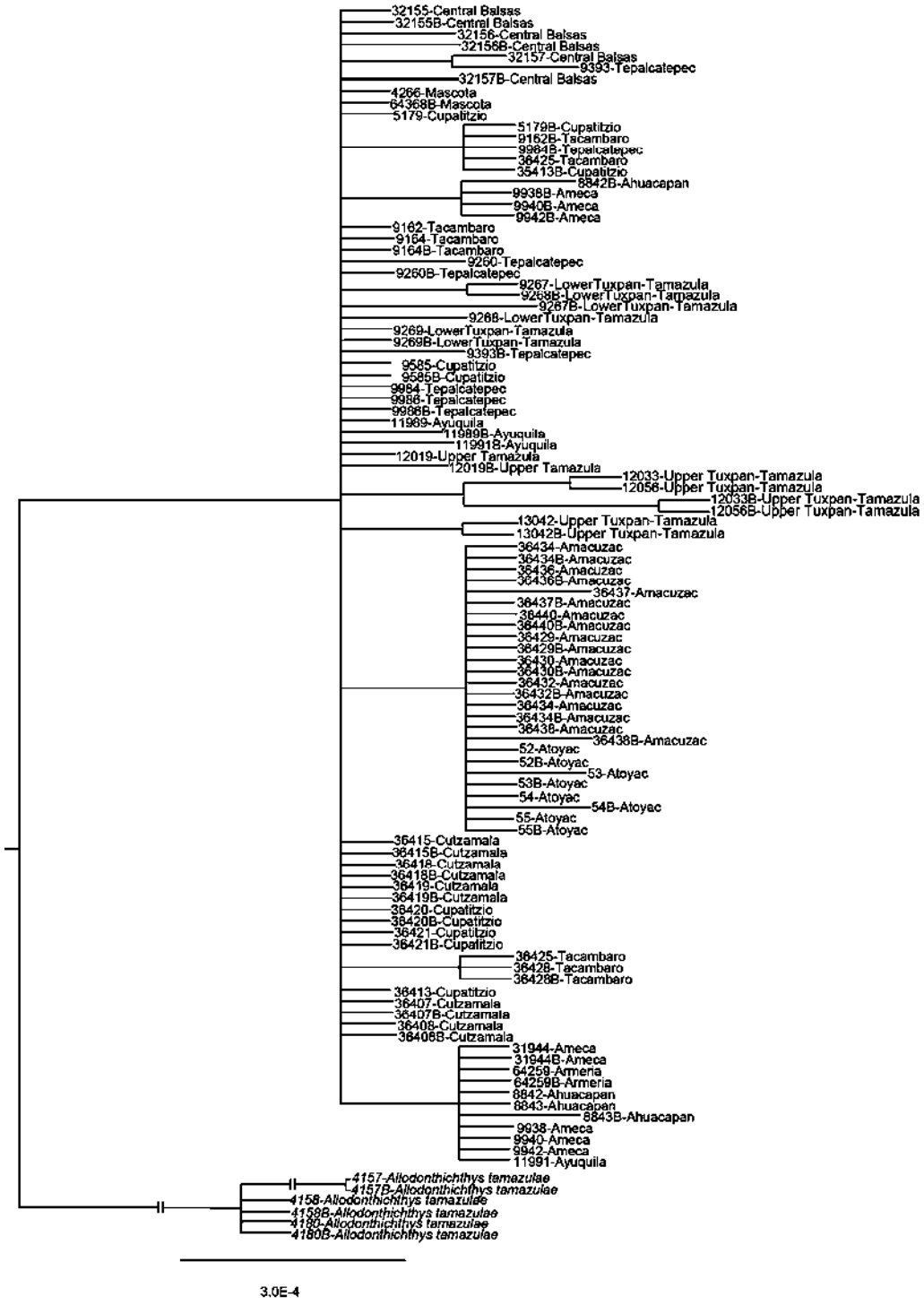


Figure S2. The Bayesian inference tree of *Ilyodon* species inferred from concatenated sequences of two nuclear genes (*ACTB* and *RAG1*; 2432 bp) concatenated. Bayesian posterior probability (>90%) are indicated.

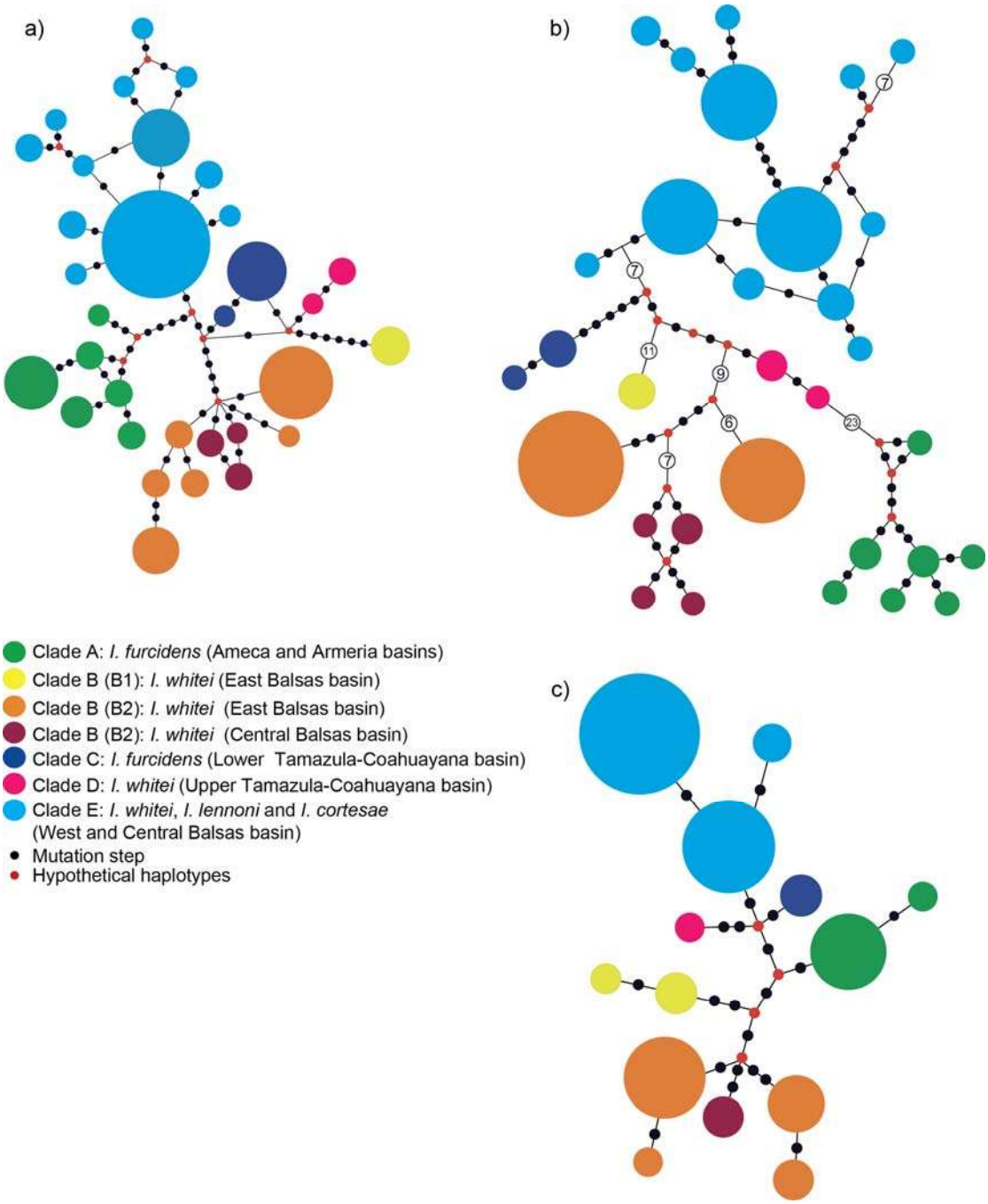


Figure S3. Haplotype networks of mitochondrial genes; a) *cytb*; b) *d-loop*; c) *coxI*; colours represent the hydrological basin; the area of the circles is proportional to haplotype frequency.

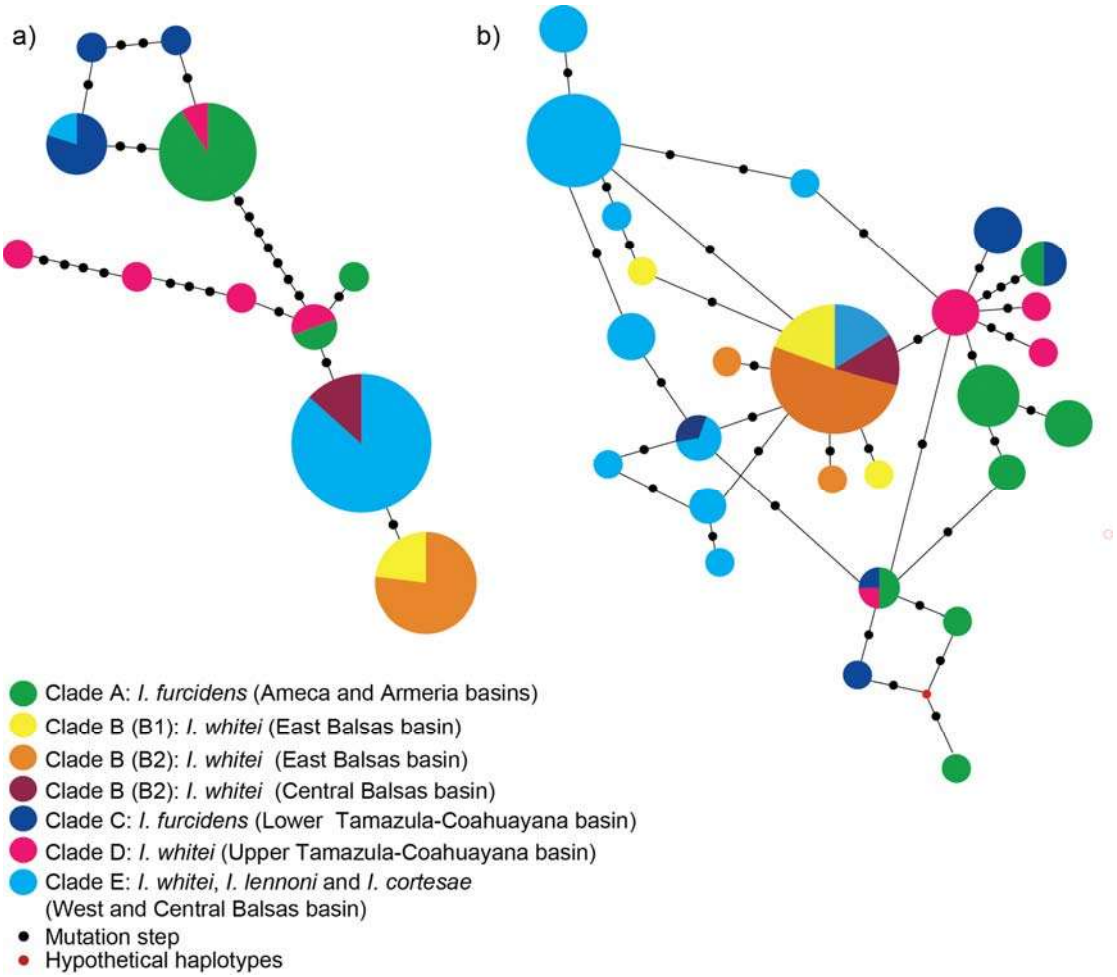


Figure S4. Haplotype networks of nuclear genes a) *ACTB*; b) *RAG1*; colours represent the hydrological basin; the area of the circles is proportional to haplotype frequency. In the case of haplotypes shared among hydrological basins, shading is proportional to the frequency of the haplotype in each region.

Table S1. Primers, PCR conditions, and References.

	Cytb	*cox1*	*d-loop*	*ACTB*	*RAG1*
Primers	Glu-F Thr-R	Fish-F1 Fish-R1	Dloop-A Dloop-E	β-actin-For β-actin-Rev	Rag-1F Rag-9R
Size (bp)	533	626	441	979	1453
Reference	Doadrio and Domínguez, (2004).	Ward *et al*. (2005).	Lee *et al*. (1995).	Robalo *et al*. (2006)	Quenouille, Bermingham & Planes (2004).
Denaturing (step 1)	94ºC, 2 min.	94ºC, 2 min.	95ºC, 1 min.	94ºC, 3 min.	94ºC, 3 min.
Cycles (step 2)	35	35	32	32	32
Denaturing	94ºC, 45 s.	94ºC, 30 s.	94ºC, 30 s.	94ºC, 45 s.	94ºC, 45 s.
Annealing	49.4ºC, 1 min.	52ºC, 30 s.	49.5ºC, 30 s.	55ºC, 45 s.	52ºC to 54.5ºC, 45 s.
Extension	72ºC, 90 s.	72ºC, 1 min.	72ºC, 1 min.	72ºC, 90 s.	72ºC, 90 s.
Final Extension (step 3)	72ºC, 5 min.	72ºC, 10 min.	72ºC, 7 min.	72ºC, 7 min.	72ºC, 7 min.

Table S2. Ambiguously Aligned Regions for *d-loop* above and *ACTB* below.

d-loop	**PB**																									
Clade	53	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	308
A	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	A	T	G	T	A	A	-
B1	A	A	G	C	A	C	A	T	A	A	G	C	A	C	A	T	A	A	G	C	A	C	A	T	A	-
B2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
C	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
D	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A
E	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

ACTB	**PB**																
Clade	86	87	88	89	90	91	93	94	95	96	97	98	99	100	101	113	114
A	G	C	A	C	A	T	-	-	-	-	-	-	-	-	-	T	T
B1	G	C	A	C	A	T	A	G	A	A	G	A	C	T	T	T	T
B2	G	C	A	C	A	T	A	G	A	A	G	A	C	T	T	T	T
C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
D	G	C	A	C	A	T	A	G	A	A	G	A	C	T	T	T	T
E	G	C	A	C	A	T	A	G	A	A	G	A	C	T	T	T	T

Table S3. Access numbers for GenBank of each sequence for all genes.

Tissue	GenBank access for *cytb*	GenBank access for *coxI*	GenBank access for *d-loop*	GenBank access for *RAG1*	GenBank access for *ACTB*
32155	KY204452	KY118827	KY204628	KY204542	KY204718
32156	KY204453	KY118828	KY204629	KY204543	KY204719
32156B				KY204544	
32157	KY204454	KY118829	KY204630	KY204545	KY204720
32157B				KY204546	
33649	KY204455	KY118830	KY204631		
64230		KY118831			
4102		KY118832			
4103	KY204456	KY118833	KY204632		KY204721
64259	KY204457	KY118834			
4133	KY204458	KY118835	KY204633	KY204547	KY204722
4266			KY204711		
64367		KY118836			
64368	KY204459	KY118837	KY204634	KY204548	KY204723
5179	KY204460		KY204635	KY204549	KY204724
5179B				KY204550	
5180	KY204461	KY118838	KY204636		
8842	KY204462	KY118839	KY204637	KY204551	KY204725
8842B				KY204552	
8843	KY204463	KY118840	KY204638	KY204553	KY204726
8843B				KY204554	KY204727
9162	KY204464	KY118841	KY204639	KY204555	KY204728
9162B				KY204556	
9164	KY204465	KY118842	KY204640	KY204557	KY204729

9166	KY204466	KY118843			
9260	KY204467	KY118844	KY204641	KY204558	KY204730
9260B				KY204559	
9267	KY204468	KY118845	KY204642	KY204560	KY204731
9267B				KY204561	KY204732
9268	KY204469	KY118846	KY204643	KY204562	KY204733
9268B				KY204563	
9269	KY204470	KY118847	KY204644	KY204564	KY204734
9269B				KY204565	
9393	KY204471	KY118848	KY204645	KY204566	KY204735
9393B				KY204567	KY204736
9584		KY118849	KY204646		
9585	KY204472		KY204647	KY204568	KY204737
9585B				KY204569	
9938	KY204473	KY118850	KY204648	KY204570	KY204738
9938B				KY204571	
9940		KY118851	KY204649	KY204572	KY204739
9940B				KY204573	
9942		KY118852	KY204650	KY204574	KY204740
9942B				KY204575	
9946			KY204651		
9984	KY204474	KY118853	KY204652	KY204576	KY204741
9984B				KY204577	
9986	KY204475	KY118854	KY204653	KY204578	KY204742
9986B				KY204579	
11989	KY204476	KY118855	KY204654	KY204580	KY204743
11989B				KY204581	
11990	KY204477	KY118856	KY204655		KY204744
11990B					KY204745
11991	KY204478	KY118857	KY204656	KY204582	KY204746

11991B				KY204583	
12018	KY204479				
12019	KY204480		KY204657	KY204584	KY204747
12019B				KY204585	
12033	KY204481	KY118858	KY204658	KY204586	KY204748
12033B				KY204587	KY204749
12056	KY204482	KY118859	KY204659	KY204588	
12056B				KY204589	
13042	KY204483	KY118860	KY204660	KY204590	KY204750
31943	KY204484	KY118861		KY204541	KY204717
31944	KY204485	KY118862	KY204661		
31945			KY204662		
31946	KY204486	KY118863	KY204663		
32158	KY204487	KY118864	KY204665		
32159	KY204488	KY118865	KY204664		
33008	KY204489	KY118866			
51		KY118900			
52	KY204490	KY118886	KY204666	KY204602	KY204772
53	KY204491	KY118887	KY204667	KY204603	KY204773
53B				KY204604	
54	KY204492	KY118888		KY204605	
54B				KY204606	
55	KY204493	KY118889	KY204710	KY204607	KY204774
36415	KY204494	KY118906	KY204677	KY204608	
36415B				KY204609	
36416	KY204495	KY118907			
36417			KY204678		
36418	KY204496	KY118908	KY204679	KY204610	KY204751
36418B				KY204611	
36419			KY204680	KY204612	KY204752

36419B				KY204613	
36420	KY204497	KY118877	KY204668	KY204614	KY204753
36421	KY204498	KY118878	KY204670	KY204615	KY204754
36421B				KY204616	
36422	KY204499	KY118879	KY204669		
36423	KY204500	KY118880	KY204671		
36424	KY204501	KY118881	KY204672		
36428				KY204619	
36410	KY204502	KY118909	KY204681		
36411	KY204503	KY118910	KY204682		
36412	KY204504	KY118911	KY204683		
36413	KY204505	KY118912	KY204684	KY204620	KY204755
36413B				KY204621	
36414	KY204506	KY118913	KY204685		
36425	KY204507	KY118890	KY204673	KY204617	KY204756
36425B				KY204618	
36426	KY204508	KY118891			
36427	KY204509	KY118892	KY204674		KY204757
36428	KY204510	KY118893	KY204675		KY204758
36406	KY204511	KY118901	KY204686		
36407	KY204512	KY118902	KY204687	KY204622	KY204759
36408	KY204513	KY118903	KY204688	KY204623	KY204760
36409	KY204514	KY118904	KY204689		KY204761
36450	KY204515	KY118905	KY204690	KY204591	KY204762
36444	KY204516	KY118895	KY204701		
36445	KY204517	KY118896	KY204702		
36446	KY204518	KY118897	KY204703	KY204592	KY204763
36447	KY204519	KY118898	KY204704	KY204593	KY204764
36447B				KY204594	
36448	KY204520	KY118899			

36449	KY204521		KY204705		
36439	KY204522	KY118882	KY204706		
36440	KY204523	KY118883	KY204707	KY204595	
36441	KY204524		KY204708		
36442	KY204525	KY118884		KY204767	KY204767
36443	KY204526	KY118885	KY204709	KY204768	KY204768
36429	KY204527	KY118867	KY204691	KY204596	KY204769
36430	KY204528	KY118868	KY204692	KY204598	KY204771
36431	KY204529	KY118869	KY204693		
36432	KY204530	KY118870	KY204694	KY204597	KY204770
36433	KY204531	KY118871	KY204695		
36434	KY204532	KY118872	KY204696	KY204599	KY204765
36435	KY204533	KY118873	KY204697		
36436	KY204534	KY118874	KY204698		
36437	KY204535	KY118875	KY204699		
36438	KY204536	KY118876	KY204700	KY204600	KY204766
36438B				KY204601	
36451		KY118894	KY204676		
1270	KY204537		KY204716	KY204624	KY204776
4157	KY204538	KY118914	KY204715	KY204626	KY204775
4158			KY204713	KY204625	KY204777
4180	KY204539		KY204714	KY204627	KY204778
4181	KY204540		KY204712		

Table S4. Models selected with Akaike information criterion and the parameters of each gene.

	Cytb	*cox1*	*d-loop*	*ACTB*	*RAG1*
Model	TrN+I+G	TrN+G	K81ug+G	HKY	TrN+I
Frec. A	0.2557	0.2357	N/A	0.2395	0.2619
Frec. C	0.3057	0.2686	N/A	0.2537	0.2416
Frec. G	0.1279	0.1668	N/A	0.2212	0.2763
Frec. T	0.3107	0.3289	N/A	0.2856	0.2203
P-inv	0.4760	N/A	N/A	N/A	0.1360
Gamma shape	0.9130	0.0660	0.0880	N/A	N/A

Table S5. Uncorrected genetic distances within recovered clades and sub-clades based on *cytb* (to the left of the diagonal) and *cox1* (to the right of the diagonal) and between recovered groups in phylogenetic analyses based on *cytb* (below the diagonal) and *cox1* (above the diagonal) genes.

cytb/*cox1*	Clade A	Clade B		Clade C	Clade D	Clade E
		Sub-clade B1	Sub-clade B2			
Clade A	0.005/0.002	0.008	0.009	0.007	0.006	0.007
Sub-clade B1	0.021	0.000/0.001	0.008	0.010	0.008	0.010
Sub-clade B2	0.019	0.016	0.005/0.005	0.011	0.010	0.011
Clade C	0.020	0.014	0.016	0.001/0.000	0.005	0.007
Clade D	0.019	0.013	0.015	0.005	0.000/0.000	0.005
Clade E	0.017	0.015	0.016	0.010	0.009	0.003/0.002

Table S6. Uncorrected genetic distances within and between recovered groups in phylogenetic analyses based on the *d-loop* gene.

d-loop	Clade A	Clade B		Clade C	Clade D	Clade E
		Sub-clade B1	Sub-clade B2			
Clade A	0.007					
Sub-clade B1	0.102	0.000				
Sub-clade B2	0.079	0.043	0.019			
Clade C	0.067	0.035	0.040	0.002		
Clade D	0.080	0.037	0.041	0.023	0.001	
Clade E	0.081	0.043	0.047	0.037	0.036	0.014

Table S7. Uncorrected genetic distances within recovered clades and sub-clades based on *ACTB* (to the left of the diagonal) and *RAG1* (to the right of the diagonal), and between recovered groups in phylogenetic analyses based on *ACTB* (below the diagonal) and *RAG1* (above the diagonal).

ACTB/*RAG1*	Clade A	Clade B		Clade C	Clade D	Clade E
		Sub-clade B1	Sub-clade B2			
Clade A	0.000/0.000	0.001	0.001	0.001	0.001	0.002
Sub-clade B1	0.001	0.000/0.000	0.000	0.001	0.001	0.001
Sub-clade B2	0.001	0.000	0.000/0.000	0.001	0.001	0.001
Clade C	0.000	0.001	0.001	0.001/0.000	0.000	0.001
Clade D	0.000	0.001	0.001	0.001	0.001/0.000	0.001
Clade E	0.000	0.001	0.001	0.000	0.000	0.000/0.000

CAPÍTULO II

Genetic differentiation among populations of the freshwater fish *Goodea atripinnis* Jordan, 1880 (Cyprinodontiformes: Goodeidae), implications for its evolutionary history and its taxonomy

Genetic differentiation among populations of the freshwater fish *Goodea atripinnis* Jordan, 1880 (Cyprinodontiformes: Goodeidae), implications for its evolutionary history and taxonomy

[1,2]Beltrán-López, Rosa Gabriela., [2,3*]Domínguez-Domínguez, Omar., [4]Piller, Kyle, R., [2,5]Mar-Silva, Adán Fernando & [6]Doadrio, Ignacio

[1]Programa Institucional de Doctorado en Ciencias Biológicas, Facultad de Biología. Universidad Michoacana de San Nicolás de Hidalgo. Michoacán. México.
[2]Laboratorio de Biología Acuática, Facultad de Biología. Universidad Michoacana de San Nicolás de Hidalgo. Michoacán. México.
[3]Laboratorio Nacional de Análisis y Síntesis Ecológica para la Conservación de Recursos Genéticos de México, Escuela Nacional de Estudios Superiores, Unidad Morelia, Universidad Nacional Autónoma de México, Michoacán, México.
*Corresponding author
[4]Department of Biological Sciences, Southeastern Louisiana University, Hammond, LA, 70402, USA. E-mail: kyle.piller@selu.edu
[5]Programa Institucional de Maestría en Ciencias Biológicas, Universidad Michoacana de San Nicolás de Hidalgo. Michoacán, México.
[6]Departamento de Biodiversidad y Biología Evolutiva. Museo Nacional de Ciencias Naturales. CSIC. Madrid. España.

ABSTRACT

Central México is characterized by a complex topography, due to tectonic and climatic history; these events have influenced the evolutionary history of numerous freshwater fishes that inhabit the area. However, recent studies have been demonstrate than when exist conection within and among water drainages, life history traits and ecological characteristics of the species may influence dispersal capabilities, which, can affect the genetic connectivity. *Goodea* is one of the most widely distributed and environmentally tolerant genera of goodeidos. We examined with *cytb* the phylogeographical relationships, genetic structure, genetic diversity, and demographic history of populations of *Goodea atripinnis* from across its range of distribution. Low genetic differentiation was found (genetic distances of 0.5 to 0.7%), and shared haplotypes between several regions, although, geographic segregation was found in samples west and east of the Lower Lerma region. Also, the AMOVA that explained better the genetic structure was when were considered two haplogroups plus Armeria region (42.18%). Several regions showed null genetic diversity suggested a dispersal mediated by humans. Finally, the BSP showed a population expansion for two haplogroups, showed the contrary to other fishes of central México, suggested to *G. atripinnis* the climate change do not have been affected its effective size populations. This phylogeographic pattern suggested a recent colonization toward the most of the biogeographic regions, due to its high dispersal capacity, and to be a very abundant species with high reproductive success, tolerant to environmental changes, monomorphic species and the generalist feeding to differences with other freshwater fishes of central México.

Keywords: Endemic fish, widest distribution, phylogeography, haplogroups

INTRODUCTION

The distribution of genetic variation in freshwater organisms around the world is mainly affected by geologic evolution of the basins where they occur and they frequently show strong genetic structure as a result of their confinement to given hydrological systems (Faulks et al., 2010; Loxterman & Keeley, 2012). For widely distributed freshwater fishes, genetic differentiation patterns are more frequently linked with historical geomorphological processes of the drainages, more than the present-day configuration (Beltrán-López et al., 2018; Bermingham & Martin, 1998; Domínguez-Domínguez et al., 2008a; Hewitt 2000; 2004; Perea et al., 2016; Pérez-Rodríguez et al., 2009).

Under this context, central México is characterized by a complex topography, due to tectonic and climatic history, been one of the most important geological features of central México the Trans Mexican Volcanic Belt (TMVB). The uplift of the TMVB since the Neogene have promoted a long history of genesis and destruction of water drainages, which has had a significant impact on the diversification of taxa in the region; due to the formation of new geographical barriers and montane habitats (Anducho-Reyes et al., 2008; Ferrusquía-Villafranca, 1993). These events have influenced the evolutionary history of numerous freshwater organisms that inhabit the area (Beltrán-López et al., 2018; Domínguez-domínguez et al., 2008a; 2010; Huidobro et al., 2006; Mateos et al., 2002; Pedraza-Lara et al., 2012; Pérez-Rodríguez et al., 2009; 2016).

But in the other hand, recent studies have been demonstrate than when exist connection within and among water drainages, life history traits and ecological characteristics of the species may influence dispersal capabilities, which, in turn, can affect the genetic connectivity (Betancourt-Resendes et al., 2018; Goto & Andoh, 1990). In this sense, dispersal and colonization capability play a very important role in occupancy of new habitats and in the geographic distribution of freshwater fishes (Clobert et al., 2001). For its high geological dynamism, México is considered a model region for studying how different ecological and biological traits affect the genetic structure or connectivity of the species (Mastretta-Yanes et al., 2015).

The goodeines (Cyprinodontiformes: Goodeidae) are one of the most prevalent groups of freshwater fishes in central México. Several species within the family are widespread, whereas most others are restricted to a few river systems or even a single spring (Miller et al., 2005). Evolutionary studies of several species have found high genetic structure and divergence among species, genera, and populations, as was the case of *Zoogoneticus quitzeonesis* (Domínguez-Domínguez et al., 2008a), *Xenotoca eiseni* (Piller et al., 2015) and *Ilyodon* species (Beltrán-López et al., 2017). Although not common, null or low genetic differentiation between freshwater fishes species, in the region, including some goodeines (Beltrán-López et al., 2018; Betancourt-Resendes et al., 2018; Corona-Santiago et al., 2015; Ornelas-García et al., 2012), in which low genetic differences, shared haplotypes, and recent isolation events were found, which could be explained by human mediated dispersion, high dispersal capability of the species, or even taxonomic unsertainity of some of the recognized species.

Within the Goodeidae, *Goodea* is one of the most widely distributed (Miller et al., 2005) and environmentally tolerant genera (De la Vega-Salazar, 2006). Morphological differences between populations are found (Miranda et al., 2010) and, at least three species have been described: *Goodea atripinnis* (Jordan, 1880), from Lerma basin, which drains to the Pacific slope, *Goodea luitpoldii* (Steindachner, 1894), from Lake Patzcuaro, and *Goodea gracilis* (Hubbs and Turner, 1939), from the Panuco River basin on the Atlantic slope. Despite of recognition of three described species, recent molecular studies using a reduced number of samples have concluded that *Goodea* is composed of only one species, *Goodea atripinnis* (Domínguez-Domínguez et al., 2010; Webb et al., 2004).

Morphological differences among populations of aquatic organisms have been postulated to be the result of divergent selection pressures of water flow, variation in dissolved oxygen, or prey type/abundance variation between lotic and lentic habitats (Collin & Fumagalli, 2011; Crispo & Chapman, 2010). Therefore, despite minimal genetic divergences between populations, a high phenotypic plasticity in *G. atripinnis* have been found, mainly related to a rapid response to a changing environment (Crispo, 2008, Foster et al., 2015; Robinson & Wilson, 1994).

The objective of this study is to examine the link between the evolutionary history of *Goodea atripinnis* and the geologic and climatic history of central México. In particular, we examined the phylogeographic relationships, genetic structure, genetic diversity, and demographic history of populations of *Goodea atripinnis* from across its range and compare with the patterns found in previous studies in goodeines and in other co-distributed species of fishes (Beltrán-López et al., 2018; Domínguez-Domínguez et al., 2008a; Pérez-Rodríguez et al., 2009; Schönhuth and Doadrio, 2003). To accomplish this, we extensively sampled throughout the distributional range of *Goodea* and included many individuals to alleviate the small sample size issues of previous studies that have included this species (Doadrio and Domínguez 2004; Domínguez-Domínguez et al., 2010; Webb et al., 2004). It is expected that inclusion of a larger number of samples from across the range will better clarify the genetic structure within this widespread species and allow us to determine if there is significant genetic structure between biogeographical regions, as has been recovered for other fishes in central México (Beltrán-López et al., 2018; Betancourt-Resendes et al., 2018; Domínguez-Domínguez et al., 2008a; 2010; Pérez-Rodríguez et al., 2009; 2015; 2016).

MATERIALS AND METHODS

Fish sampling and DNA isolation

Three hundred eighteen specimens were collected from 72 localities in 23 biogeographical regions proposed to central México by Domínguez-Domínguez et al., (2006). The samples include rivers and lakes and cover the entire distributional range of *Goodea atripinnis* (Figure 1; Table 1). Fish were captured with permission of the authorities by electrofishing and seine nets and anesthetized with tricaine-mesylate (MS-222). Pectoral fin clips were obtained and preserved in absolute ethanol and frozen at -75°C, and deposited in the tissue bank of the Aquatic Biology laboratory of the Universidad Michoacana de San Nicolas de Hidalgo.

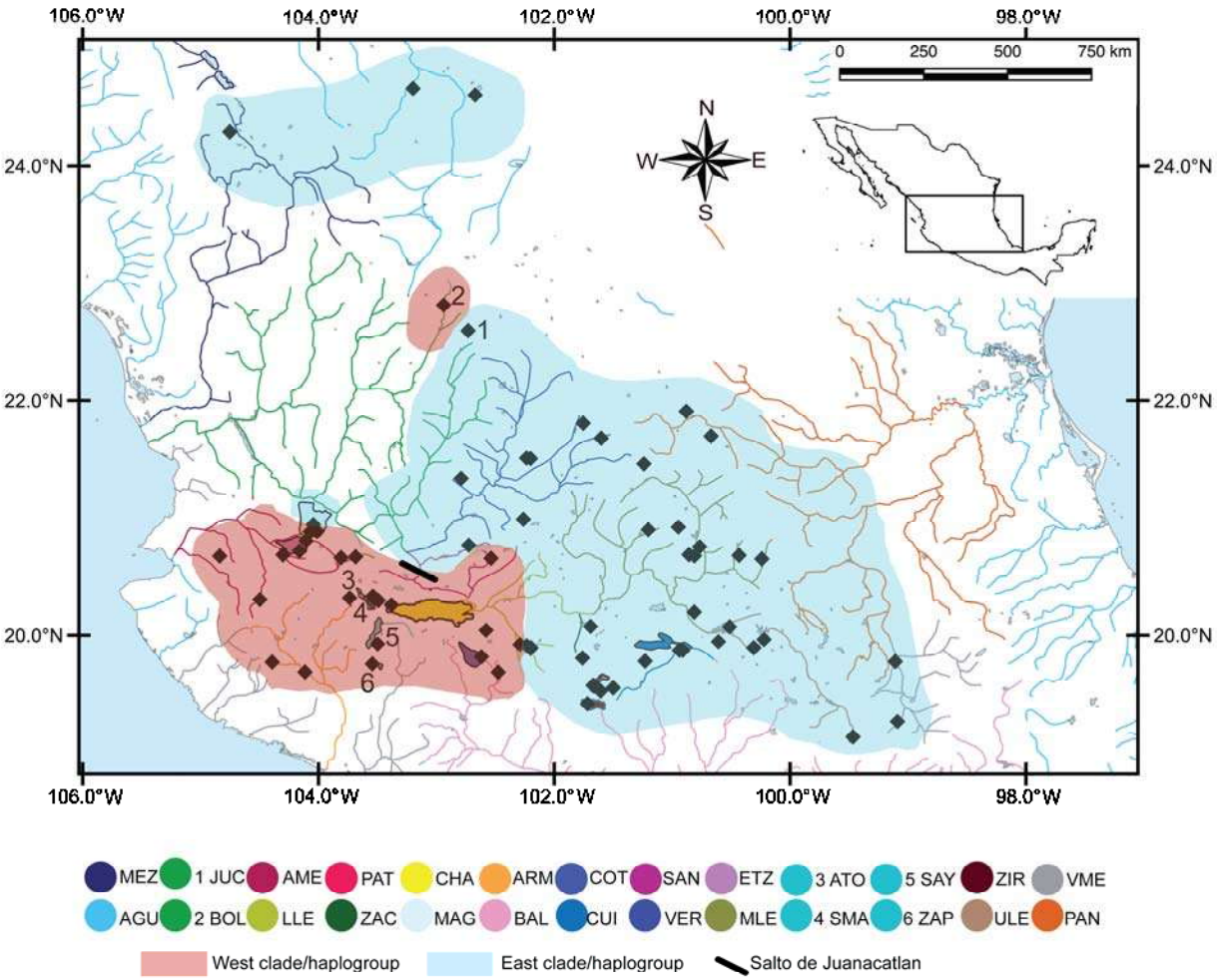


Figure 1 Distribution map of *Goodea atripinnis* samples, diamonds shown the localities sampled for the present work, the shaded part show the geographic distribution of the two recovered clades/haplogroups obtained in phylogenetic analyses. The abbreviations of the biogeographical regions are: MAG: Lake Magdalena, LLE: Lower Lerma river, AME: Ameca river, VME: Valley of México, PAN: Panuco river, AGU: Aguanaval river, CUI: Lake Cuitzeo, VER: Verde river, BOL: Bolaños river, MLE: Middle Lerma river, CHA: Lake Chapala, ULE: Upper Lerma river, PAT: Lake Patzcuaro, ZAC: Lake Zacapu, ETZ: Etzatlan-San Marcos, JUC: Juchipila river, ARM: Armeria river, COT: Lake Cotija, ZIR: Lake Zirahuen, MEZ: Mezquital river, SAN: Santiago river, BAL: Balsas river, ATO (3): Lake Atotonilco, SMA (4): Lake San Marcos, SAY (5): Lake Sayula and ZAP (6): Lake Zapotlan, the last four numbers corresponded with the Sayula region.

Table 1 Samples localities and sequences information.

Site	Locality	Basin	Biogeographic region	Sequences number	GPS Coordinates
1	Presa la Quemada	Lake Magdalena	Magdalena	5	20° 57´ 51.05´´ N, 104° 3´8.34´´ W
2	Los Venados	Lake Magdalena	Magdalena	7	20° 54´ 14.2´´ N, 104° 1´ 11.6´´ W
3	Lago de Magdalena	Lake Magdalena	Magdalena	9	20° 54´ 3.69´´ N, 104° 4´ 45.62´´ W
4	El Moloya	Lake Magdalena	Magdalena	11	20° 57´ 51.05´´ N, 104° 3´ 8.34´´ W
5	Presa La Luz	Duero river	Lower Lerma	2	19° 56´ 12.4´´ N, 102° 17´ 55.3´´ W
6	Las adjuntas	Duero river	Lower Lerma	8	19° 54´ 39.9´´ N, 102° 12´ 20´´ W
7	El Platanal	Duero river	Lower Lerma	6	19° 55´ 40.7´´ N, 102° 14´ 30.3´´ W
8	Pateo		Lower Lerma	1	19° 54´ 33.1´´ N, 100° 19´ 5.1´´ W
9	Camecuaro	Lake Camecuaro	Lower Lerma	3	19° 54´ 9.8´´ N, 102° 12´ 35.27´´ W
10	Amatlán de cañas	Ameca	Ameca	2	20° 42´ 13.7´´ N, 104° 18´ 34.4´´ W
11	El Rincón	Ameca	Ameca	1	20° 41´ 34.8´´ N, 104° 50´ 47.1´´ W
12	Tala, Río Salado	Ameca	Ameca	5	20° 41´ 12.1´´ N, 103° 41´ 36.3´´ W
13	Teuchitlán	Cocula-La Vega	Ameca	5	20° 40´ 46.8´´ N, 103° 50´ 59.2´´ W
14	Xochimilco	Canal de Xochimilco	Valle de México	2	19° 17´ 3´´ N, 99° 6´ 7.62´´ W
15	Presa del Carmen		Panuco	6	20° 40´ 8.1´´ N, 100° 15´ 6.73´´ W
16	Zumpango		Panuco	4	19° 47´ 31.6´´ N, 99° 7´ 1´´ W
17	San Ildefonso		Panuco	1	19° 57´ 43.2´´ N, 100° 37´ 1.27´´ W
18	Tierra Quemada		Panuco	1	21° 42´ 53.9´´ N, 100° 41´ 15.9´´ W
19	Jesús María		Panuco	5	21° 55´ 24.3´´ N, 100° 54´ 39.9´´ W
20	Jesús María-Villa de Reyes		Panuco	5	21° 55´ 32.3´´ N, 100° 54´ 39.5´´ W
21	Santa Clara		Aguanaval	12	24° 40´ 47.5´´ N, 103° 13´ 36.2´´ W
22	Presa el Tecolote	Verde	Verde	1	21° 31´ 25.1´´ N, 102° 13´ 25.2´´ W
23	Arroyo el Tecolote	Verde	Verde	3	21° 31´ 38.4´´ N, 102° 15´ 22.1´´ W
24	Presa La Paz	Verde	Verde	1	21° 49´ 21.0´´ N, 101° 46´ 13.7´´ W
25	Guadalupe Victoria	Verde	Verde	2	21° 41´ 49.3´´ N, 101° 37´ 21.3´´ W
26	San Julián-San Miguel	Verde	Verde	2	21° 0´ 32.7´´ N, 102° 17´ 47.3´´ W
27	Bordo en Chimaliquin	Verde	Verde	1	21° 21´ 24´´ N, 102° 48´ 6.8´´ W
28	Presa El Tesorero	Bolaños	Bolaños	3	22° 49´ 53.6´´ N, 102° 57´ 12.7´´ W
29	Urideo, ojo de agua el Capulin		Middle Lerma	2	20° 12´ 52.2´´ N, 100° 50´ 43.4´´ W
30	Arroyo Neutla		Middle Lerma	2	20° 41´ 41.5´´ N, 100° 50´ 48.8´´ W
31	Carretera San Miguel-Comonfort		Middle Lerma	2	20° 46´ 17.8´´ N, 100° 47´ 30.0´´ W
32	Afluente Neutla		Middle Lerma	1	20° 42´ 16.5´´ N, 100° 52´ 1.6´´ W
33	Río San José del Rodeo		Middle Lerma	7	20° 55´ 12.2´´ N, 101° 13´ 19.9´´ W

34	Río Xoconostle-San Juan	Laja	Middle Lerma	1	20° 56´ 31.5´´ N, 100° 58´ 38´´ W
35	Manantial Andrés-Figueroa	San Marcos	Sayula	5	20° 20´ 0.4´´ N, 103° 34´ 97.6´´ W
36	Depósito Santa Catarina	San Marcos	Sayula	1	20° 21´ 3.6´´ N, 103° 33´ 11.6´´ W
37	Canal Presa Buena Vista	Atotonilco	Sayula	7	20° 20´ 5.8´´ N, 103° 45´ 19.7´´ W
38	Lago Atotonilco	Atotonilco	Sayula	4	20° 18´ 45.6´´ N, 103° 31´ 59.3´´ W
39	Presa Buena Vista	Atotonilco	Sayula	4	20° 20´ 5.2´´ N, 103° 45´ 20.2´´ W
40	Lago de Zapotlán	Zapotlán	Sayula	9	19° 44´ 44.9´´ N, 103° 28´ 22.4´´ W
41	Lago de Sayula	Sayula	Sayula	10	19° 56´ 32´´ N, 103° 31´ 36.7´´ W
42	Lago de Chapala	Chapala	Chapala	2	20° 16´ 30.4´´ N, 103° 24´ 37.1´´ W
43	Los Negritos	Chapala	Chapala	2	20° 3´ 36.3´´ N, 102° 36´ 46.1´´ W
44	Chiquimitio	Cuitzeo	Cuitzeo	4	19° 47´ 44.6´´ N, 101° 15´ 41.1´´ W
45	Río Querendaro	Cuitzeo	Cuitzeo	7	19° 53´ 13.9´´ N, 100° 56´ 52.2´´ W
46	Desembocadura Río Querendaro	Cuitzeo	Cuitzeo	2	19° 53´ 14.5´´ N, 100° 57´ 7.1´´ W
47	Manantial Chapultepec	Pátzcuaro	Pátzcuaro	8	19° 34´ 19.1´´ N, 101° 31´ 29.3´´ W
48	Embarcadero Principal	Pátzcuaro	Pátzcuaro	3	19° 32´ 42.9´´ N, 101° 37´ 4.2´´ W
49	Erongacícuaro	Pátzcuaro	Pátzcuaro	2	19° 35´ 12.6´´ N, 101° 41´ 50.21´ W
50	Napizaro	Pátzcuaro	Pátzcuaro	3	19° 35´ 30´´ N, 101° 41´ 16.5´´ W
51	Presa Melchor Ocampo	Angulo-Lerma	Zacapu	12	20° 5´ 36.5´´ N, 101° 43´ 57.4´´ W
52	La Zarcita	Angulo-Lerma	Zacapu	2	19° 49´ 19´´ N, 101° 47´ 51´´ W
53	Tocumbo		Balsas	7	19° 42´ 7´´ N, 102° 30´ 60´´ W
54	Presa San Juanico	Cotija	Cotija	3	19° 49´ 57.4´´ N, 102° 38´ 25.8´´ W
55	San Sebastián	Etzatlan-San Marcos	Etzatlan-San Marcos	9	20° 49´ 25´´ N, 104° 7´ 10.8´´ W
56	Presa San Rafael	Etzatlan-San Marcos	Etzatlan-San Marcos	4	20° 44´ 8.3´´ N, 104° 11´ 49.7´´ W
57	Puente Malpaso	Juchipila	Juchipila	11	22° 36´ 57.2´´ N, 102° 45´ 39.5´´ W
58	Tepatitlán	Santiago-Chapala	Grande de Santiago	4	20° 47´ 6.8´´ N, 102° 45´ 58.7´´ W
59	San Antonio, Tepatitlán	Santiago-Chapala	Grande de Santiago	7	20° 40´ 27.2´´ N, 102° 33´ 19.4´´ W
60	Presa Garabato, Totolán	Santiago-Chapala	Grande de Santiago	12	24° 37´ 28.4´´ N, 102° 41´ 15.6´´ W
61	Joya Grande		Upper Lerma	2	20° 5´ 25.3´´ N, 100° 32´ 40.8´´ W
62	Manantial del seminario		Upper Lerma	1	21° 28´ 42.2´´ N, 101° 15´ 6.6´´ W
63	Presa Juriquilla		Upper Lerma	3	20° 41´ 56.3´´ N, 100° 27´ 31.7´´ W
64	Laguna de Almoloya		Upper Lerma	12	19° 9´ 8.4´´ N, 99° 29´ 30.6´´ W
65	Pateo-Contepec		Upper Lerma	1	19° 54´ 33.1´´ N, 100° 19´ 5.1´´ W
66	Tepuxtepec		Upper Lerma	6	19° 58´ 50.1´´ N, 100° 14´ 12.6´´ W
67	Achacales		Armeria	3	19° 42´ 14.1´´ N, 104° 8´ 37.9´´ W
68	El grullo	Ayuquila	Armeria	1	19° 42´ 16.6´´ N, 104° 31´ 41.1´´ W
69	Atenguillo		Armeria	5	20° 19´ 16.6´´ N, 104° 31´ 41.1´´ W
70	Opopeo		Zirahuen	3	19° 26´ 12.2´´ N, 101° 44´ 23.9´´ W
71	Río San Pedro		Mezquital	8	24° 18´ 24´´ N, 104° 46´ 20.38´´ W

Some specimens were deposited in the fish collection of the Universidad Michoacana de San Nicolás de Hidalgo (registration by the Ministry of Environment and Natural Resources for México SEMARNAT: MICH-PEC-227-07-09), the remaining specimens were returned to the water. Tissues for DNA extraction were digested with ATL QIAGEN Buffer and Proteinasa K and purified with BioSprint DNA Blood Kit QIAGEN according to the manufacturer instructions.

Locus amplification and sequencing

Polymerase Chain Reaction (PCR) was performed to amplify the cytochrome b gene (*cytb*: 1112 bp) with the primers GLuDG (Palumbi et al., 1991) and H16460 (Perdices et al., 2002). The PCR reaction consisted of a 12.5 µl volume reaction containing 4.25 µl ultrapure water, 0.5 µl of each 0.2 µM primer, 6.25 µl Dream *Taq* Green PCR Master Mix 2x (Thermo Scientific, Waltham, MA, USA), and 1 µl (ca 10-100 ng) of DNA template. The PCR procedure consisted of 2 min at 94°C followed by 35 cycles of 45 s at 94°C for DNA denaturation, 1 min at 46.5°C for primer alignment, 1.5 min at 72°C for synthesis, and a final extension of 5 min at 72°C. After checking PCR products by electrophoresis in agarose gel of 1.5%, amplicons were purified using ExoSAP-IT (USB Corp. Cleveland, OH, USA), and submitted to Macrogen Inc. (Netherlands) for sequencing. Manual alignment of the sequences was implemented in Mega v7.0 (Kumar et al., 2016).

Phylogeny, divergence time estimation, population structure and genetic distances

Sequences were collapsed to haplotypes using the web-based program ALTER (González-Peña et al., 2010). The best-fit model of molecular evolution selected by the Akaike information criterion (AIC), using jModelTest v2 (Santorum et al., 2014), was general time reversible + invariable site proportion + gamma (GTR+I+G). Phylogenetic reconstruction was implemented using Maximum likelihood (ML) and Bayesian inference (BI). Maximum likelihood was carried out using RAxMLGUI v1.3.1 (Silvestro and Michalak, 2012; Stamatakis, 2014), with the substitution model selected by jModelTest and 10,000 bootstrap

replicates. Bayesian analysis was implemented in MrBayes v3.2.1 (Ronquist et al., 2012). The analysis was run for five million of generations, with two independent runs implementing four Markov chain Monte Carlo (MCMC) processes, sampling every 500 generations. The model of molecular evolution was the selected by jModelTest v2 (GTR+I+G). We evaluated chain convergence with the log-likelihood (-lnL) values of the runs on Tracer v1.5 (Rambaut and Drummond, 2007). After discarding the first 10% of generations as burn-in, the 50% majority rule consensus tree and posterior probabilities were obtained.

To determine the geographic correspondence with the genetic structure, a haplotype network was constructed using HaploViewer (available at http://www.cibiv.at/~%20greg/haploviewer).

The divergence time for populations of *G. atripinnis* was estimated using the program BEAST v1.8.1 (Drummond et al., 2012), and a lognormal relaxed clock (uncorrelated) model on branch length (Drummond et al., 2006). The molecular clock was calibrated using the mutation rate of *cytb* in teleosts of 0.76-2.2%/million years (Near and Benard, 2004; Zardoya and Doadrio, 1999). The substitution model was selected by jModelTest. We used the tree prior: Coalescent Bayesian Skyline (Drummond et al., 2005) and estimated a starting tree using the random method. The MCMC analysis was conducted with 30 million generations, sampling every 500 generations. We assessed whether parameters values had reached effective sample size and convergence in Tracer v1.5 (Rambaut and Drummond, 2007). Finally, the maximum clade credibility tree was built discarding the first 10% of the trees as burn-in using Tree-Anotator v1.8.1 (Drummond et al., 2012).

In order to analyse the genetic structure of populations of *G. atripinnis*, we conducted analyses of molecular variance (AMOVA) at three hierarchical levels: 1) considering the biogeographic regions proposed by Domínguez-Domínguez et al., (2006), that were based in distribution and evolution of the species of the family Goodeidae, 2) according with recovered haplogroups in haplotype network (two and three haplogroups), and 3) an analysis without *a priori* grouping. Components of the fixation index Φ_{CT}, Φ_{ST} and Φ_{SC} were also calculated using Arlequin v3.5.1.3 (Excoffier and Lischer 2010).

Uncorrected *p*-distances were estimated in MEGA v7.0 (Kumar et al.,

2016) in order to quantify genetic differences between recovered haplogroups of *G. atripinnis* (east and west + Armeria River).

Genetic diversity and historical demography

We estimated levels of genetic diversity including the number of haplotypes (H), polymorphic sites, nucleotide diversity (π), and haplotype diversity (*h*) for all biogeographic regions (*sensu* Domínguez-Domínguez et al., 2006) of *G. atripinnis* using Arlequin v3.5.1 (Excoffier & Lischer, 2010). Population size fluctuations through time were examined with a Coalescent Bayesian skyline plot analysis (Drummond et al., 2005) implemented in BEAST v1.8.1 (Drummond et al., 2012). This analysis was implemented to infer the historical demography of the recovered haplogroups in the haplotypes network. The substitution model was that obtained in jModeltest. The molecular clock was calibrated using the mutation rate used in the divergence time estimation, using an uncorrelated relaxed clock model for 50 million of generations and sampling every 500 generations. Convergence was assessed with Tracer v1.5 (Rambaut & Drummond, 2007), and were discarded the first 10% of the trees as burn-in.

RESULTS

We obtained sequences of the mitochondrial gene cytochrome b (*Cytb*: 1112 bp) from 318 individuals. Of this, 1,053 sites were invariable, while 52 were variable, 19 were singleton variable sites, and 33 were parsimony informative.

Phylogeny, haplotype network and divergence time estimation

Although the dicotomous tree obtained from the Molecular Clock Analyses in Beast recovered two genetic groups, that have correspondence with geographical distribution (west and east of the Lower Lerma region) of populations of *Goodea atripinnis*, the ML and BI analyses not support this relationship, and even the Beast tree only support the existence of the West Group (Fig. 2). The west group, cluster individuals of Bolaños, Ameca, Magdalena, Chapala, Sayula, Etzatlan-San Marcos, Cotija, Balsas, Armeria and

Santiago regions, while the not supported east group clustered individuals of the follow biogeographic regions (*sensu* Domínguez-Domínguez et al., 2006): Mezquital, Verde, Aguanaval, Panuco, Zirahuen, Zacapu, Patzcuaro, Upper Lerma, Middle Lerma, Valley of México and Juchipila. Both genetic groups have correspondence with the east of the Lower Lerma region and the west of the Lower Lerma region (Figure 1), but samples of Cuitzeo, Magdalena, Santiago and Lower Lerma regions are clustered in both clades (Figures 2 and 3).

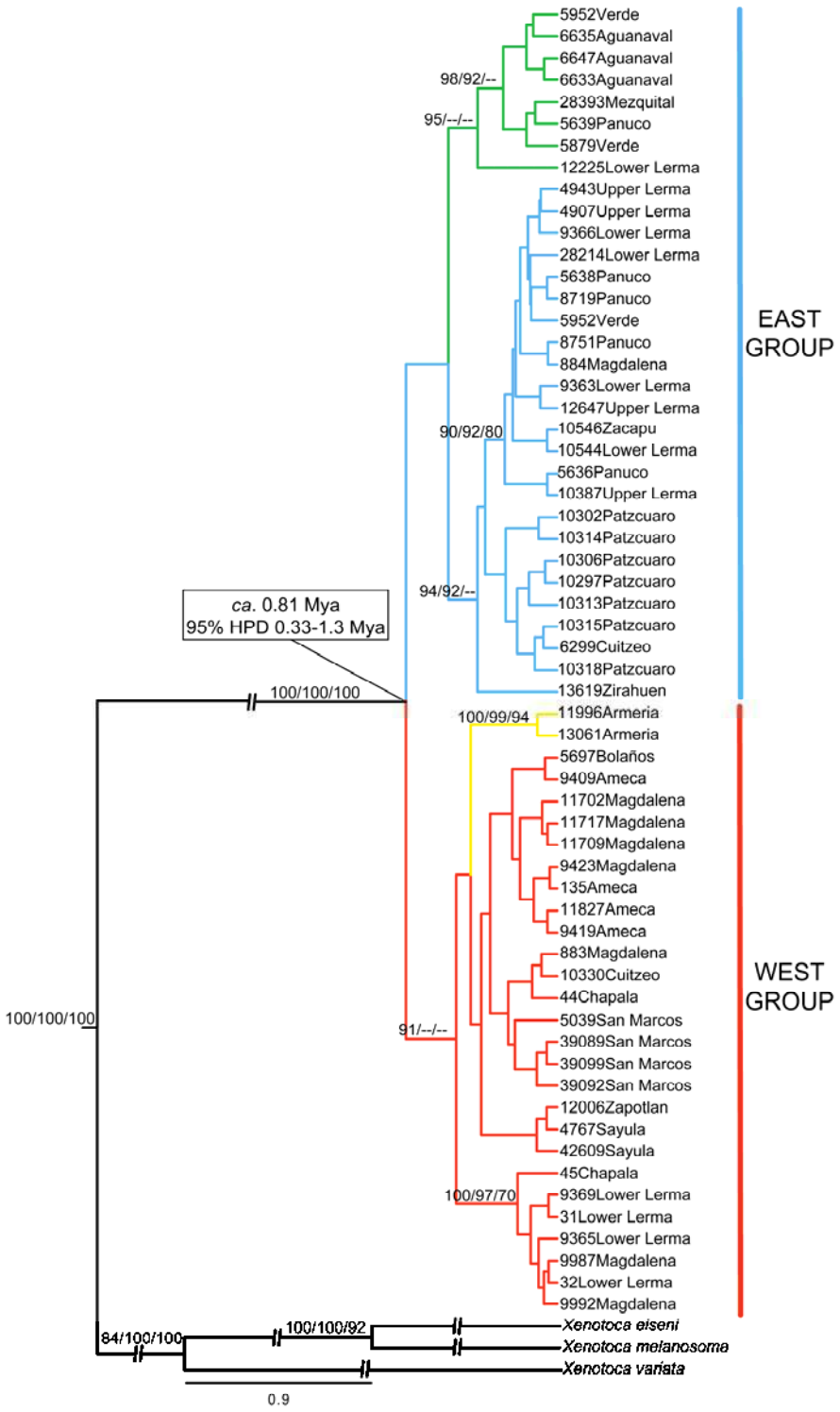


Figure 2 Haplotype phylogenetic tree obtained in BEAST. Support values are shown above the branch in the following order: 1) posterior probabilities obtained from BEAST, 2) posterior probabilities obtained from Mr. Bayes, 3) bootstrap values from a Maximum Likelihood analysis, and 4) the divergence time between the two clades/haplogroups.

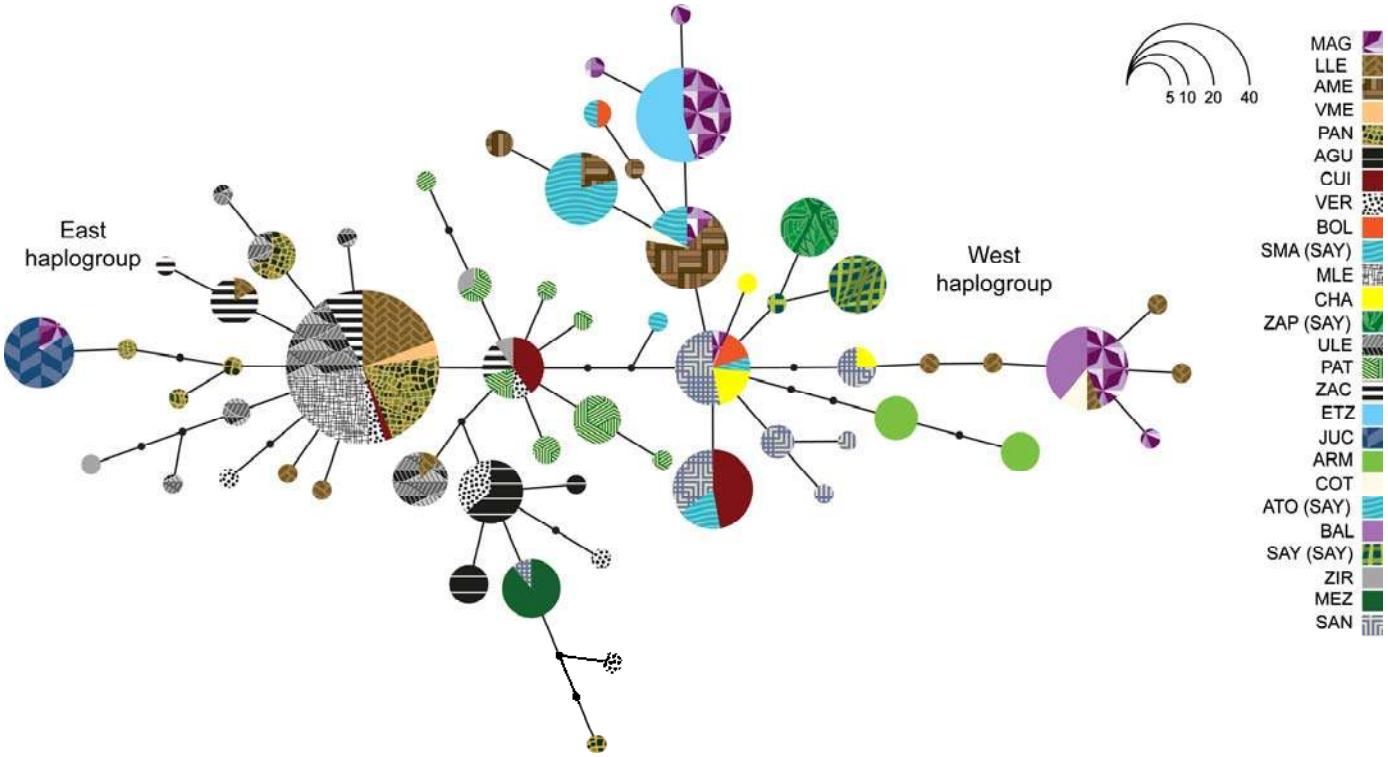


Figure 3 Haplotypes network for all sampled populations, the circles sizes corresponded with the number of samples belonging to each haplotype. Black circles represent the mutational steps.

In the haplotypes network, several biogeographic regions (*sensu* Domínguez-Domínguez et al., 2006), share haplotypes. As was explained in the phylogenetic analyses, geographic segregation was found in samples west and east of the Lower Lerma region. For instance, two main haplogroups were found and were separated by two mutational steps. One cluster the samples distributed west to the Lower Lerma region (Ameca, Bolaños, San Marcos, Zapotlan and Atotonilco, Sayula, Chapala, Etzatlan-San Marcos, Armeria, Cotija, and Balsas regions) (Figure 3). Within this west group, the Armeria, Zapotlan and Sayula samples were grouped in exclusive haplotypes separated by two mutational and one mutational step respectively.

The other haplogroup is comprised mainly by the samples collected east of the Lower Lerma region (Upper and Middle Lerma, Valley of México, Panuco, Patzcuaro, Zacapu, Zirahuen, Juchipila, Verde, Mezquital and Aguanaval regions). The samples from the northwest drainages of Juchipila, Verde,

Mezquital and Aguanaval regions belong to the east haplogroup. The samples of Cuitzeo, Magdalena, Santiago and Lower Lerma regions share haplotypes between west and east haplogroups (Figures 1 and 3).

The divergence time estimation for populations of *G. atripinnis* showed that the divergence time between the two recovered haplogroups was *ca*. 0.81 Mya (95% HPD: 0.33-1.3) during the Pleistocene (Figure 2).

Population structure and genetic distances

The analysis of molecular variance implemented by biogeographic regions showed that the percentage of molecular variance among groups was of 19.09% (Φ_{CT} = 0.19, Φ_{SC} = 0.63 and Φ_{ST}: 0.70). Furthermore, the analyses of molecular variance implemented with the two inferred haplogroups showed that the percentage of molecular variance among groups was of 38.77% (Φ_{CT} = 0.38, Φ_{SC} = 0.55 and Φ_{ST}: 0.72). When a third haplogroup that was separated by two mutational steps were added (Armeria River), the percentage of molecular variance among groups was higher than when only two groups were included (42.18%, Φ_{CT} = 0.42, Φ_{SC} = 0.52 and Φ_{ST}: 0.72). Also, in the arrangement of no grouping *a priori*, the highest percentage of variation was among populations with 67.84% (Φ_{ST}: 0.67). The only comparison that was not significant was when the samples were grouped according to the biogeographical regions (Table 2).

The genetic distances, based on the three genetic groups were of 0.5% between west *vs*. east and Armeria, and of 0.7% of Armeria *vs*. east group (Table 3).

Table 2 Analyses of molecular variance for groups according to: 1) 23 biogeographic regions considered in the present work; 2) the two recovered haplogroups; 3) The two recovered haplogroups + Armeria River; and 4) Without grouping *a priori*. Ns = not significant.

		Cytb		
Testing assumptions	Source of variation	% of variance	Fixation index	*P*- value
Groups	Among groups	19.09	Φ_{CT}: 0.19	Ns
According with	Among populations within groups	51.19	Φ_{SC}: 0.63	<0.0001
biogeographical regions	Within populations	29.71	Φ_{ST}: 0.70	<0.0001
	Total	100		
According with the two	Among groups	38.77	Φ_{CT}: 0.38	<0.0001
recovered haplogroups	Among populations within groups	33.81	Φ_{SC}: 0.55	<0.0001
West and East	Within populations	27.42	Φ_{ST}: 0.72	<0.0001
	Total	100		
According with the two	Among groups	42.18	Φ_{CT}: 0.42	<0.0001
recovered haplogroups	Among populations within groups	30.46	Φ_{SC}: 0.52	<0.0001
West and East + Armeria	Within populations	27.35	Φ_{ST}: 0.72	<0.0001
River	Total	100		
No grouping *a priori*	Among populations	67.84	Φ_{ST}: 0.68	<0.0001
	Within populations	32.16		
	Total	100		

Table 3 Uncorrected genetic distances based on *cytb* within and between the three genetic groups recovered for *G. atripinnis*.

cytb	West haplogroup	East haplogroup	Armeria River
West haplogroup	0.003		
East haplogroup	0.005	0.002	
Armeria River	0.005	0.007	0.001

Genetic diversity and historical demography

In general, the genetic diversity of all populations of *Goodea atripinnis* are moderate to low, with haplotype diversity in almost all biogeographic regions

being <0.7, while the nucleotide diversity show values less than 0.001 in most of the biogeographical regions.

The highest haplotype diversity was found in Patzcuaro (*h*= 0.850), followed by Verde (*h*=0.844), both biogeographic regions belonging to the east haplogroup. Null haplotype and nucleotide diversity was found in Valley of México, Middle Lerma, Etzatlan-San Marcos, Juchipila, Balsas, Zapotlan and Mezquital regions, while Magdalena, Cotija and Zirahuen regions, as well as Verde and Lower Lerma regions exhibit the highest nucleotide diversity (*π*= 0.002; Table 4).

Table 4 Genetic diversity for each biogeographic region of *Goodea atripinnis* using mtDNA data. N, sample size, S, polymorphic sites, H, number of haplotypes, *π,* nucleotide diversity *h,* haplotype diversity.

	Cytb				
Biogeographic region	N	S	H	Π	*h*
Magdalena	32	16	9	0.002+/-0.001	0.800+/-0.04
Ameca	13	3	4	0.0006+/-0.000	0.602+/-0.130
Valley of Mexico	2	0	1	0.000+/-0.000	0.000+/-0.000
Panuco	22	12	6	0.001+/-0.000	0.580+/-0.111
Aguanaval	12	2	3	0.0005+/-0.0005	0.590+/-0.107
Middle Lerma	15	0	1	0.000+/-0.000	0.000+/-0.000
San Marcos (Sayula)	6	4	4	0.001+/-0.001	0.800+/-0.172
Atotonilco (Sayula)	15	2	3	0.0004+/-0.0004	0.447+/-0.134
Sayula (Sayula)	10	1	2	0.0001+/-0.0002	0.200+/-0.154
Cuitzeo	13	5	3	0.001+/-0.001	0.564+/-0.111
Patzcuaro	16	8	8	0.001+/-0.001	0.850+/-0.075
Zacapu	14	3	4	0.0008+/-0.0006	0.714+/-0.078
Etzatlan-San Marcos	13	0	1	0.000+/-0.000	0.000+/-0.000
Chapala	4	3	3	0.001+/-0.001	0.833+/- 0.222
Juchipila	11	0	1	0.000+/-0.000	0.000+/-0.000
Balsas	7	0	1	0.000+/-0.000	0.000+/-0.000
Cotija	3	4	2	0.002+/-0.002	0.666+/-0.314
Verde	10	6	8	0.002+/-0.001	0.844+/-0.102
Bolaños	3	1	2	0.0006+/-0.0007	0.666+/-0.314
Zapotlan	9	0	1	0.000+/-0.000	0.000+/-0.000
Upper Lerma	25	8	7	0.001+/-0.001	0.740+/-0.066
Armeria	9	2	2	0.000+/-0.000	0.555+/-0.090
Zirahuen	3	4	3	0.002+/-0.002	1.00+/-0.272
Mezquital	8	0	1	0.000+/-0.000	0.000+/-0.000
Grande de Santiago	23	10	7	0.001+/-0.001	0.806+/-0.052
Lower Lerma	20	10	8	0.002+/-0.001	0.589+/-0.129

The BSP analysis implemented for the two haplogroups showed that the west group showed a stable effective population size through time, with a population expansion calculated to be <30,000 years. While for the eastern group, population expansion occurred after the long stable effective population size and much older that the western group (≈125,000 years ago) (Figure 4).

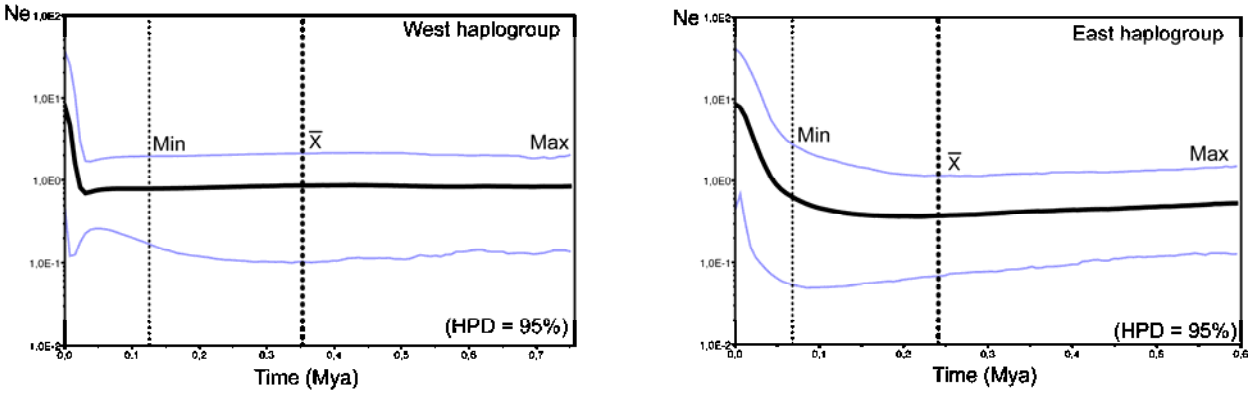


Figure 4 Bayesian Skyline Plots according to the two recovered haplogroups of the haplotype network (west clade/haplogroup and east clade/haplogroup).

DISCUSSION

The results obtained herein, using mDNA data for populations of *G. atripinnis* along its wide distributional range, reveal minimal genetic differentiation within and between biogeographic regions (Figures 2 and 3; Tables 2 and 3). These results are not in agreement with previous phylogeographic studies of freshwater fishes in central México, which generally showed high genetic structure associated with volcanic and tectonic activities, and quaternary climate change (Beltrán-López et al., 2018; Domínguez-Domínguez et al., 2008a; 2010). Some differentiation is significant enough that it has warranted the description of new species (Domínguez-Domínguez et al., 2008b; 2016; Pérez-Rodríguez et al., 2009). The lack of structure in co-distributed species indicates that *G. atripinnis* follows an independent evolutionary history that could be related to the biological and ecological characteristics of the species. *Goodea atripinnis* is a generalist feeder and environmentally tolerant species (De la Vega-Salazar, 2006; Miller et al., 2005)

that has the ability to disperse between drainages and survive and persist in new environments (Avise et al., 1987; McGlashan & Hughes, 2000).

Two main genetic groups

The haplotypes network recovered two geographically segregated haplogroups, separated by two mutational steps (Figures 1 and 3). The highest percentage of molecular variation is found when populations were grouped according with phylogenetic analyses and haplotype network and not according with hydrological basins or biogeographical regions (Beltrán-López et al., 2017; 2018; Domínguez-Domínguez et al., 2008a). These two haplogroups have geographical correspondence with one of them mainly distributed to the west and the other to the east of the Lower Lerma region, with low genetic distances of 0.5% between them (Table 3), and recent divergence times (*ca*. at 0.81 Mya) (Figure 2). This contradicts previous genetic analyses in this species, in which no genetic differences were found in 11 studied populations, with limited sample representation (Doadrio & Domínguez, 2004; Domínguez-Domínguez et al., 2010).

The biogeographic break between eastern and western haplogroups of the Lerma River was previously found in *Zoogoneticus quitzeoensis* (Goodeidae) but with a much higher level of mean genetic distance for *cytb* (3.05%) and a much older divergence time estimate of *ca*. 3.3 Mya between these two lineages (Domínguez-Domínguez et al., 2008a). Similarly, *Alloophorus robustus* had divergence times estimated for both groups at *ca*. 2.6 Mya (Domínguez-Domínguez et al., 2010). The differences in the phylogeographic patterns among these species could be related to the ecological and biological aspects among the three genera.

Previous works with goodeids has been established that the more sexually dimorphic species showed greater genetic differentiation, with much larger values of F_{ST} than monomorphic species, implying lower gene flow between populations, as predicted if stronger female-biased sexual selection promoted population differentiation (Ritchie et al., 2007), also, monomorphic species are amongst the most widespread geographic ranges of all goodeids (Ritchie et al., 2007), since *Goodea atripinnis* is a monomorphic species being the goodeid with the widest geographic distribution, also, with capability to

survive to the different conditions of the water bodies along the distribution area of the species (Miller et al., 2005), and, with the capacity to disperse between areas in different times (Domínguez-Domínguez et al., 2010), these could be reasons of its phylogeographic pattern. Another reason for the low genetic differentiation between populations of this species could be related with the high abundance that present the populations and the high tolerance to environmental degradation (De la Vega-Salazar, 2006), while *Z. quitzeoensis* and *A. robustus* are considered as environmentally sensitive species, with a lost of 25% and 33% of their historical localities respectively (De la Vega-Salazar, 2006).

Another plausible explanation for the difference in the phylogeographical pattern in these three genera is the related with the reproductive potential, reproduction studies of goodeids have been established that *G. atripinnis* has a higher reproductive potential than *Z. quitzeoensis* and *A. robustus*, has been show that while *G. atripinnis* could has an offspring number between 19.1 (Mendoza, 1962) and the maximum embryos reported is of 167 (Miller et al., 2005; Uribe et al., 2005), for *Z. quitzeoensis* has been reported between 3.33 and 13 embryos (Ramírez-Herrejón et al., 2007), For *A. robustus* has been reported between 20 and 38 embryos (Meek, 1904; Mendoza, 1962). Moreover, the reproductive epoch appear to be longer for *G. atripinnis* compare with *A. robustus* and *Z. quitzeoensis* (Barragán and Magallón, 1994; Mendoza, 1962; Miller et al., 2005; Ramírez-Herrejón et al., 2007). The above mentioned, may be related with to favoring the increased of population effective size, decresaded the genetic drift, which also, due to the high dispersal capability reported for *G. atripinnis*, is related to the homogenized among populations of one species (Hartl and Clark, 1997), as is the case of the present work.

Finally, the other plausible explanation for the differences in the phylogeographic pattern could be related with the feeding, *Z. quitzeoensis* has carnivorous feeding habits (Miller et al., 2005). Moreover, *A. robustus* seems to be the carnivorous most specialized of the family Goodeidae (Miller et al., 2005), while *Goodea atripinnis* is considered a generalist feeder that may change according with the food availability (De la Vega-Salazar, 2006; Ramírez-Herrejón et al., 2014; Soto-Galera et al., 1998; 1999). The fact that *G. atripinnis* is a generalist, could be the reason for which is easier for this species

established in the new environments, could explained its high dispersal capacity, which could explained the low genetic differentiation among all populations. Due to the above mentioned, the geological events and climate cycles could affect populations of *G. atripinnis* of a different manner respect to other goodeids.

Domínguez-Domínguez et al., (2010) proposed that the *G. atripinnis* evolved in the Santiago River, and later dispersed to Lerma River at *ca*. <1.7 Ma. The colonization of the Lerma system was followed by the dispersal and colonization of several water bodies along the TMVB. This scenario is not supported by the results shown herein, since the haplotypes from the Santiago River (BOL and JUC) are peripheral and not the central haplotypes.

Similarily that the geographic pattern showed in the distribution of the haplotypes and the genetic structure of *Goodea atripinnis*, for plants species distributed in the TMVB, a clear separation of haplotypes between east and west regions indicating a lack of gene flow between these two regions (Pérez-Crespo et al., 2017; Ruiz-Sanchez & Specht, 2013). This same pattern has been found in animals, including birds (Kingston et al., 2014), lizards and snakes (Bryson et al., 2011a,b,c,d) for which genetic structure is correlated with the western, central, and eastern regions of the TMVB. New evidence, however, suggests that climatic cycles may have had less of an effect on species-level diversification in Middle America and more of an effect on intraspecific genetic structuring (Castoe et al., 2009; Daza et al., 2009, 2010). New evidence, however, suggests that climatic cycles may have had less of an effect on species-level diversification in Middle America and more of an effect on intraspecific genetic structuring (Castoe et al., 2009; Daza et al., 2009, 2010). New evidence, however, suggests that climatic cycles may have had less of an effect on species-level diversification in Middle America and more of an effect on intraspecific genetic structuring (Castoe et al., 2009; Daza et al., 2009, 2010).

Other divergent groups

The population of *G. atripinnis* from the Armeria River shows some differentiation from other populations, with mean genetic distances between 0.5 to 0.7% with respect to west and east haplogroups respectively. Also, the

haplotype network showed unique haplotypes separated by two mutational steps from the other populations. The colonization of the Armeria River for other fish species including *Poeciliopsis infans* (Beltrán-López et al., 2018; Mateos et al., 2002) and *Moxostoma austrinum* (Pérez-Rodríguez et al., 2016), have been hypothesized to result of a unidirectional faunal exchange due to a river capture. In addition, other freshwater fishes species characteristic of the Armeria River are absent in central México, as is the case of *Allodontichthys zonistius, Ilyodon furcidens* and *Poecilia butleri* (Miller et al., 2005).

Recent dispersal events

The presence of *G. atripinnnis* in some peripheral basins such as the Sayula and Panuco regions correspond with the distribution of other fish species as *Ameca splendens*, *Xenotoca melanosoma* and *Xenotoca variata* (Domínguez-Domínguez et al., 2010), *Yuririra alta* (Domínguez-Domínguez et al., 2007), *Algansea tincella* (Pérez-Rodríguez et al., 2009) and *Poeciliopsis infans* (Beltrán-López et al., 2018; Mateos et al., 2002) and events related with river piracy during recent geological times (less than 1 Mya). River capture is considered to be the main mode of dispersal that has influenced the recent evolutionary history of goodeines, at least during the last 2 Mya (Domínguez-Domínguez et al., 2010). This, together with the high dispersal capacity for *G. atripinnis* could explain the low genetic distances between haplogroups, and the shared haplotypes between several biogeographical regions. The more common haplotype grouped the following biogeographic regions: Upper, Middle and Lower Lerma River, Zacapu and Cuitzeo Lakes, Panuco River and Valley of México River.

Also, populations of *G. atripinnis* from the Santiago River region were grouped in both haplogroups, the samples collected upstream of Salto de Juanacatlan were grouped in the east haplogroup, while populations downstream of Salto de Juanacatlan were grouped in the west haplogroup, indicating that the Salto de Juanacatlan act as a geographic barrier for different species of freshwater fishes (Figure 1). Similar results were found for a sympatric species, *Poeciliopsis infans* (Beltrán-López et al., 2018).

Genetic diversity and demographic history

The genetic diversity in general was moderate to low (Table 4), except for the west drainages Magdalena, San Marcos of the Sayula region and Chapala, as well as the east drainage Patzcuaro, Verde and Santiago regions, that was high, with h >0.8 and π >0.001.

For Magdalena region, it was reported that *P. infans* shows minimal genetic diversity which was attributed to extreme and intermittent periods of flooding and drying (Beltrán-López et al., 2018). The results for *G. atripinnis* presented herein, do not agree with these results, since individuals from Magdalena Lake show high levels of genetic diversity (Table 4).

The high genetic diversity for these populations of *G. atripinnis* could be explained the same manner that for other goodeids as a response to the gradual reduction in population sizes, which allowing to the species to retain high levels of genetic diversity along the time (Ornelas-García et al., 2012), the second and most plausible explanation to this high genetic diversity is due to the migration between populations, which is an important source of new alleles and limit the loss of genetic diversity (Vega et al., 2007), considering that *G. atripinnis* is highly abundant in the areas were is present (Miller et al., 2005) and that is widely distributed with a high colonization success (De la Vega-Salazar, 2006).

In contrast, seven biogeographical regions show null genetic diversity (Valley of México, Middle Lerma, Juchipila, Mezquital, Balsas, Etzatlan-San Marcos and Zapotlan Lake of Sayula region) (Table 4). The most plausible reason that explains the null genetic diversity for *G. atripinnis* in these biogeographic regions is a founder effect due to recent human mediated dispersal or recent connection of the hydrological basins (Beltrán-López et al., 2018; Corona-Santiago et al., 2015). In the case of some drainage where *G. atripinnis* shows null genetic diversity (i.e. Valley of México, Mezquital and Balsas Rivers), a recent natural dispersal event is not supported since these drainages have a long history of isolation and no recent connections between these basins and other basins have been found (Domínguez-Domínguez et al., 2006; 2010). Furthermore, the shared haplotypes in the network seems to indicate that populations of this species in these hydrological basins could be related to human introduction, possibly acompaning the introductions of *Chirostoma* and *Tilapia* (*Oreochromis* and *Tilapia*).

This scenario has been proposed in previous works in which the distribution in some hydrological basins of some species of goodeids and populations of *Poeciliopsis infans*, were the result of human introductions which promoted a founder effect and subsequent rapid fixation of a single haplotype in populations (Beltrán-López et al., 2018; Corona-Santiago et al., 2015; Galindo-Villegas & Sosa-Lima, 2002; Ramírez-Herrejón et al., 2013).

Goodeines have had great importance as food for indigenous populations since their establishment 3,000 years ago (Parsons, 2010). Even today, these fishes are important part of the food supply for natives of central México. Archeological remains of *G. atripinnis* have been found in several parts of central México, reinforcing the idea that humans could have moved specimens between basins (Bravo-Espinosa et al., 2009; Guzmán et al., 2001). In other cases, such as Etzatlan-San Marcos, where null genetic diversity was found, this could be related with the irrigation channels that were built between the Etzatlan-San Marcos endorrehic basin and Magdalena Lake. This is also congruent with the haplotypes network, since the only haplotype of Etzatlan-San Marcos endorrehic basin is shared with Magdalena Lake.

The demographic history found for the two haplogroups of *G. atripinnis* in the BSP analyses recovered stable effective population size through time with a recent population expansion in <30,000 years for the west haplogroup (LGM: Last Glacial Maximum), while for the east haplogroup, the population expansion was <125,000 years (LIG: Last Inter Glacial) (Figure 4). These results contrast the result of other analyses implemented for freshwater fishes of central México, which show a demographic decline in the last 0.150-0.100 Mya (LIG), as is the case of *P. infans* (Beltrán-López et al., 2018). For *G. atripinnis*, glacial and inter-glacial climate change did not seem to have a negative effect in the effective population size of *G. atripinnis* populations through time.

Taxonomic implications

The taxa with *Goodea* has not been consistent, with three described species (*Goodea atripinnis*, *Goodea luitpoldii* and *Goodea gracilis)* being recognized at one time or another. However, the samples of the three species included herein were not recovered as monophyletic and shared haplotypes between them were found (Figures 2 and 3) suggesting that a three taxa

system is not taxonomic accurate. These three species were described based on morphology (Miranda et al., 2010), but it has been shown that the goodeids exhibit high morphological variation that does not agree with genetic data, as is the case for *Ilyodon*, in which three described species clustered in the same clade in phylogenetic analyses, and also the existence of shared haplotypes between valid species (Beltrán-López et al., 2017).

Moreover, populations of the genus *Goodea* show similar phenotypic responses to similar environmental gradients and flow regimes, despite the lack of intraspecific genetic variation across their respective range, also showed that populations of *Goodea atripinnis* may show adaptive responses to divergent habitats, including mouth position, dorsal fin position, anal fin position, and caudal peduncle (length and width), considering that phenotypic plasticity may promote morphological response as a result of environmental changes (Crispo, 2008; Foster et al., 2015). Also, the morphological differences might have arisen as a result of divergent selection pressures of water flow differences, dissolved oxygen variation, or prey type/abundance variation between lotic and lentic habitats (Crispo and Chapman, 2010). Accordingly, based on the result of this study, we consider all populations analyzed in the present work to belong to a single taxonomic unit: *Goodea atripinnis* (Jordan, 1880).

Acknowledgements

This work was conducted as part of the doctoral studies of RGBL, and she was supported with a scholarship from the Consejo Nacional de Ciencia y Tecnología (CONACYT: scholarship number 329883). We thank Berenice García and Isaí Betancourt for their help with sampling. This work was supported, in part, by PRODEM, CIC-UMSNH 2015-2017 to ODD, as fund from the U.S. National Science Foundation grant (DEB 1354930) to KRP, and the Ministerio de Economía y Competitividad y FEDER, Spain (CGL2016-75262-P) to ID.

Figure and Tables legends

Figure 1. Distribution map of *Goodea atripinnis* samples, diamonds shown the localities sampled for the present work, the shaded part show the geographic

distribution of the two recovered clades/haplogroups obtained in phylogenetic analyses. The abbreviations of the biogeographical regions are: MAG: Lake Magdalena, LLE: Lower Lerma river, AME: Ameca river, VME: Valley of México, PAN: Panuco river, AGU: Aguanaval river, CUI: Lake Cuitzeo, VER: Verde river, BOL: Bolaños river, MLE: Middle Lerma river, CHA: Lake Chapala, ULE: Upper Lerma river, PAT: Lake Patzcuaro, ZAC: Lake Zacapu, ETZ: Etzatlan-San Marcos, JUC: Juchipila river, ARM: Armeria river, COT: Lake Cotija, ZIR: Lake Zirahuen, MEZ: Mezquital river, SAN: Santiago river, BAL: Balsas river, ATO (3): Lake Atotonilco, SMA (4): Lake San Marcos, SAY (5): Lake Sayula and ZAP (6): Lake Zapotlan, the last four numbers corresponded with the Sayula region.

Figure 2. Haplotype phylogenetic tree obtained in BEAST. Support values are shown above the branch in the following order: 1) posterior probabilities obtained from BEAST, 2) posterior probabilities obtained from Mr. Bayes, 3) bootstrap values from a Maximum Likelihood analysis, and 4) the divergence time between the two clades/haplogroups.

Figure 3. Haplotypes network for all sampled populations.

Figure 4. Bayesian Skyline Plots according to the two recovered haplogroups of the haplotype network (west clade/haplogroup and east clade/haplogroup).

Table 1. Samples localities and sequences information.

Table 2. Analyses of molecular variance for groups according to: 1) 23 biogeographic regions considered in the present work; 2) the two recovered haplogroups; 3) the two recovered haplogroups + Armeria River, and 4) without grouping *a priori.*

Table 3. Uncorrected genetic distances based on *cytb* within and between the recovered genetic groups of *Goodea atripinnis*.

Table 4. Genetic diversity for each biogeographic region of *Goodea atripinnis* using mtDNA data. N, sample size, S, polymorphic sites, H, number of haplotypes, *π,* nucleotide diversity *h,* haplotype diversity

REFERENCES

Anducho-Reyes, M.A., Cognato, A.I., Hayes, J.L. & Zuñiga, G. (2008). Phylogeography of the bark beetle *Dendroctonus mexicanus* Hopkins (Coleoptera: Curculionidae: Scolyti- nae). *Molecular Phylogenetics and Evolution*, 49, 930–940.

Avise, J. C., Arnold, J., Ball, R. M., Bermingham, E., Lamb, T., Neigel, J. E., Reeb, C. A. & Saunders, N. C. (1987). Intraspecific phylogeography: The mitochondrial DNA bridge between population genetics and systematics. *Annual Review in Ecology and Systematics*, 18, 489-522.

Barragán, J. & Magallón, S. B. (1994). Peces Dulceacuícolas Mexicanos X. *Goodea atripinnis* (Cyprinodontiformes: Goodeidae). Zoología Informa, 28, 27-36.

Beltrán-López, R. G., Domínguez-Domínguez, O., Guerrero, J. A., Corona-Santiago, D. K., Mejía-Mojica, H., & Doadrio, I. (2017). Phylogeny and taxonomy of the genus *Ilyodon* Eigenmann, 1907 (Teleostei: Goodeidae), based on mitochondrial and nuclear DNA sequences. *Journal of Zoological Systematics and Evolutionary Research*, 55(4), 340-355. DOI: 10.1111/jzs.12175

Beltrán-López, R. G., Domínguez-Domínguez, O., Pérez-Rodríguez, R., Piller, K. R., & Doadrio, I. (2018). Evolving in the highlands: The case of the Neotropical Lerma Live-bearing *Poeciliopsis infans* (Woolman, 1894) (Cyprinodontiformes: Poeciliidae) in central México. *BMC Evolutionary Biology*. 18:56, https://doi.org/10.1186/s12862-018-1172-7.

Bermingham, E. & Martin, A. P (1998). Comparative mtDNA phylogeography of neotropical freshwater fishes: testing shared history to infer the evolutionary landscape of lower Central America. *Molecular Ecology*, 7(4), 499-517.

Betancourt-Resendes, I., Pérez-Rodríguez, R., & Domínguez-Domínguez, O. (2018). Speciation of silverside *Chirostoma attenuatum* (Pisces:

Atheriniformes) in Central Mexico. *Journal of Zoological Systematics and Evolutionary Research*, 56(3), 323-334. DOI: 10.1111/jzs.12216

Bravo-Espinosa, M., Mendoza-Cantú, M. E. & Medina-Orozco, L. E. (2009). Escenarios de erosión bajo diferentes manejos agrícolas en la cuenca del lago de Zirahuén, Michoacán, México. *Investigaciones geográficas*, 68, 73–84.

Bryson, R. W. Jr., García-Vázquez, U. O. & Riddle, B. R. (2011a). Phylogeography of Middle American gopher-snakes: mixed responses to biogeographical barriers across the Mexican Transition Zone. *Journal of Biogeography*, 38, 1570–1584.

Bryson, R. W. Jr., García-Vázquez, U. O. & Riddle, B. R. (2011b). Relative roles of Neogene vicariance and Quaternary climate change on the historical diversification of bunchgrass lizards (*Sceloporus scalaris* group) in Mexico. *Molecular Phylogenetics and Evolution*, 62, 447–457.

Bryson, R. W. Jr., Murphy, R. W., Lathrop, A. & Lazcano-Villareal, D. (2011c). Evolutionary drivers of phylogeographical diversity in the highlands of Mexico: a case study of the *Crotalus triseriatus* species group of montane rattle-snakes. *Journal of Biogeography*, 38, 697–710.

Bryson, R. W. Jr., Murphy, R. W., Graham, M. R., Lathrop, A. & Lazcano-Villareal, D. (2011d). Ephemeral Pleistocene woodlands connect the dots for highland rattlesnakes of the *Crotalus intermedius* group. *Journal of Biogeography*, 38, 2299–2310.

Clobert, J., Danchin, E., Dhondt, A. A. & Nichols, J. D. (2001). Dispersal. Oxford University Press, New York. pp. 452.

Collin, H. & Fumagalli, L. (2011). Evidence for morphological and adaptive genetic divergence between lake and stream habitats in European minnows (*Phoxinus phoxinus*, Cyprinidae). *Molecular Ecology*, 20, 4490–4502.

Corona-Santiago, D. K., Doadrio, I., & Domínguez-Domínguez, O. (2015). Evolutionary history of the livebearing endemic *Allotoca diazi* species complex (Actinopterygii, Goodeinae): Evidence of founder effect events in the Mexican Prehispanic period. *PLoS ONE*, 10, e0124138. https://doi.org/10.1371/journal.pone.0124138. PMID: 25946217.

Crispo, E. (2008). Modifying effects of phenotypic plasticity on interactions among natural selection, adaptation and gene flow. *Journal of Evolutionary Biology*, 21: 1460–1469.

Crispo, E. & Chapman, L. J. (2010). Geographic variation in phenotypic plasticity in response to dissolved oxygen in an African cichlid fish. *Journal of Evolutionary Biology*, 23, 2091-2103, doi:10.1111/j.1420-9101.2010.02069.x

De la Vega-Salazar, M. Y. (2006). Estado de Conservación de los peces de la familia Goodeidae (Cyprinodontiformes) en la mesa central de México. *Revista de Biología Tropical*, 54(1), 163-177.

Doadrio, I., & Domínguez, O. (2004). Phylogenetic relationships within the fish family Goodeidae based on cytochrome b sequence data. *Molecular Phylogenetics and Evolution*, 31, 416–430.

Domínguez-Domínguez, O., Pompa-Domínguez, A. & Doadrio, I. (2007). A new species of the genus *Yuriria* Jordan & Evermann, 1896 (Actinopterygii, Cyprinidae) from the Ameca basin of the central Mexican plateau. *Graellsia*, 63(2), 259–271.

Domínguez-Domínguez, O., Alda, F., Pérez-Ponce de León, G., García-Garitagoitia, J. L., & Doadrio, I. (2008a). Evolutionary history of the endangered fish *Zoogoneticus quitzeoensis* (Bean, 1898) (Cyprinodontiformes: Goodeidae) using a sequential approach to phylogeography base on mitochondrial and nuclear DNA data. *BMC Evolutionary Biol- ogy*, 8, 161. https://doi.org/10.1186/1471-2148-8-161.

Domínguez-Domínguez, O., Bernal-Zuñiga, D. M., & Piller, K. R. (2016). Two new species of the genus *Xenotoca* Hubbs and Turner, 1939 (Teleostei, Goodeidae) from central-western México. *Zootaxa*, 4189, 81-98.

Domínguez-Domínguez, O., Doadrio, I., Martínez-Meyer, E., Zambrano, L., & Pérez-Ponce de León, G. (2006). Using ecological niche modeling as a conservation tool for freshwater species: The live-bearing fishes (Goodeidae) in Central México. *Conservation Biology*, 20, 1730–1739.

Domínguez-Domínguez, O., Pedraza-Lara, C., Gurrola-Sánchez, N., Perea, S., Pérez-Rodríguez, R., Israde-Alcántara, I., ... Brooks, D. R. (2010). Historical biogeography of the goodeinae (Cyprinodontiformes). In M. C.

Uribe-Aranzabal, & H. J. Grier (Eds.), *Viviparous fishes II* (pp. 13– 30). Florida: New Life Publications.

Domínguez-Domínguez, O., Pérez-Rodríguez, R. & Doadrio, I. (2008b). Morphological and genetic comparative analyses of populations of *Zoogoneticus quitzeoensis* (Cyprinodontiformes: Goodeidae) from Central México, with description of a new species. *Revista Mexicana de Biodiversidad*, 79, 373-383.

Drummond, A. J., Rambaut, A., Shapiro, B. & Pybus, O. G. (2005). Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and Evolution*, 22, 1185-1192.

Drummond, A. J., Ho, S. Y. W., Phillips, J. & Rambaut, A. (2006). Relaxed phylogenetics and dating with confidence. *PLoS Biology*, 5:e88.

Drummond, A. J., Suchard, M. A., Xie, D., & Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, 29, 1969–1973.

Excoffier, L., & Lischer, H. E. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, 10, 564–567.

Faulks, L. K., Gilligan, D. M. & Beheregaray, L. B. (2010). Evolution and maintenance of divergent lineages in an endangered freshwater fish, *Macquaria australasica*. *Conservation Genetics*, 11(3), 921-934.

Ferrusquía-Villafranca, I. (1993). Geología de México: una sinopsis. In T. P. Ramamoorthy, R. Bye, A. Lot, & J. Fa (Eds.), Biological diversity of Mexico: Origins and distribution. (pp. 3–108). Nueva York: Oxford.

Foster, K., Bower, L., & Piller, K. (2015). Getting in shape: habitat-based morphological divergence for two sympatric fishes. *Biological Journal of the Linnean Society*, 114, 152-162.

Galindo-Villegas, J., & Sosa-Lima, E. (2002). Gonopodial system review and a new record of *Poeciliopsis infans* (Cyprinodontiformes: Poeciliidae) for Lake Patzcuaro, Michoacán, central México. *Revista de Biología Tropical*, 50, 3-4.

González-Peña, D., Gómez-Blanco, D., Reboiro-Jato, M., Fernández-Riverola, F. & Posada, D. (2010). ALTER: program-oriented conversion of DNA and protein alignments. *Nucleic Acids Research*, 38, 14-18.

Goto, A. & Andoh, T. (1990). Genetic divergence between the sibling species of river-sculpin, *Cottus amblystomopsis* and *C. Nozawae*, with special reference to speciation. *Environmental Biology of Fishes*, 28, 257-266.

Guzmán, A. F., Polaco, Ó. J. & Pollard, H. P. (2001). Ofrendas de peces asociadas a entierros del Clásico-Epiclásico en Urichu, Michoacán, México. *Archaeofauna*, *International Journal of Archaeozoology*, 10, 149–162.

Hartl, D. L. & Clark, A. G. (1997). Principles of Population Genetics. Sinauer Assciates. 545 pp.

Hewitt, G. M. (2000). The genetic legacy of the quaternary ice ages. *Nature*. 405(6789), 907–913.

Hewitt, G. M. (2004). Genetic consequences of climatic oscillations in the quaternary. Philos Trans R Soc B, 359(1442):183–195.

Huidobro, L., Morrone, J. J., Villalobos, J. L. & Álvarez, F. (2006). Distributional patterns of freshwater taxa (fishes, crustaceans and plants) from the Mexican transition zone. *Journal of Biogeography*, 33, 731-741.

Kingston, S. E., Navarro-Sigüenza, A. G., García-Trejo, E. A., Vázquez.Miranda, H., Fagan, W. F., Braun, M. J. (2014). Genetic differentiation and habitat connectivity across towhee hybrid zones in Mexico. *Ecology and Evolution*, 28, 277-297.

Kumar, S., Stecher, G. & Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Molecular Biology and Evolution, 33(7), 1879-1874, doi: 10.1093/molbev/msw054.

Loxterman, J. L. & Keeley, E. R. (2012). Watershed boundaries and geographic isolation: patterns of diversification in cutthroat trout from western North America. BMC Evolutionary Biology, 12:38, https://doi.org/10.1186/1471-2148-12-38

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, 1586-1600.

Mateos, M., Sanjur, O. I., & Vrijenhoek, R. C. (2002). Historical biogeography of the livebearing fish genus *Poeciliopsis* (Poeciliidae: Cyprinodontiformes). *Evolution*, 56, 972–984.

McGlashan, D. J. & Hughes, J. M. (2000). Reconciling patterns of genetic variation with stream structure, earth history and biology in the Australian freshwater fish *Craterocephalus stercusmuscarum* (Atherinidae). *Molecular Ecology*, 9, 1737-1751.

Meek, S. E. (1904). The fresh-water fishes of Mexico north of the Isthmus of Tehuantepec. Field Columbian Mus. Publ. 93, Zool. Ser. 5: i-lxiii, 1-252, 72 figs., 17 láms., 1 mapa.

Mendoza, G. (1962). The reproductive cycles of three viviparous teleosts, *Allophorus robustus*, *Goodea luitpoldii* and *Neoophorus diazi*. *The Biological Bulletin*, 123, 351-365, doi:10.2307/1539280

Miranda, R., Galicia, D., Monks, S., & Pulido-Flores, G. (2010). First record of *Goodea atripinnis* (Cyprinodontiformes: Goodeidae) in the state of Hidalgo (México) and some considerations about its taxonomic position. *Hidrobiológica*, 20, 185-190.

Miller, R. R., Minckley, W. L. & Norris, S. M. (2005). Freshwater Fishes of México. Chicago, IL: The University of Chicago Press.

Near, T. J., & Benard, M. F. (2004). Rapid allopatric speciation in logperch darters (Percidae: Percina). *Evolution*, 58, 2798–2808.

Ornelas-García, C. P., Alda, F., Díaz-Pardo, E., Gutiérrez-Hernández, A. & Doadrio I. (2012). Genetic diversity shaped by historical and recent factors in the live-bearing twoline skiffia *Neotoca bilineata*. *Journal of Fish Biology*, 81, 1963-1984.

Palumbi, S., Martin, A., Romano, S., McMillan, W. O., Stice, L., Grabowski, G. (1991). Simple fool´s guide to PCR. University of Hawaii, Honolulu.

Parsons, J. R. (2010). The pastoral niche in pre-Hispanic Mesoamerica. In *Pre-Columbian Foodways* (pp. 109-136). Springer, New York, NY.

Pedraza-Lara, C., Doadrio, I., Breinholt, J. W. & Crandall, K. A. (2012). Phylogeny and Evolutionary Patterns in the Dwarf Crayfish Subfamily (Decapoda: Cambarellinae). *Plos One*, 7, e48233.

Perdices, A., Bermingham, E., Montilla, A., & Doadrio, I. (2002). Evolutionary history of the genus Rhamdia (Teleostei: Pimelodidae) in Central America. *Molecular Phylogenetics and Evolution*, 25, 172-189.

Perea, S., Cobo-Simo, M. & Doadrio, I. (2016). Cenozoic tectonic and climatic events in southern Iberian peninsula: implications for the evolutionary history of freshwater fish of the genus *Squalius* (Actinopterygii, Cyprinidae). *Molecular Phylogenetics and Evolution*, 97:155–169.

Pérez-Crespo, M. J., Ornelas, J. F., González-Rodríguez, A., Ruiz-Sanchez, E., Vásquez-Aguilar, A. A. & Ramírez-Barahona, S. (2017). Phylogeography and population differentiation in the *Psittacanthus calyculatus* (Loranthaceae) mistletoe: a complex scenario of climate-volcanism interaction along the Trans-Mexican Volcanic Belt. *Journal of Biogeography*, 44, 2501-2514.

Pérez-Rodríguez, R., Domínguez-Domínguez, O., Pérez-Ponce de León, G., & Doadrio, I. (2009). Phylogenetic relationships and biogeography of the genus *Algansea* Girard (Cypriniformes: Cyprinidae) of central México inferred from molecular data. *BMC Evolutionary Biology*, 9, 223.

Pérez-Rodríguez, R., Domínguez-Domínguez, O., Doadrio, I., Cuevas- García, E., & Pérez-Ponce de León, G. (2015). Comparative historical biogeography of three groups of Neartic freshwater fishes across central México. *Journal of Fish Biology*, 86, 993–1015.

Pérez-Rodríguez, R., Domínguez-Domínguez, O., Mar-Silva, F., Doadrio, I. & Pérez-Ponce de León, G. (2016). The historical biogeography of the southern group of the sucker genus *Moxostoma* (Teleostei: Catostomidae) and the colonization of Central México. *Zoological Journal of the Linnean Society*, 177, 633-647.

Piller, K. R., Kenway-Lynch, C. S., Camak, D. T., & Domínguez-Domínguez, O. (2015). Phylogeography and population structure of the imperilled redtail splitfin (Goodeidae: *Xenotoca eiseni*): Implications for Conservation. *Copeia*, 103, 440–454. https://doi.org/10.1643/CI-14- 067.

Ramírez-Herrejón, J. P., Medina-Nava, M., Salazar-Tinoco, C. I. & Zubieta, T. L. E. (2007) Algunos aspectos reproductivos de *Zoogoneticus quitzeoensis* Hubbs y Turner (1939) (Osteichtyes-Goodeidae) en la represa La Mintzita Morelia, Michoacán, México. *Biológicas Revista de la*

DES Ciencias Biológico Agropecuarias Universidad Michoacána de San Nicolás de Hidalgo, 9,63-71.

Ramírez-Herrejón J. P., Castañeda-Sam, L. S., Moncayo-Estrada, R., Caraveo-Patiño, J. & Balart, E. F. (2013). Trophic ecology of the exotic Lerma livebearer *Poeciliopsis infans* (Cyprinodontiformes: Poeciliidae) in the Lago de Pátzcuaro, Central Mexico. *Revista de Biologia Tropical*, 61(3), 1289–1300.

Ramírez-Herrejón, J. P., Zambrano, L., Mercado-Silva, N., Torres-Téllez, A., Pineda-García, F., Caraveo-Patiño, J. & Balart, E. F. (2014). Long term changes in the fish fauna of Lago de Pátzcuaro in Central México. Latin American Journal of Aquatic Research, 42(1), 137-149.

Ritchie, M. G., Hamill, R. M., Graves, J. A., Magurran, A. E., Webb, S. A. & Macías-García, C. (2007). Sex and differentiation: population genetic divergence and sexual dimorphism in Mexican goodeid fish. *Journal of Evolutionary Biology*, 20, 2048-2055.

Robinson, B. W. & Wilson, D. S. (1994). Character release and displacement in fishes: a neglected literature. *American Naturalist*, 144, 596–627.

Ruiz-Sanchez, E. & Specht, Ch. 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, 1336-1347.

Schönhuth, S., & Doadrio, I. (2003). Phylogenetic relationships of Mexican minnows of the genus *Notropis* (Actinopterygii, Cyprinidae). *Biological Journal of the Linnean Society*, 80, 323–337. https://doi.org/10.1046/j.1095-8312. 2003.00241.x.

Soto-Galera, E., Paulo-Maya, J., López-López, E., Serna-Hernández, J. A. & Lyons, J. (1998). Change in fish fauna as indication of aquatic ecosystem condition in Río Grande Morelia-Lago de Cuitzeo basin, México. *Environmental Management*, 24, 133-140.

Soto-Galera, E., Díaz-Pardo, E., López-López, E. & Lyons, J. (1999). Fish as indicators of environmental quality in the Río Lerma Basin, México. *Aquatic Ecosystems Health & Management*, 1, 267-276.

Rambaut, A., & Drummond, A. (2007). Tracer v1.5. Available: http://tree. bio.ed.ac.uk/software/tracer

Ronquist, F., Teslenko, M., Van der Mark, P., Ayres, D. L., Darling, A., Höhna, S., . . . Huelsenbeck, J. P. (2012). MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, 61, 539–542.

Santorum, J. M., Darriba, D., Taboada, G. L., & Posada, D. (2014). Jmodeltest.org: selection of nucleotide substitution models on the cloud. *Bioinformatics*, 30, 14-18.

Silvestro, D., & Michalak, I. (2012). raxmlGUI: A graphical frontend for RAxML. *Organisms, Diversity, and Evolution*, 12, 335–337. https://doi. org/10.1007/s13127-011-0056-0.

Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30, 1312–1313.

Uribe, M. C., De la Rosa-Cruz, G. & García-Alarcón, A. (2005). The ovary of viviparous teleost. Morphological differences between the ovaries of *Goodea atripinnis* and *Ilyodon whitei* (Goodeidae). In: Uribe, M. C. & Grier, H. J. (Eds), Viviparous Fishes, New Life Publications, Florida, pp: 217-236.

Vega, R., Vazquez-Dominguez, E., Mejia-Puente, A. & Cuaron, A. D. (2007). Unexpected high levels of genetic variability and the population structure of an island endemic rodent (*Oryzomys couesi cozumelae*). *Biological Conservation*, 137, 210–222.

Webb, S. A., Graves, J. A., Macías-García, C., Magurran, A. E., Foighil, D. O., & Ritchie, M.G. (2004). Molecular phylogeny of the livebearing Goodeidae (Cyprinodontiformes). *Molecular Phylogenetics and Evolution*, 30, 527–544.

Zardoya, R., & Doadrio, I. (1999). Molecular evidence on the evolutionary and biogeographical patterns of European Cyprinids. *Journal of Molecular Evolution*, 49, 227–237.

CAPÍTULO III

Processes determining the cladogenetic events among populations of *Xenotoca variata* (Bean, 1887) (Cyprinodontiformes: Goodeidae)

Abstract

Geological and climatic events represent the primary explanation to describe evolutionary processes among species. Freshwater fish groups have been used as a model species to uncover evolutionary and historical biogeography patterns in central´s Mexico hydrographic systems. *Xenotoca variata* was studied, being one of the most widely distributed species among biogeographic regions of central Mexico, it has ecological particularities and represent a highly dimorphic species. In this study, the phylogeographic patterns were described in order to evaluate the role of geological events, climatic fluctuations and intrinsic characteristics on lineage diversification. One mitochondrial (*cytb*) and three nuclear loci (*S7*, *RHO* and *RAG1*) were analyzed. Phylogeographical structure, divergence time and demographic history were inferred. Were recovered two well-defined and highly supported clades, that were separated by 49 mutational steps and a divergence time of *ca*. 2.87 Mya. Genetic distance between the two clades with *cytb* was of 5.0% prove that *X. variata* is a genetically structured species. Ecological segregation and their sexual selection are determining factors of the complex phylogeographic and biogeographic history of the two highly divergent clades recovered in the present work.

Keywords: Geological history, intrinsic characteristics, endemic fish, Cuitzeo, dimorphic species

Introduction

Freshwater fishes are usually restricted to water bodies which are surrounded by biogeographic barriers (land areas). As a result, it is possible to invoke an island-like biogeographic model to explain the possible exchange of biotic components among “islands” (water bodies). This biotic exchange is only possible when a physical connection has occurred (e.g., river capture, piracy). Thus, it has been argued that such connection events, which are mainly induced by geological forces or climatic changes, have consequently influenced the exchange of fish fauna along the new dispersion route (Bermingham & Martin, 1998).

From the evolutionary perspective, geological and climatic events represent the primary explanation to describe evolutionary processes among species (Beltrán-López et al., 2018; Bermingham & Martin, 1998; Domínguez-Domínguez et al., 2008; 2010; Pérez-Rodríguez et al., 2015; 2016). In freshwater fishes, these events have an strong impact on population differentiation being linked to the evolution of hydrographic systems, which have some direct bearing on the historical biogeography of fishes due to events such as vicariance, dispersal and range expansion (Burridge et al., 2006), for which, the main explanation for phylogeographic patterns frequently lies in geological or climatic hypotheses (Beltrán-López et al., 2018; Casal-López & Doadrio, 2018; Domínguez-Domínguez et al., 2010; Edwards et al., 2005; Perea & Doadrio, 2015; Perea et al., 2016; Pérez-Rodríguez et al., 2016;).

Freshwater fish groups have been used as a model to uncover evolutionary and historical biogeography patterns in central Mexico hydrographic systems as in godeines (Beltrán-López et al., 2017; Doadrio and Domínguez, 2004; Domínguez-Domínguez et al*.,* 2006; 2008; 2010; Ritchie et al., 2005; 2007; Webb et al. 2004), cyprinids (Pérez-Rodríguez et al., 2009), atherinids (Betancourt-Resendes et al., 2018; Bloom et al., 2009) catostomids (Pérez-Rodríguez et al., 2016) and poeciliids (Beltrán-López et al., 2018; Mateos et al., 2002). Concordant and not concordant patterns have been found even within the same area, as the close relationship and even shared haplotypes within several isolated fish populations [Lerma River and Zacapu lake *versus* Cuitzeo lake populations] (Domínguez-Domínguez et al., 2007a; b;

2010; Pérez-Rodríguez et al., 2009; 2015), or the endemic fish species of Zacapu lake strongly differentiated from its sister taxa (Domínguez-Domínguez et al., 2009; 2010).

For this purpose, fishes of the subfamily Goodeinae are a good model to study the evolution of its species within a geological, climatic and biogeographical context. The Goodeinae include about 42 species, apparently originated in the Middle Miocene, followed by a diversification process in river basins from central Mexico in the last 11-15 million years (Mya) (Doadrio and Domínguez, 2004; Domínguez-Domínguez et al., 2010).

Within the Goodeinae, the species *Xenotoca variata* is one of the most widely distributed, among biogeographic regions of central Mexico (Domínguez-Domínguez et al., 2006; 2010; Fitzsimons, 1972). This widespread capability seems to be linked to the wide range of environments that occupy, its tolerance to environmental conditions, feeding habits plasticity, consuming a variety of items which can change in space and time, and its long and multiple reproductive cycles (Duarte-Sánchez, 1981; Miller *et al,* 2005). The *Xenotoca variata* and its sister species have been recognized as the most dimorphic lineage within the family Goodeidae (Ritchie et al., 2005). Previous research have been show that in the dimorphic species (i. e. *Xenotoca melanosoma*) the genetic differentiation between populations (F_{ST}) is higher than in monomorphic species, implying lower gene flow between populations, despite that males are moving between populations, but reproducing less successfully, and this support the mechanism by which sexual isolation can influence faster speciation in dimorphic species by sexual selection (Ritchie et al., 2007). Moreover, sexual selection seems to be associated with greater species diversity, encouraging the reputation of sexual selection as an engine of speciation for a broad range of organisms (Arnqvist et al., 2000; Gavrilets et al., 2001; Martin and Hosken, 2003).

In this study, the phylogeographic patterns of *Xenotoca variata* are described, to evaluate the role of geological events and climatic fluctuations on lineage diversification. Taking into account the ecological aspects of this species, their wide geographic distribution, high tolerance to environmental conditions, and generalist feeding; we can expect a lack of genetic structure

and low genetic variability between populations, as has been the case for other goodeids (Beltrán-López et al., *in prep*). An alternative hypothesis is that populations exhibit high genetic structure and high genetic differentiation, despite ecological traits, which could be related to sexual selection being a dimorphic species, besides the geologic events and climatic change during the quaternary as is a common pattern in other freshwater fishes of central México (Beltrán-López et al., 2018; Betancourt-Resendes et al., 2018; Domínguez-Domínguez et al., 2008; 2010; Pérez-Rodríguez et al., 2009; 2015; 2016).

To respond the above mentioned, one mitochondrial and three nuclear loci were analyzed. Phylogeographical structure was inferred and divergence times estimated. Finally, the demographic history and ancestral area were reconstructed. The results are discussed within a geological and climatic context, and the species intrinsic characteristics.

Methods

Sample collection

Fish exemplars were sampled at 20 sites in 13 water drainages and 7 biogeographically discrete regions (*sensu* Domínguez-Domínguez et al., 2006) throughout the range of the species in central Mexico, using seine nets and electrofishing. In total, 215 fin clippings were obtained, preserved in absolute ethanol and stored at -80°C (Table 1, Fig. 1). Some specimens were fixed in formaline for identification and deposited in the fish collection at the Universidad Michoacana de San Nicolas de Hidalgo-CPUM, Mexico (SEMARNAT registration number MICH-PEC-227-07-09). The rest of the fishes were released to the water. The number of individuals sampled per population, varied in accordance to their abundance (Table 1).

Table 1 List of sampled localities with geographic coordinates and number of individuals analyzed for each locality. The numbers in locality refers to numbers in Figure 1. *N* = sample size for genetics *cytb*/*S7*/*RAG1*/*RHO*.

Discrete biogeographic region		Locality (abbreviation)	*N*	Coordinates
Panuco	Santa Maria River	1. Tierra Quemada (TQ)	7/2/2/2	21º42’53.2” N 100º41’13.7” W
	Santa Maria River	2. Jesús María (JM)	12/1/2/2	21º55’30.5” N 100º53’57.6” W
Chapala Lake	Chapala Lake	3. La Alberca (LA)	26/6/6/6	20º03’32.9” N 102º36’33.1” W
Middle Lerma	Turbio River	4. San Francisco del Rincon (SF)	6/1/1/1	21º02’ 47.3” N 101º50’9.3” W
	Laja River	5. Gallinero Dam (Ga)	5/2/2/2	20°56'31.50"N 100°58'38.00"W
	Yuriria Dam	6. Yuriria (Yu)	7/1/1/1	20°14’13.6” N 101°09’25.5” W
	Laja River	7. Ignacio Allende Dam (IA)	12/7/8/7	20°53’11.4” N 100°47’24.2” W
	Lerma River	8. Urideo Sprig (Ur)	4/2/2/2	20°12’52.20”N 100°50’43.40”W
	Yuriria Dam	9. Joya Grande JG	11/5/5/5	20° 7’43.94”N 101° 0’24.39”W
	Turbio River	10. San Jose del Rodeo (JR)	10/3/3/3	20°55’12.20”N 101°13’19.90”W
	Laja River	11. Afluente Neutla (Ne)	3/2/2/2	20°42’16.50”N 100°52’1.60”W
Cuitzeo Lake	Cuitzeo Lake	12. Belisario (Be)	11/0/0/0	19º53’42.41” N 101º04’16.8” W
	Cuitzeo Lake	13. San Cristóbal (SC)	16/5/5/6	19º57’42” N 101º18’6” W
	Morelia River	14. La Mintzita (Mi)	10/2/4/4	19º38’40” N 101º16’28” W
	Cuitzeo Lake	15. Salvador Escalante base de datos dice Mariano Escobedo (SE)	7/2/2/2	19º57’57.4” N 101º03’19.2” W
Zacapu region	Zacapu Lake	16. Zacapu (Zc)	16/8/8/8	19º49’35” N 101º47’10” W
	Duero River	17. Melchor Ocampo Dam	12/6/6/6	20° 5’40.53”N 101°44’18.71”W

		(MO)		
Verde River	Verde River	18. La Estancia Stream (Es)	14/5/7/6	21°24'36.30"N 102°44'15.00"O
Aguanaval	Aguanaval River	19. Atotonilco (At)	11/6/6/6	23º33'33.8" N 103º15'50.3" W
	Aguanaval River	20. Valenciana (Va)	14/4/3/3	24º10'25.3" N 103º17'33.5" W

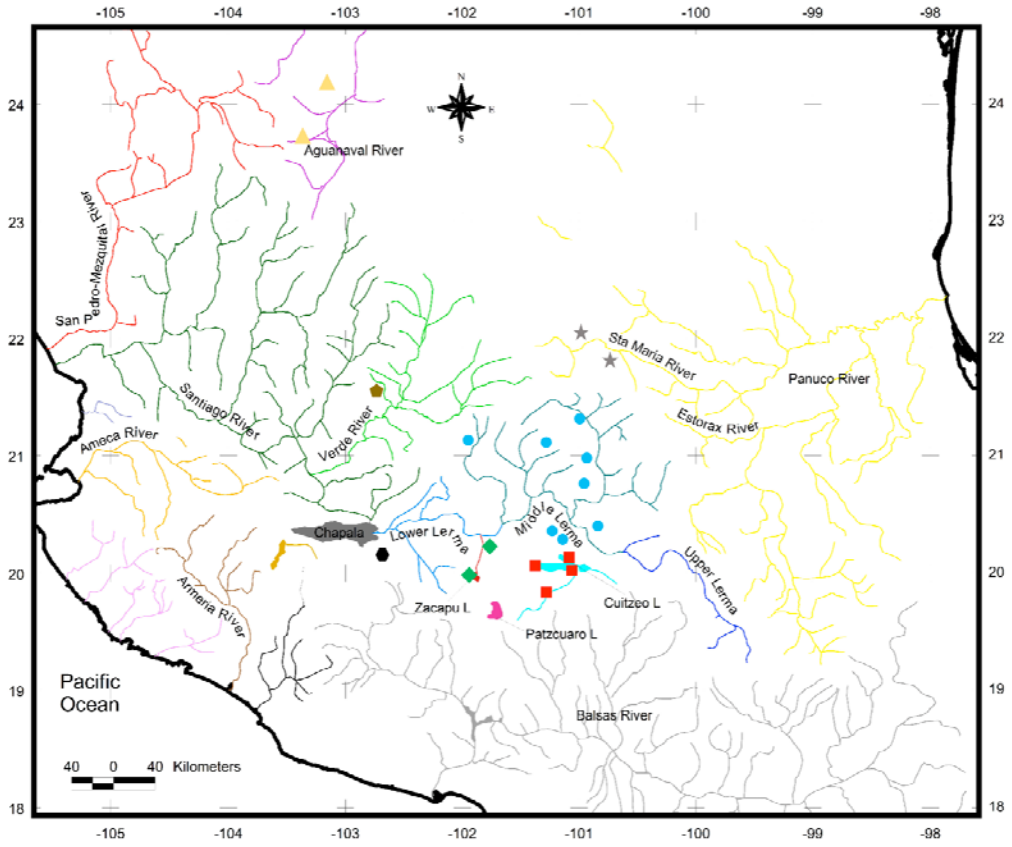


Figure 1 Sampling locations and discrete biogeographical regions where *Xenotoca variata* is distributed. The numbers of the locations has correspondence with the locations showed in table 1.

DNA extraction, PCR amplification and sequencing

Isolation of genomic DNA was performed using the QIAGEN BioSprint Dneasy Tissue and Blood Kit (Qiagen, Valencia, Ca, USA) following manufacturer's instructions. Two hundred and fifteen specimens were amplified for the mitochondrial Cytochrome b gene (*cytb*: 1,000 bp) with the primers GLU-F and Thr-R (Doadrio & Domínguez, 2004). A subset of 76 specimens were

selected to represent the genetic variation obtained with *cytb*; including all provenances and haplotypes. These were amplified for the first intron and the second exon of the gene coding for the S7 ribosomal protein (*S7*: 722 bp), the exon 3 of the recombination-activating gene 1 (*RAG1*: 1465 bp), and the gene coding for the Rhodopsin protein (*RHO*: 847 bp), a total of 4,034 bp were amplified for the four genes. The primers details are showed in Table S1.

Each fragment was individually amplified using the Polymerase Chain Reaction (PCR) for a total volume of 12.5 µl, containing 4.25 µl ultrapure water, 0.5 µl of each 0.2 µM primer, 6.25 µl Dream Taq Green PCR Master Mix 2x (Thermo Scientific), and 1 µl (ca. 10-100 ng) of DNA template. The specific PCR protocols of each gene are provided in Table S1. After checking PCR products on 1.5% agarose gels, the PCR products were purified using ExoSAP-IT (USB Corp.) and sent for sequencing at htSEQ Inc (High-Throughput Sequencing, University of Washington, USA) and Macrogen Inc. (The Netherlands). Sequences were manually aligned in Mega v7.0 (Kumar et al., 2016) and examined using chromatograms. For the nuclear genes, a Bayesian computational inference of nuclear genes gametic phase was performed using DNAsp v5.10 (Librado & Rozas, 2009), using the algorithm provided by PHASE v2.0 (Stephens & Donelly, 2003). Codification in aminoacids was used to verify the alignment and the absence of stop codons. The obtained sequences were deposited in GenBank under the follow access number (*cytb*: xxx-xxx; *S7*: xxx-xxx; *RAG1*: xxx-xxx; and *RHO*: xxx-xxx) (see Table S2).

Phylogenetic analyses and haplotype networks

Recombination of nuclear genes was tested carrying out a PHI test in Splitstree v4.13 (Huson & Bryant, 2006) the test for the three genes did not find significant statistical evidence for recombination (*S7, P* = 1.0; *RAG1, P* = 0.24; *RHO, P* = 1.0). DNA sequences of each of the four genes (*cytb*, *S7*, *RAG1* and *RHO*) were collapsed to haplotypes using the web-based program ALTER (González-Peña et al., 2010). Model selection was based on Akaike information criteria (AIC) and optimal partition setting analysis was conducted using PartitionFinder v1.1.0 (Lanfear et al., 2012). Optimal partition setting was obtained by assigning one substitution model to each gene, using model

GTR+G for *cytb*, model TVM+I+G for *S7*, model HKY+I for *RAG1*, and model HKY+I+G for *RHO*.

The phylogenetic analyses were conducted with each gene separately, with concatenated datasets for concatenated nuclear genes and for the four concatenated genes (mtDNA + nDNA)

Phylogenetic reconstructions were assessed using two building algorithms. Maximum likelihood (ML) reconstruction was conducted in RAxMLGUI v1.3.1 (Silvestro & Michalak, 2012; Stamatakis. 2014) using the GTR+Gamma+I substitution model, and 10,000 bootstrap replicates. A Bayesian inference (BI) reconstruction was conducted with the software MrBayes v3.2.1 (Ronquist et al., 2012), for 15 million generations, with two independent runs implementing four Markov Chain Monte Carlo (MCMC) processes and sampling every 500 generations. The substitution model was the assigned for each gene according to PartitionFinder v1.1.0 (Lanfear et al., 2012). The chains convergence was evaluated with the log-likelihood (-lnL) values of two independent runs on Tracer v1.5 (Rambaut and Drummond, 2007), discarding the initial 10% of the generations (burn-in) to construct the consensus tree (σ = 0.0002). According with previous phylogenetic analyses, *Ameca splendens* and *Allophorus robustus* were used as the outgroup in both implemented algorithms (ML and BI) (Domínguez-Domínguez et al., 2010).

Four independent networks were reconstructed using Haploviewer (http://www.cibiv.at/~%20greg/haploviewer) to determine the geographic distribution of the haplotypes for all populations and the four genes.

Divergence time estimation and genetic distances

The program BEAST v1.8.1 (Drummond et al., 2012), and a lognormal relaxed clock (uncorrelated) model on branch length (Drummond et al., 2006) were used to estimate the most recent common ancestor (MRCA) for the different clades and sub-clades identified within *X. variata*. This analysis was carried out with a subset of 80 sequences that include all different haplotypes for all genes. The molecular clock was calibrated using the mutation rate of *cytb* in teleosts of 0.76%-2.2%/million years (Machordom & Doadrio, 2001; Near and Benard, 2004; Zardoya & Doadrio, 1999). Since the mutation rate is not

available for the other genes, they were included in the analysis without calibration information. The model parameters were unliked across *cytb*, *S7*, *RAG1* and *RHO* genes and substitution models were set according to the selected model for each gene by PartitionFinder v1.1.0 (Lanfear et al., 2012). In a first analysis, was selected the tree prior Coalescent: Bayesian Skyline Plot (Heled & Drummond, 2008), and estimated a starting tree using the random method. A MCMC analysis with 70 million of generations was conducted, and sampled every 1000 generations. We assessed whether parameter values had reached effective sample size and convergence in Tracer v.1.6 (Rambaut et al., 2014). Finally, the maximum clade credibility tree was built, discarding the first 10% of the trees as burn-in, using Tree Annotator v1.8.1 (Drummond et al., 2012).

Because were found two highly divergent clades within *X. variata*, in a second analysis we used a Yule tree prior to the date of the cladogenetic event, in this case, to estimate divergence times and their credibility intervals, a MCMC analysis with 70 million of generations was conducted, and sampling every 1000 generations. We assessed whether parameter values had reached effective sample size and convergence in Tracer v.1.6 (Rambaut et al., 2014). Finally, the maximum clade credibility tree was built, discarding the first 10% of the trees as burn-in, using Tree Annotator v1.8.1 (Drummond et al., 2012).

The genetic uncorrected-*p* distances (Dp) were calculated with three different grouping 1) considering the recovered major clades, 2) according to the recovered sub-clades within clades and, 3) according with the biogeographic regions proposed by Domínguez-Domínguez et al., (2006) for the four genes separately. All the analyses were carried out in the software Mega v7.0 (Kumar et al., 2016), implementing a bootstrapping process with 500 repetitions.

Genetic structure among populations

Genetic differentiation level among all geographic populations of *X. variata* was estimated with paired test fixation indices (Φ_{ST}). A Bonferroni correction (Rice, 1989) was applied to each *p*-value obtained in the paired test of genetic differentiation.

To examine genetic differentiation at different hierarchical levels, genetic structure of *X. variata* was estimated via analysis of molecular variance (AMOVA) with 10,000 permutations. The two implemented analyses were conducted for the four separate genes and grouping according to different criteria. The first analysis 1) was considering the recovered major clades, 2) according to the recovered sub-clades within clades, 3) according with the biogeographic regions proposed by Domínguez-Domínguez et al., (2006). All paired test fixation indices and AMOVAs were conducted with the software Arlequin v3.5.1.3 (Excoffier & Lischer, 2010).

Genetic diversity and historical demography

For each gene (*cytb*, *S7*, *RAG1* and *RHO*), the number of haplotypes (H), polymorphic sites (S), nucleotide (π) and haplotype (h) diversities were obtained to estimate genetic diversity levels to different grouping of *X. variata*, the same manner that the genetic distances and genetic structure, the genetic diversity was calculated for three different grouping, 1) considering the recovered major clades, 2) according to the recovered sub-clades within clades, 3) according with the biogeographic regions proposed by Domínguez-Domínguez et al., (2006).

In order to infer the population size fluctuations through time of each of the recovered groups in phylogenetic analyses of *X. variata*, a Coalescent Bayesian Skyline Plot (BSP) analysis (Drummond et al., 2005) was implemented in BEAST v1.8.1 (Drummond et al., 2012). This analysis only was implemented with sequences of *cytb* due the higher number of available sequences. The substitution model used was that obtained by PartitionFinder. The molecular clock was also calibrated using the mutation rate of *cytb* in teleosts of 0.76-2.2%/million years (Machordom & Doadrio, 2001; Near & Benard, 2004; Zardoya & Doadrio, 1999), an uncorrelated relaxed clock model was set a priori, and 70 million generations were run, sampling every 500 generations. Convergence was assessed with Tracer v1.5 (Rambaut & Drummond, 2007). The first 10% of the states were discarded as burn-in.

Results

Sequence variation

By sequencing 1,000 bp of the cytochrome *b* gene in 215 specimens of *Xenotoca variata* from 20 sampling sites in 13 water drainages and 7 biogeographically discrete regions, 60 haplotypes were identified. For *cytb* data set 121 characters were variable (12.1%), 121 polymorphic segregating sites, 70 were parsimony informative and 51 were singleton variable sites. For the nuclear gene *S7* (722 bp), of the 140 sequences that included the two alleles, 18 haplotypes were identified, for this gene, six polymorphic segregation sites; six were parsimony informative and no one singleton variable sites. For the nuclear gen *RAG1* (1463 bp), of the 150 sequences that included the two alleles, 28 haplotypes were identified, for this gene, 35 polymorphic segregation sites, 26 were parsimony informative and nine singleton variable sites. Finally, for the nuclear gen *RHO* (847 bp), of the 148 sequences that included the two alleles, seven haplotypes were identified, for this gene, six polymorphic segregation sites, five were parsimony informative and one site was a singleton variable sites.

Phylogenetic reconstruction and haplotype networks

Phylogenetic analyses with *cytb* and with the four concatenated genes (*cytb* + *S7* + *RAG1* + *RHO*) result in the formation of two well defined and highly supported clades. Clade I cluster haplotypes samples from the 16 locations from the six biogeographic regions, within this, four sub-clades were recovered, the first one (sub-clade 1) clustered haplotypes sampled from Middle Lerma, Zacapu (Duero river samples) and Panuco biogeographic regions, the sub-clade 2 clustered haplotypes sampled from Chapala lake biogeographic region, the sub-clade 3 grouped haplotypes sampled from Aguanaval and Verde (Santiago river) biogeographic region, and, the last sub-clade (sub-clade 4) grouped haplotypes sampled in Zacapu biogeographic region (Zacapu lake), (Fig. 2 and S1). Whereas, the clade II clustered haplotypes sampled in the four

locations from Cuitzeo biogeographic region (Cuitzeo lake). The phylogenetic trees for the three nuclear genes concatenated show a politomy (Fig. S2).

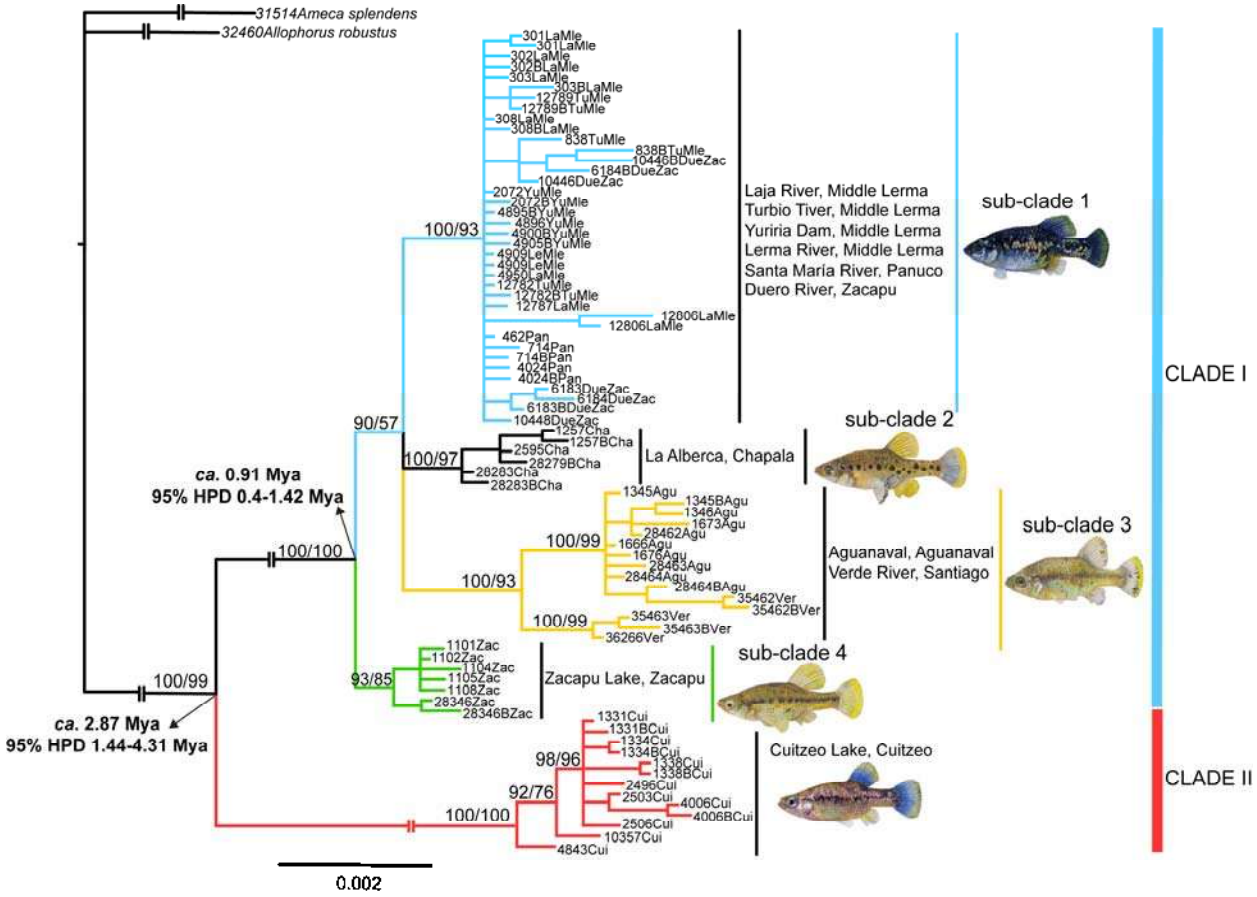


Figure 2 Bayesian inference tree of Xenotoca variata from concatenated sequences of the four genes (4,034 bp). Bayesian posterior probabilities (>90%) and maximum likelihood bootstrap values (>80%) (BBP/MLB) are indicated. The divergence time estimations are shown with 95% HPD. Also, the morphological differences between clades and lineages are shown.

The haplotype networks for *X. variata* showed a general congruence with the phylogenetic reconstruction (Figs. 2 and 3, respectively). For *cytb* gene, the haplotype network recovered six haplogroups. The most divergent haplogroup (HgVI; Clade II) comprise the samples belonging to Cuitzeo, which was separated by 49 mutation steps from haplogroup V (sub-clade 4). The haplogroup V included samples of Zacapu Lake, within the Zacapu biogeographic region.

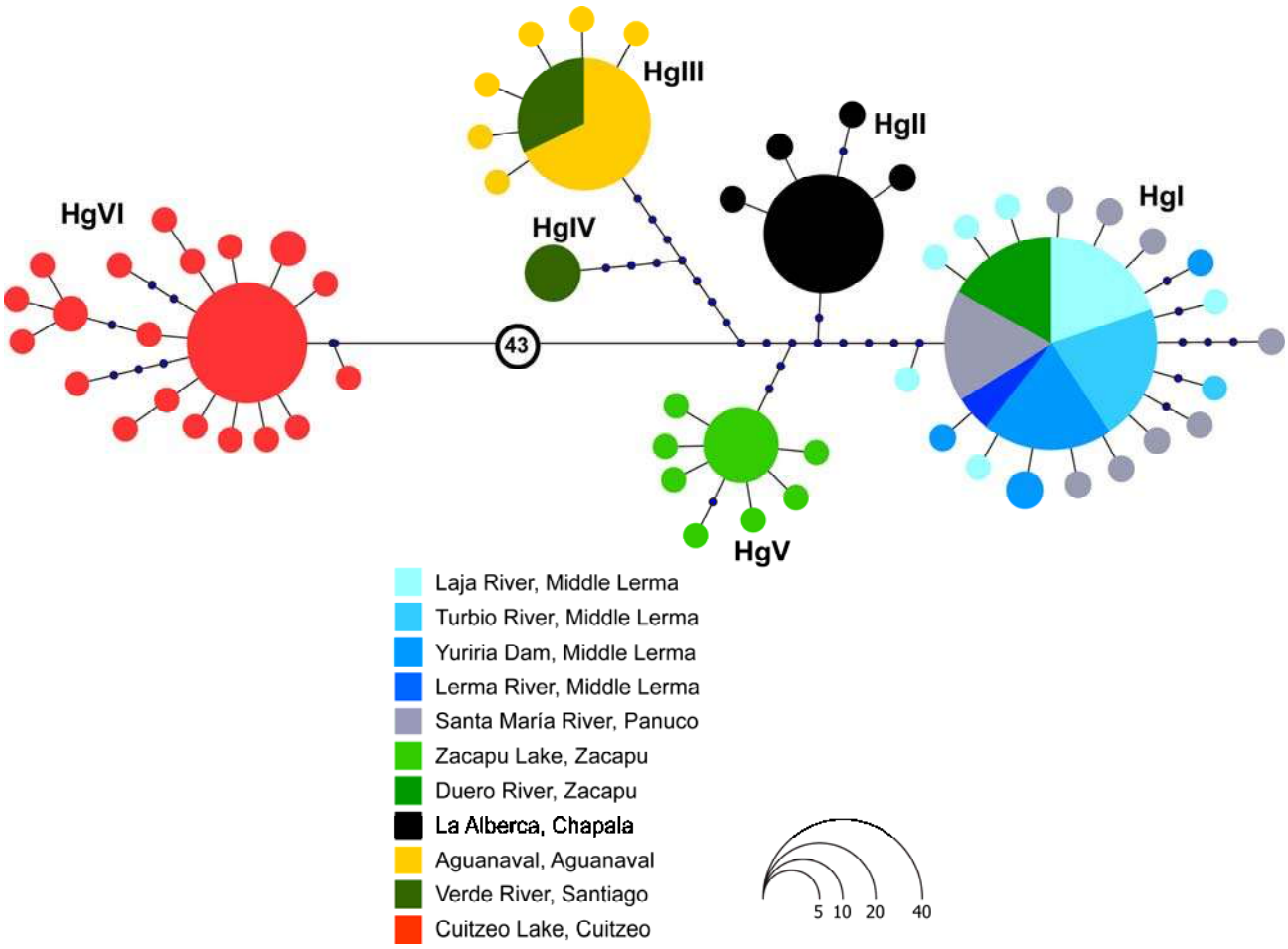


Figure 3 Haplotypes network of *cytb*. Size of the circles indicates the frequency of the haplotype and colors correspond to the discrete biogeographic regions described by Domínguez-Domínguez et al., (2006) were the samples belong as show. Numbers inside circles correspond to mutation steps. The smallest circles correspond to mutation steps.

Haplogroup I (sub-clade 1) was conformed by specimens collected in location along the Middle Lerma, Panuco and Zacapu biogeographic regions (Melchor Ocampo Dam in the Duero river basin). Haplogroup II (sub-clade 2) was conformed by haplotypes from Chapala Lake biogeographic region collected in one location (La Alberca). Haplogroup III (sub-clade 3) is represented by specimens collected in Aguanaval and Verde (Santiago river) biogeographic regions. The HgIV are represented for five specimens from Verde river in a single haplotype, this were separated from the other samples from the same location and Aguanaval samples by seven mutation steps.

Between five and 15 mutation steps separate the HgI to HgV. The most common haplotype corresponded to the HgI, which includes samples of Middle Lerma and Duero River of Zacapu biogeographic region. Moreover, Melchor Ocampo location in Zacapu biogeographic region not share haplotypes with specimens sampled in Zacapu lake, and are separated by eight mutation steps between them (Fig. 3).

For the nuclear genes, haplotypes from Panuco, Aguanaval, Middle Lerma, Verde (Santiago river) and Zacapu (Duero river) biogeographic regions appeared mixed in at least one haplotype. Samples from Zacapu (Zacapu lake) appeared also mixed in *S7* and *RHO* genes, but conformed a single haplotype in *RAG1*. Samples from Chapala appeared mixed in *RAG1* and *RHO* genes. As in the *cytb*, most of the samples from Cuitzeo region were grouped together and separated from the rest of the samples for at least one mutation step, only in the *RHO* gene, one specimen was grouped with samples from the other biogeographic region. For *S7*, one deletion with five exclusive gaps were found in the Cuitzeo samples (Fig. 4).

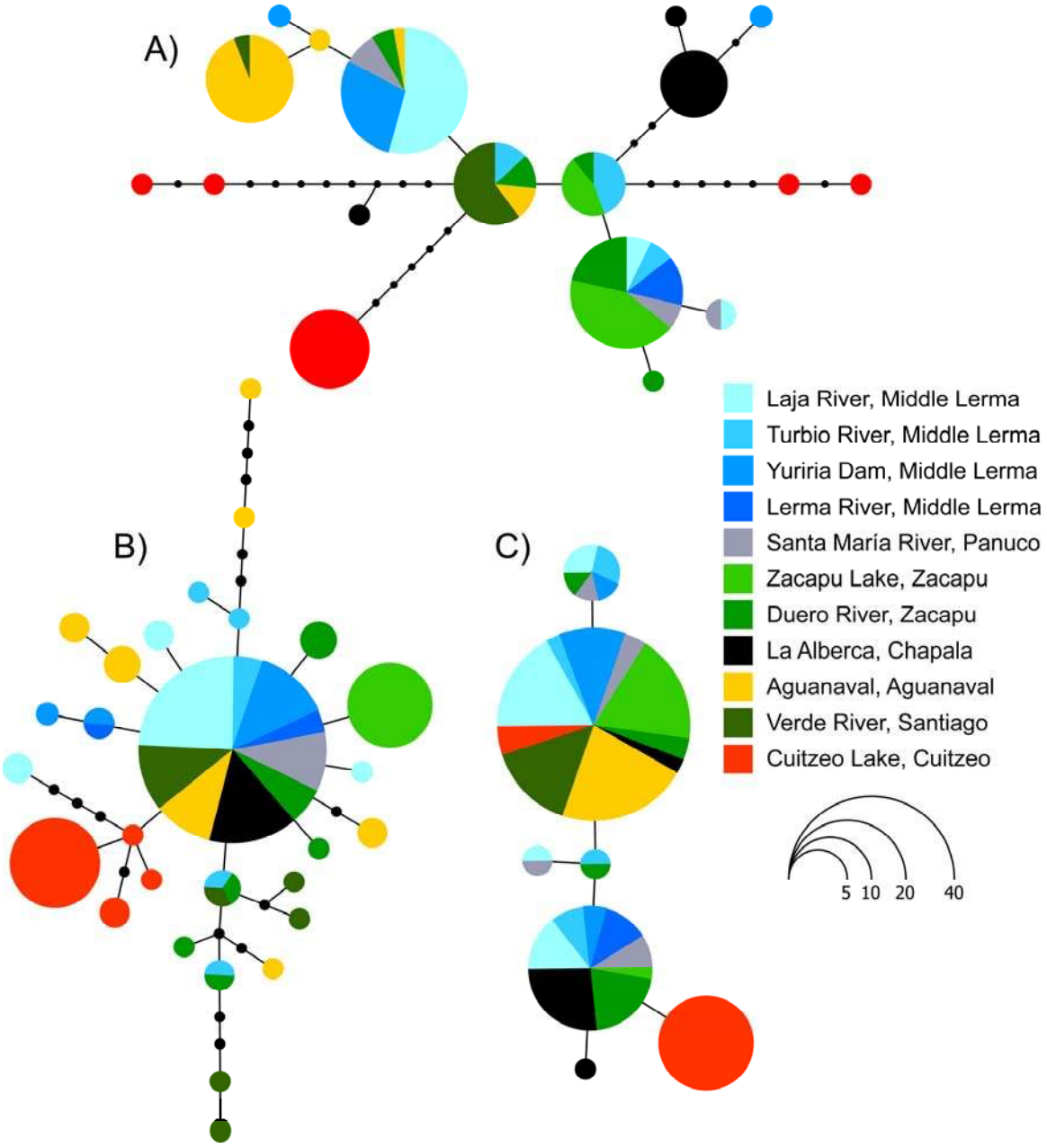


Figure 4 Haplotypes network, A) with *S7* gene, B) with *RAG1* gene, C) with *RHO* gene. Size of the circles indicates the frequency of the haplotype and colors correspond to the discrete biogeographic regions described by Domínguez-Domínguez et al., (2006) were the samples belong as show. The smallest and black circles correspond to mutation steps.

Divergence time estimation and genetic distances

The mutation rate-calibration tree suggested that the first isolation event in *X. variata*, that separated clades I and II, probably occurred near the middle Pliocene and middle Pleistocene *ca*. 2.87 Mya (95% HPD: 1.44-4.31 Mya).

Whereas, the separation between sub-clades within clade I was estimated to have occurred during the Pleistocene period *ca*. 0.91 Mya (95% HPD: 0.4-1.42 Mya) (Fig. 2).

The genetic distance between the two highly divergent clades (Cuitzeo *vs*. the rest of biogeographic regions) was of 5.0% with *cytb*, while for the nuclear genes were of 0.2% for *S7* and *RAG1*, and 0.3% for *RHO*. When the grouping was according with recovered sub-clades within clade I, the highest genetic distance with *cytb* was of sub-clade III (Aguanaval and Verde rivers (this last belonging to Santiago river) *vs*. clade II (Cuitzeo lake) 5.3%, while for the nuclear genes the genetic distance was from 0.1% to 0.4%. Finally for the third grouping according with biogeographic region, for *cytb* the highest genetic distance was between Aguanaval river and Cuitzeo lake (5.4%), while for the nuclear genes the maximum genetic distance was between biogeographic regions clustered in clade I for *S7* and *RAG1*, for the *RHO* the highest genetic distance was between Cuitzeo (clade II), and the rest of biogeographic regions (Table 2).

Table 2 Uncorrected *p* genetic distances in percentage within (bold numbers) and according with the same grouping that the used in genetic structure and genetic diversity for the four genes for *Xenotoca variata*.

		Cytb		
Testing assumptions	Source of variation	% of variance	Fixation index	*P*- value
Groups	Among groups	19.09	Φ_{CT}: 0.19	Ns
According with	Among populations within groups	51.19	Φ_{SC}: 0.63	<0.0001
biogeographical regions	Within populations	29.71	Φ_{ST}: 0.70	<0.0001
	Total	100		
According with the two	Among groups	38.77	Φ_{CT}: 0.38	<0.0001
recovered haplogroups	Among populations within groups	33.81	Φ_{SC}: 0.55	<0.0001
West and East	Within populations	27.42	Φ_{ST}: 0.72	<0.0001
	Total	100		
According with the two	Among groups	42.18	Φ_{CT}: 0.42	<0.0001
recovered haplogroups	Among populations within groups	30.46	Φ_{SC}: 0.52	<0.0001
West and East + Armeria	Within populations	27.35	Φ_{ST}: 0.72	<0.0001

River	Total	100		
No grouping *a priori*	Among populations	67.84	Φ_{ST}: 0.68	<0.0001
	Within populations	32.16		
	Total	100		

Genetic structure among populations

The pairwise Φ_{ST} with *cytb* and *S7* when compared clade I *vs*. clade II shown high and significant genetic differentiation (>0.75), while for *RAG1* and *RHO* although the differences are not very high, were significant (Table 3).

When the arrangement was according with the mtDNA + nDNA phylogeny the pairwise Φ_{ST} were high, although not all were significant. For nuclear genes for this arrangement, *S7* showed the highest genetic differentiation (Table 3). Finally, when comparisons were performed according with biogeographic regions, for all genes, all comparisons of Cuitzeo respect to the other biogeographic regions were highest, but not all were significant (Table 3).

Table 3 Genetic differentiation using pairwise Φ_{ST} for *cytb, S7*, *RAG1* and *RHO*, 1) according with recovered lineages, 2) according with clades within lineages, and, 3) according with biogeographic regions. The comparisons significant after Bonferroni correction ($p < 0.05$) are in bold.

cytb	West haplogroup	East haplogroup	Armeria River
West haplogroup	0.003		
East haplogroup	0.005	0.002	
Armeria River	0.005	0.007	0.001

AMOVA analysis for *cytb* revealed the most significant genetic structure when grouped according with phylogenetic trees and haplotype networks (sub-clades within clade I and clade II) (Φ_{ST} = 0.96, Φ_{SC} 0.12 and Φ_{CT} 0.95, $p < 0.001$) (Table 4). The AMOVAs when the arrangements were according with

sub-clades and clades, and for biogeographic regions, recovered a high genetic structure among groups, however, were not significant (Table 4).

For the nuclear genes, the *S7*, although showed high genetic structure when groups were set according to clades and biogeographic regions, but were not significant, while for *RAG1* and *RHO*, the highest genetic structure was found when samples were grouping according with the two clades, but were not significant (Table 4).

Table 4 Analyses of molecular variance for groups according to: 1) the recovered lineages; 2) clades and lineages recovered in phylogenetic analyses (within lineage I: Clade 1 = Middle Lerma, Panuco, Zacapu (Duero River); Clade 2 = Chapala Lake biogeographic region; Clade 3 = Aguanaval and Verde river (Santiago); Clade 4 = Cuitzeo Lake; and, lineage II = Cuitzeo biogeographic region); and 3) according to biogeographic regions proposed by Domínguez-Domínguez et al., (2010), for the four genes. Ns = Not significant.

		Cytb		
Testing assumptions	Source of variation	% of variance	Fixation index	*P*- value
1) Lineage I and lineage II	Among groups	83.24	Φ_{CT}: 0.83	Ns
	Among populations within groups	14.64	Φ_{SC}: 0.87	<0.001
	Within populations	2.12	Φ_{ST}: 0.97	<0.001
	Total	100		
2) Clades within lineages	Among groups	95.73	Φ_{CT}: 0.95	<0.001
	Among populations within groups	0.55	Φ_{SC}: 0.12	Ns
	Within populations	3.72	Φ_{ST}: 0.96	<0.001
	Total	100		
3) Biogeographic regions	Among groups	89.01	Φ_{CT}: 0.89	Ns
proposed by Domínguez-	Among populations within groups	6.78	Φ_{SC}: 0.61	<0.001
Domínguez et al., 2010.	Within populations	4.21	Φ_{ST}: 0.95	<0.001
	Total	100		
		S7		
1) Lineage I and lineage II	Among groups	69.20	Φ_{CT}: 0.67	Ns
	Among populations within groups	21.89	Φ_{SC}: 0.50	<0.001
	Within populations	8.91	Φ_{ST}: 0.84	<0.001
	Total	100		
2) Clades within lineages	Among groups	67.78	Φ_{CT}: 0.39	Ns
	Among populations within groups	16.32	Φ_{SC}: 0.01	<0.001

	Within populations	15.90	Φ_{ST}: 0.40	<0.001
	Total	100		
3) Biogeographic regions proposed by Domínguez-Domínguez et al., 2010.	Among groups	64.90	Φ_{CT}: 0.64	Ns
	Among populations within groups	17.84	Φ_{SC}: 0.50	<0.001
	Within populations	17.26	Φ_{ST}: 0.82	<0.001
	Total	100		
RAG1				
1) Lineage I and lineage II	Among groups	48.90	Φ_{CT}: 0.48	Ns
	Among populations within groups	13.28	Φ_{SC}: 0.25	<0.001
	Within populations	37.82	Φ_{ST}: 0.62	<0.001
	Total	100		
2) Clades within lineages	Among groups	40.99	Φ_{CT}: 0.40	Ns
	Among populations within groups	6.55	Φ_{SC}: 0.11	<0.001
	Within populations	52.47	Φ_{ST}: 0.47	<0.001
	Total	100		
3) Biogeographic regions proposed by Domínguez-Domínguez et al., 2010.	Among groups	23.53	Φ_{CT}: 0.23	Ns
	Among populations within groups	20.45	Φ_{SC}: 0.26	<0.001
	Within populations	56.02	Φ_{ST}: 0.43	<0.001
	Total	100		
RHO				
1) Lineage I and lineage II	Among groups	49.76	Φ_{CT}: 0.49	Ns
	Among populations within groups	16.62	Φ_{SC}: 0.33	<0.001
	Within populations	33.62	Φ_{ST}: 0.66	<0.001
	Total	100		
2) Clades within lineages	Among groups	43.75	Φ_{CT}: 0.43	Ns
	Among populations within groups	9.30	Φ_{SC}: 0.16	Ns
	Within populations	46.95	Φ_{ST}: 0.53	<0.001
	Total	100		
3) Biogeographic regions proposed by Domínguez-Domínguez et al., 2010.	Among groups	24.42	Φ_{CT}: 0.23	Ns
	Among populations within groups	25.32	Φ_{SC}: 0.33	<0.001
	Within populations	50.26	Φ_{ST}: 0.49	<0.001
	Total	100		

Genetic diversity and historical demography

The genetic diversity for *X. variata* when considered the two recovered clades for the four genes for the clade I were h between 0.529 and 0.824 and π between 0.0009 and 0.007, higher than the obtained for the clade II (h between 0.289 and 0.730; π between 0.0005 and 0.001) (Table 5). When the sub-clades within clade I were considered separately and when grouping according to the biogeographic regions, the genetic diversity for *cytb* was highest for clade II

(Cuitzeo biogeographic region) (Table 6). For the nuclear genes when the arrangement was according with sub-clades and clades the highest genetic diversity was for sub-clade 1 with *S7* and *RHO* and for sub-clade 3 with the *RAG1* (Table 6). Finally when the grouping was according to the biogeographic regions for nuclear genes the highest genetic diversity for *S7* and *RHO* was for Panuco biogeographic region, and for *RAG1* for Aguanaval (Table 7).

Table 5 Genetic diversity for each recovered lineage of *Xenotoca variata* for the four genes. N, sample size, S, polymorphic sites, H, number of haplotypes, *π,* nucleotide diversity *h,* haplotype diversity.

		cytb			
Lineages	**N**	**S**	**H**	***π***	***h***
Lineage I	171	63	40	0.007+/-0.003	0.783+/-0.026
Lineage II	44	26	20	0.001+/-0.001	0.730+/-0.075
		S7			
Lineage I	122	8	13	0.003+/-0.001	0.824+/-0.016
Lineage II	18	4	5	0.001+/-0.0009	0.405+/-0.142
		RAG1			
Lineage I	128	31	24	0.0009+/-0.0006	0.613+/-0.048
Lineage II	22	4	4	0.0005+/-0.0004	0.333+/-0.124
		RHO			
Lineage I	124	5	6	0.001+/-0.0008	0.529+/-0.038
Lineage II	24	3	2	0.001+/-0.0008	0.289+/-0.102

Table 6 Genetic diversity according with the phylogenetic arrangement (including sub-clades within clade I) of *Xenotoca variata* for the four genes. N, sample size, S, polymorphic sites, H, number of haplotypes, *π,* nucleotide diversity *h,* haplotype diversity.

			cytb		
Phylogenetic arrangement	**N**	**S**	**H**	**Π**	***h***
Clade 1 (lineage I)	90	26	19	0.0006+/-0.0005	0.379+/-0.066
Clade 2 (lineage I)	26	5	5	0.0003+/-0.0004	0.289+/-0.114
Clade 3 (lineage I)	39	13	8	0.002+/-0.001	0.476+/-0.094
Clade IV (lineage I)	16	8	8	0.001+/-0.0007	0.700+/-0.127
Lineage II	44	26	20	0.001+/-0.001	0.730+/-0.75
			S7		
Clade 1 (lineage I)	64	5	8	0.002+/-0.001	0.653+/-0.050
Clade 2 (lineage I)	12	3	3	0.0006+/-0.0007	0.318+/-0.163
Clade 3 (lineage I)	30	2	4	0.0008+/-0.0007	0.560+/-0.058
Clade IV (lineage I)	16	1	2	0.0005+/-0.0006	0.400+/-0.113
Lineage II	18	4	5	0.001+/-0.0009	0.405+/-0.142
			RAG1		
Clade 1 (lineage I)	68	16	13	0.0006+/-0.0005	0.480+/-0.075
Clade 2 (lineage I)	12	0	1	0.000+/-0.000	0.000+/-0.000
Clade 3 (lineage I)	32	19	12	0.001+/-0.001	0.715+/-0.085
Clade IV (lineage I)	16	0	1	0.000+/-0.000	0.000+/-0.000
Lineage II	22	4	4	0.0005+/-0.0004	0.333+/-0.124
			RHO		
Clade 1 (lineage I)	66	4	5	0.001+/-0.001	0.643+/-0.035
Clade 2 (lineage I)	12	3	3	0.0009+/-0.0008	0.439+/-0.158
Clade 3 (lineage I)	30	0	1	0.000+/-0.000	0.000+/-0.000
Clade IV (lineage I)	16	2	2	0.0002+/-0.0003	0.125+/-0.106
Lineage II	24	3	2	0.001+/-0.0008	0.289+/-0.102

Table 7 Genetic diversity according with the biogeographic regions proposed by Domínguez-Domínguez et al., (2010) of *Xenotoca variata* for the four genes. N, sample size, S, polymorphic sites, H, number of haplotypes, *π,* nucleotide diversity *h,* haplotype diversity.

			cytb		
Phylogenetic arrangement	**N**	**S**	**H**	**Π**	***h***
Middle Lerma	58	14	11	0.0005+/-0.0004	0.345+/-0.081
Panuco	20	12	9	0.001+/-0.0008	0.652+/-0.122
Zacapu	28	15	9	0.004+/-0.002	0.730+/-0.062
Chapala	26	5	5	0.0003+/-0.0004	0.289+/-0.114
Aguanaval	25	6	7	0.0004+/-0.0004	0.430+/-0.123
Verde	14	8	2	0.003+/-0.002	0.494+/-0.087
Cuitzeo	44	26	20	0.001+/-0.001	0.730+/-0.075
			S7		
Middle Lerma	46	5	7	0.001+/-0.001	0.573+/-0.075
Panuco	6	3	3	0.002+/-0.001	0.733+/-0.155
Zacapu	28	3	5	0.001+/-0.0009	0.563+/-0.097
Chapala	12	3	3	0.0006+/-0.0007	0.318+/-0.163
Aguanaval	20	2	4	0.0006+/-0.0006	0.363+/-0.130
Verde	10	1	2	0.0002+/-0.0004	0.200+/-0.154
Cuitzeo	18	4	5	0.001+/-0.0009	0.405+/-0.142
			RAG1		
Middle Lerma	48	14	10	0.0006+/-0.0005	0.438+/-0.090
Panuco	8	0	1	0.000+/-0.000	0.000+/-0.000
Zacapu	28	6	7	0.0008+/-0.0006	0.648+/-0.089
Chapala	12	0	1	0.000+/-0.000	0.000+/-0.000
Aguanaval	18	13	7	0.001+/-0.001	0.784+/-0.084
Verde	14	8	6	0.001+/-0.0009	0.604+/-0.149
Cuitzeo	22	4	4	0.0005+/-0.0004	0.333+/-0.124
			RHO		
Middle Lerma	46	4	5	0.001+/-0.0009	0.612+/-0.052
Panuco	8	4	4	0.001+/-0.001	0.785+/-0.112
Zacapu	28	3	4	0.001+/-0.0008	0.521+/-0.079
Chapala	12	3	3	0.0009+/-0.0008	0.439+/-0.158
Aguanaval	18	0	1	0.000+/-0.000	0.00+/-0.000
Verde	12	0	1	0.000+/-0.000	0.00+/-0.000
Cuitzeo	24	3	2	0.001+/-0.0008	0.289+/-0.102

The BSP analyses for *cytb* for clades within clade I showed the same pattern of a slightly demographic decline followed of a population expansion in the last 0.04 Mya for sub-clade 1 (Middle Lerma, Duero River of Zacapu biogeographic region and Panuco), sub-clade 2 (Chapala Lake) and sub-clade 4 (Zacapu Lake). For sub-clade 3 (Aguanaval and Verde Rivers) a demographic decline was detected, and, although a recent population expansion is showed (<0.04 Mya) at the same time that the others sub-clades, this is out of the confidence intervals. For lineage II (Cuitzeo biogeographic region) a stable population size was detected until the last 0.200 Mya, in which a population expansion was detected (Fig. 5).

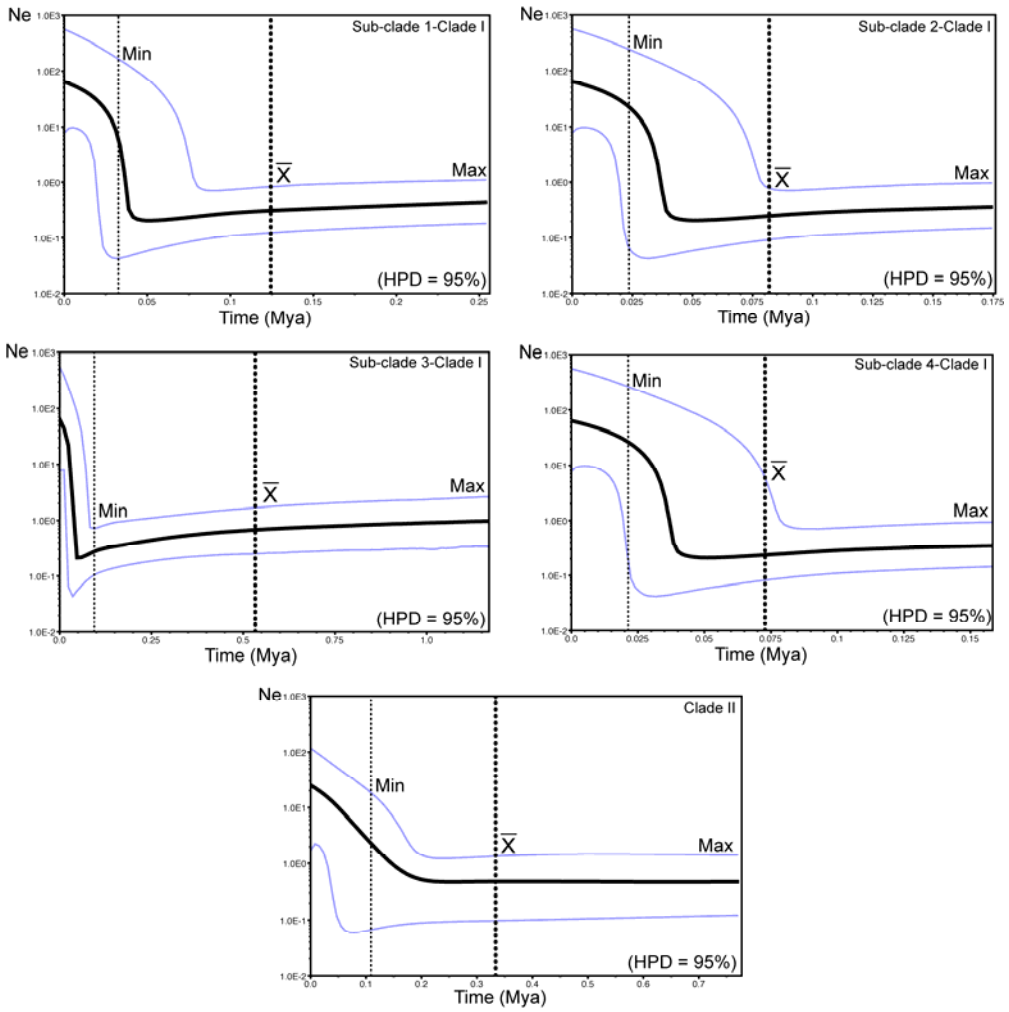


Figure 5 Demographic history of populations of each recovered clade within lineage I and the lineage II using Bayesian Skyline Plot (BSP) from *cytb* sequences. Dotted lines represent the location of the upper bound (Max), the mean (X) and lower bound (Min) of the HPD = 95%.

Discussion

Freshwater fishes are usually restricted to water bodies, which are surrounded by biogeographic barriers. As a result, it is possible to invoke an island-like biogeographic model to explain the possible exchange of biotic components among “islands” (water bodies). This exchange is only possible when a physical connection occurred between them (e. g. river capture, water bodies connection). In this way, it has been argued that these connection events, mainly induced by geologic forces or climatic changes, influenced in the same way the exchange of fish fauna along the new dispersion route (Bermingham and Martin, 1998).

The evolutionary history of the freshwater ecosystems of central Mexico has been addressed by studying different taxa, such as goodeids (Beltrán-López et al., 2017; Doadrio and Domínguez, 2004; Domínguez-Domínguez et al., 2006; 2008; 2010; 2016; Piller et al., 2015), cyprinids (Pérez-Rodríguez et al., 2009) poeciliids (Beltrán-López et al., 2018; Mateos et al., 2002) catostomids (Pérez-Rodríguez et al., 2016), snakes (Bryson et al., 2011), crayfish (Pedraza-Lara et al., 2012), ambystomatid salamanders (Parra-Olea et al., 2012), helminth parasites (Mejía-Madrid et al., 2007) and even by considering a combination of different taxa (Huidobro et al., 2006; Pérez-Rodríguez et al., 2015).

These studies have revealed several vicariant events that have produced congruent biogeographic histories shared by different co-distributed taxa, in which the main explanation for phylogeographic patterns lies in geological or climatic hypotheses, also, the results showed in the present work within clade I of *Xenotoca variata* support previous vicariance-dispersal hypothesis results, that lies in the geologic and climatic explanation. But in contrast to previous results, the analyses including all the genes surveyed and all the molecular analyses conducted, show the Cuitzeo populations as a well differentiated genetic clade with the highest genetic divergence found for the comparation of population of fish distributed in the Cuitzeo and surrounding drainage (*p*-uncorrected distances = or > to 5.0%; Table 2) (Beltrán-López et al., 2018; *en preparación*; Domínguez-Domínguez et al., 2008; Ornelas-García et al., 2012; Pérez-Rodríguez et al., 2009).

In this scenario, although Cuitzeo and Zacapu-Lerma are now independent drainages, several geological and fish biogeographic studies have been argued that this drainages suffer several events of connection and disconnection historically (Israde-Alcantara, 1999; Moncayo-Estrada et al., 2001), this pattern has been corroborated in studies that involve fish species distributed in Zacapu, Cuitzeo and Lerma basin, frequently showing shared or closely related haplotypes between them (Beltrán-López et al., 2018; Domínguez-Domínguez et al., 2008; 2010; Pérez-Rodríguez et al., 2009; 2016), explained as a recent connection and interchange of individuals with genetic homogenization (Domínguez-Domínguez et al., 2009; 2010), connection that is highly probable to use by species with high dispersed capabilities, as is expected in widely distributed species, as *X. variata,* that is found in seven drainages along central Mexico.

Early divergences in two main clades

Xenotoca variata is a genetically structured species. In this case, clade I comprised specimens widely distributed in six biogeographic areas; Aguanaval, Verde, Chapala, Middle Lerma, Zacapu and Panuco (*sensu* Domínguez-Domínguez et al., 2006), most of them actually separated drainages and with distances ~ 470 km between width localities. Clade II is composed by specimens collected in three locations in the Cuitzeo biogeographic region, which represents a contiguous basin with respect to Zacapu and Middle Lerma populations (~ 30 km of distance) that belong within clade I (Figs. 1, 2 and 3).

The high genetic distances in mtDNA (4.7 to 5.3%) between both recovered clades, the segregation of samples from Cuitzeo in an independent haplogroup for nDNA, the high and significant Φ_{ST} values for all the genes when compared the two clades, the high genetic structure was according with sub-clades and clades, and the divergence time between both clades (*ca*. 2.87 Mya) support the two clades as two independent evolutionary entities (see Figs. 2, 3 and 4; Tables 2, 3 and 4).

Moreover, the genetic distances in the *cytb* gene is higher than the lower genetic distances observed in other goodeido species pairs, which have interspecific genetic distances within the same genera of 0.5 to 11% (Beltrán-

López et al., 2017; Doadrio & Domínguez, 2004; Piller et al., 2015), the minimum value proposed as the limit for sibling species of fishes with the *cytb* marker (Avise, 2000; Bradley & Baker, 2001) and even largely exceeds the 2% sequence divergence generally accepted as a cut-off value between sister species of vertebrates (Avise, 1998).

The isolation of the two clades is in agreement with previous results showing the Cuitzeo lake as an important area for the early radiation of fish species, in which volcanic, tectonic and occasional paleoclimatic factors could be involved (Domínguez-Domínguez et al., 2010).

Ancient connection of the Cuitzeo paleolake and regions of the Lerma-Zacapu River basin have been documented to occur during the early Miocene and Pliocene (Domínguez-Domínguez et al., 2010). The divergence time between the two clades seems to be related with the geological activity during Pliocene-Quaternary that have placed in this part of the country due to faults around Cuitzeo lake, as Morelia fault zone and the Penjamillo graben, which were active as extensional systems in Plio-Quaternary times, and which locally reactivate older faults (Ferrari et al., 1994). Also, this divergent event coincides with the formation of monogenetic volcanic fields. The most prominent is the Michoacán-Guanajuato volcanic field (MGVF) with *ca*. 1000 volcanic centers distributed over 40,000 Km^2 in the central Trans Mexican Volcanic Belt (TMVB) sector (Hasenaka, 1994; Hasenaka & Carmichael, 1985). Vulcanism in the MGVF started in the late Pliocene (2.78 Mya; Hasenaka & Carmichael, 1985), this date coincides with the divergence of the two lineages of *X. variata*, that could interrupted the connection between Cuitzeo lake and Middle Lerma and Zacapu biogeographic regions, added to this, faults in the northwestern part of the system (Cuitzeo lake) affect Miocene to Pliocene volcanic sequences (Garduño-Monroy et al., 2001), which could be related with the divergence of the two lineages found herein (Figs. 2 , 3 and 6).

Additionally, it has been demonstrated that during the late Miocene and early Pliocene, Cuitzeo Lake had a much larger extension. It was a deeper lake, and it experienced a strong level decrease during the Pliocene as a result of the geological faults in the region (Israde-Alcántara & Garduño-Monroy, 1999). This is in agreement with the faunal exchange and dispersal and subsequent isolation of the two lineages of *X. variata* during Pliocene (Fig. 2).

Onces the *X. variata* was isolated in Cuitzeo, its differentiation starts, but the maintainence of this differentiation through time is no in agreement with other studies using molecular data for phylogenetic reconstructions in goodeids (Doadrio and Domínguez, 2004; Domínguez-Domínguez et al., 2008) cyprinids (Domínguez-Domínguez et al., 2007b; Pérez-Rodríguez et al., 2009; Schönhuth et al., 2008), and poeciliids (Beltrán-López et al., 2018), in which a close relationship was found among the populations of different co-distributed species in the Cuitzeo and Zacapu-Lerma River basins, which supports the hypothesis of recent faunal exchange among these areas that are currently isolated. The recent connection between Cuitzeo and contiguous basins are also in agreement with geological and paleoenvironmental data (diatomits distribution) that support the existence of a connection between Zacapu and Cuitzeo lakes that is promoted by the activity of the northeast-southwest fault system of the area, disrupted by the activity of the Ventanas Volcano *ca.* 0.7 to 0.5 Mya (Israde-Alcántara, 1999; Israde-Alcántara & Garduño-Monroy, 1999; Israde-Alcántara et al., 2008).

The historical connection among these areas has also been suggested by other authors who based their conclusions on the description of distribution patterns of the freshwater fish fauna (Domínguez-Domínguez et al., 2010; Ornelas-García et al., 2012). Moreover, populations of *Zoogoneticus quitzeoensis* (Domínguez-Domínguez et al., 2008), *Alloophorus robustus* (Domínguez-Domínguez et al., 2008), *Poeciliopsis infans* and *Goodea atripinnis* (Beltrán-López et al., 2018; *en preparación*) share haplotypes with the Cuitzeo and Zacapu-Lerma areas. Considering all this information together, there is no doubt that a recent connection occurred between these areas with a concomitant fish exchange that according with our data must to have included *X. variata,* which is the second most widespread species of goodeine, which is indicative of high dispersal capability.

In contrast, the populations of *X. variata* did not share haplotypes among these areas (Cuitzeo and Zacapu-Lerma), moreover show a high and significant genetic differentiation (Tables 2, 3 and 4) results for all the genes and high genetic distances with *cytb* when Cuitzeo population is compared with Zacapu (4.7%) demonstrates that the two clades of *X. variata* have a long history of reproductive isolation (Table 2). Furthermore, this divergence value is high in

comparison to general results reported for *cytb* genetic differentiation between species of vertebrates and even for species within the Goodeinae (Beltrán-López et al., 2017; Doadrio and Domínguez, 2004; Domínguez-Domínguez et al., 2010).

Nevertheless, following Ritchie et al., (2005), the information we provide herein allows us to propose that a vicariant event was more important than sexual selection in the early radiation of *X. variata,* resulting in the splitting of the ancestral population of the two divergent lineages of *X. variata*. Additionally, because faunal exchange has been commonly observed between the Cuitzeo and Zacapu areas in recent geological times (less than 1 Ma), we suggest that contact between the two clades after the isolation event is likely to have occurred, but the occurrence of the two divergent clades in sympatry could have been prevented by competition (ecological equivalence) and sexual selection. A plausible explanation for this ancient divergence between populations of Cuitzeo *vs*. Zacapu - Middle Lerma, that have been in continuous cycles of connection and isolations until recently, could be due to the sexual dimorphism. In previous works it has been shown, that the most sexually dimorphic species have greater genetic differentiation between populations, as is the case of *Xenotoca melanosoma*, and suggested that males are moving between populations of the dimorphic species, but reproducing less successfully. This supports the mechanism by which sexual isolation can influence speciation. If male traits have diverged more, and females are more choosy, the effective gene flow would be reduced (Ritchie et al., 2007), which could be why Cuitzeo maintains its divergence; although Cuitzeo has been connected to Zacapu and Middle Lerma recently, due to *X. variata* represents a dimorphic species, the same manner that *X. melanosoma*.

Although this is just one possible explanation and it needs further corroboration, recent studies have been demonstrate that under controlled conditions with males and females adults of two populations of *X. variata* (Cuitzeo and Zacapu), showed that the selection of males by females responses to different factors for each one of the worked populations as could be the body and fin size of males, also, the courtship in the two populations is different in terms of the number of patterns that they performed for each event, the frequency with which they performed these patterns and the time it takes to

performed them, also, in this work, they considering that the sexual selection is the mechanism that have been separated the two genetic groups of *X. variata* (Villa-Villaseñor, 2013). Moreover, it has been found that other species of freshwater fishes, including goodeines, exhibit equivalent species or lineages that do not overlap in their distributional ranges even though they inhabit the same hydrological system after secondary contact. Such a phenomenon is seen in *Zoogoneticus quitzeoensis* (Domínguez-Domínguez et al., 2008), *Notropis calientis-N. marhabatiensis-N. grandis* (Domínguez-Domínguez et al., 2009), and the species pairs *Skiffia multipunctata-S. lermae* and *X. variata-Chapalichthys encaustus* (Domínguez-Domínguez et al., 2010).

Sub-clades within Clade I

In the case of clade I, the phylogenetic trees, haplotype networks results, high and significative Φ_{ST} values for all pairwise comparisons (>0.9 and p<0.001) and the high genetic distances with *cytb* (0.7% to 1.4%) support the existence of four highly supported and well-structured sub-clades for *cytb* database. The nuclear genes show contradictory results, showing somewhat structure and significative Φ_{ST} values depending on the gene tested (Fig. 4; Tables 2 and 3), differences that can be attributed to the low mutation rate for the nuclear genes, and that the time since isolation is not enough to segregate, giving an incomplete lineage sorting for some population in a certain nDNA genes.

Sub-clades within clade I diversified during the Pleistocene, between 1.42 and 0.4 Mya (*ca*. 0.91 Mya), which indicates that all clades originated at almost the same period, probable as a response to the same event around 1 Mya. This period is in accordance with the estimated time when goodeids experienced their highest diversification during a dry period that caused isolation by basin fragmentation in central Mexico (Domínguez-Domínguez et al., 2008) and that also corresponds to diversification patterns in other clades of freshwater fishes such as Cyprinids (Domínguez-Domínguez et al., 2007b; Pérez-Rodríguez et al., 2009), Characids (Ornelas-García et al., 2008; Strecker et al., 2004), and Poecilids (Beltrán-López et al., 2018; Mateos et al., 2002). The AMOVA results also support this interpretation (Table 4); the test for

phylogenetic arrangements maximizes the percentage of the variance explained between groups, which indicate that the actual configuration of the drainages (biogeographic arrangement) does not explain the divergence levels in this clade (Table 4).

The results for the *cytb* network indicates that the Zacapu Lake population (HgV) is the most close related group with respect to Clade I. The genetic distinctiveness of Zacapu with respect to the rest of the populations is not surprising, although Zacapu lake is connected with Lerma River trough Angulo river, the aquatic fauna from Zacapu is highly characteristic, with several endemic species, as *Notropis grandis* (Domínguez-Domínguez et al., 2009), *Allotoca zacapuensis* (Meyer et al., 2001), *Ambystoma andersoni* (Krebs & Brandon, 1984) and also highly divergent population as in *Cambarellus* (Pedraza-Lara et al., 2012) and several species of goodeines which show closer relationship with respect to Cuitzeo populations than Lerma-Chapala populations (Domínguez-Domínguez et al., 2010). Apparently, during this period, the activity of the so-called Tarascan Corridor was also intense (Israde-Alcántara, 1999), which could have promoted the isolation of Zacapu population through the formation of a biogeographical barrier between the Zacapu and Lerma River basins.

Some authors argued the possibility that the Zacapu paleolake was endhorreic or at least with a very limited superficial outflow during unusually rainy years (Arnauld et al., 2014; Correa-Metrio et al., 2012), that could be the reason of the particular aquatic fauna in the area, since not all the species could follow the same dispersion route do to differences in ecological or biological traits (Beltrán-López et al., 2018; en preparación; Betancourt-Resendes et al., 2018).

Moreover, other barrier that could isolate the Zacapu population is represented by the waterfall of approximately 6 m in height that is documented to have been formed in Angulo River, at the mid-way point between Zacapu Lake and Lerma River, although there are no data on the time when this waterfall was formed.

The haplogroup I comprise samples from three different biogeographic regions (Middle Lerma, Panuco and Angulo River in the Zacapu region). The close relationship between Panuco and Lerma river faunal populations have

been largely discussed for goodeines Domínguez-Domínguez et al., 2010), poecilids (Mateos et al., 2002), cyprinids (Domínguez-Domínguez et al., 2007b; Pérez-Rodríguez et al., 2009) and crayfish (Pedraza-Lara et al., 2012), and is explained by river piracy events during recent geological times (less than 1 Mya). West (1964) suggested that the San Juan del Río River was part of the Lerma system before its capture by the Panuco River during the Pleistocene, which could explained the recurrent pattern in several freshwater organism in which Panuco and Lerma shared haplotypes, as is the case of *X. variata* (Figs. 2, 3 and 4).

The population from Melchor Ocampo Dam (Zacapu biogeographic region) that belong with the Lerma-Panuco samples is explained because this dam is located in the lower part of the Duero River, and more influenced by Lerma fauna than the isolated Zacapu lake, so the aquatic fauna from this location is more closely related with Lerma river than with Zacapu lake (Fig 1), in this case the limits of the Zacapu biogeographic region proposed by Domínguez-Domínguez et al. (2006) must to be considered.

Chapala population (sub-clade 2/HgII) is the group with the low genetic diversity (Tables 6 and 7). La Alberca population belong in to the Chapala biogeographic region, but inhabit an isolated lake of 35 hectares that was part of the Chapala lake until 1900, when the Chapala lake water surface was reduce drastically (Sandoval, 1981). This reduction of water surface and the isolation of the population in a reduced isolate lake could be the cause of the ancient demographic decline and the low genetic diversity found within this haplogroup, although a recent population expansion was detected (Fig. 5).

Verde River samples (HgIV) were collected in a tributary of the Santiago river which ichthyofaunal have been largely argued to be more close related with Chapala-Lerma river than Santiago river (Domínguez-Domínguez et al., 2010). Surprisingly, the samples from this location belong to two differentiated clades, five specimens conform its own haplogroup (HgIV) whereas the rest of samples conform one haplotype that is shared by the most common haplotype from Aguanaval drainage (HgIII). *Xenotoca variata* and *Goodea atripinnis* are the only species of fish inhabiting the Aguanaval drainage which origin is linked with central Mexico drainage, the other approximately ten fish species, except *Astyanax mexicanus,* in the Aguanaval river are related with north basin

species, so the divergences present in population of *X. variata* in this norther basin is expected, but unexpected is the presence of haplotypes distributed in the Santiago basin.

Demographic history

All sub-clades within clade I, except Santiago group (HgIII) exhibit short external branches in the phylogeny and a star-like network (Figs. 2 and 3) with the most frequent haplotype in a central position and surrounded for a low frequency haplotypes. But the low genetic diversity in all the groups (Tables 5, 6 and 7), except Zacapu (clade 4/HgV) may indicate of significant bottlenecks before expansion, which partially corresponded with the BSP analyses that showed for the four clades, a recent population expansion, after of demographic decline (Fig. 5).

Conservation Implications

The freshwater ecosystems in central Mexico have been significantly affected by habitat degradation, the introduction of alien species, desiccation of water bodies, drainage deforestation, overharvesting and water diversion (De la Vega-Salazar, 2006; Domínguez-Domínguez et al., 2005). Thus, some species have drastically reduced populations, and in some cases, they have been now extirpated from certain areas (De la Vega-Salazar, 2006).

In general, the genetic diversity in *X. variata* is lower than *Zoogoneticus quitzeoensis* (Domínguez-Domínguez et al., 2007a; 2008) and *Neotoca bilineata* (Ornelas-Garcia et al., 2012), this is unexpected since *X. variata* is more abundant than *Z. quitzeoensis and N. bilineata* in all the sites were are co-distributed, as is expected for a more generalist species (De la Vega-Salazar, 2006; Domínguez-Domínguez et al., 2005; Ornelas-Garcia et al., 2012).

The data provided herein support conservation management decisions by providing an analysis of the pattern of geographical distribution of genetic variability within populations and the differentiation among populations. The mtDNA data presented herein suggest significant geographical structure of genetic variation across the range of *X. variata*, and our phylogenetic and

phylogeographic analyses point to a long-term historical isolation among the seven geographical areas where this species occurs. In this case, five main operational conservation units (OCUs) (Doadrio et al., 1996) should be considered for *X. variata*: Zacapu, Lerma-Panuco, Chapala, Aguanaval and Cuitzeo. Therefore, the management of central Mexico freshwater ecosystems requires a great amount of information on the organisms occurring there before decisions are made. For instance, data provided in this study show that *X. variata* exhibits strong differentiation between basins and that conservation efforts should carried out that consider populations in each drainage.

Conclusions

This study demonstrates that the dynamics of the genesis and destruction of central Mexico river drainages induced by the high tectonic and volcanic activity since the Miocene, the climatic change events, and also, the ecological segregation joined to sexual selection present in *X. variata* because it is a high dimorphic species have been determining factors of the complex phylogeographic and biogeographic history of the two highly divergent clades recovered in the present work, especially these ecological and biological characteristics are decisive for genetic isolation of clade II (Cuitzeo) respect to the rest of biogeographic regions, despite to the recent connection and exchange of organism in several species distributed there.

Authors' contributions

ODD, RGBL, GPPDL, ID conceived the ideas, ODD, conducted the fieldwork and collected the samples, ODD, RGBL, obtained the molecular data, ODD, RGBL analyzed the data and ODD, RGBL, KRP, SP, GPPDL, ID write the manuscript. All authors have read and approved the final manuscript.

Acknowledgements

We thank the following people for their support during the fieldwork: I. Dibble, R. Pérez, L. Escalera, A. Pompa, H. Mejía, Nayeli Gurrola and R. Rosas. L.

Alcaraz assisted in the laboratory work. This study was partially supported by the Chester Zoo UK and Cientific Research Council of the UMSNH project 8.33. ODD thanks the Consejo Nacional de Ciencia y Tecnología (CONACYT) for providing a scholarship. RGBL thanks the Consejo Nacional de Ciencia y Tecnología (CONACYT: scholarship number 329883) for providing a scholarship Fieldwork was partially funded by the project CGL2006-12325/BOS and the program PAPIIT-UNAM-IN209608 and CONACYT (grant number 47233) to GPPDL.

Figure legends

Figure 1. Sampling locations and discrete biogeographical regions where *Xenotoca variata* is distributed. The numbers of the locations has correspondence with the locations showed in table 1.

Figure 2. Bayesian inference tree of Xenotoca variata from concatenated sequences of the four genes (4,034 bp). Bayesian posterior probabilities (>90%) and maximum likelihood bootstrap values (>80%) (BBP/MLB) are indicated. The divergence time estimations are shown with 95% HPD. Also, the morphological differences between clades and lineages are shown.

Figure 3. Haplotypes network of *cytb*. Size of the circles indicates the frequency of the haplotype and colors correspond to the discrete biogeographic regions described by Domínguez-Domínguez et al., (2006) were the samples belong as show. Numbers inside circles correspond to mutation steps. The smallest circles correspond to mutation steps.

Figure 4. Haplotypes network, A) with *S7* gene, B) with *RAG1* gene, C) with *RHO* gene. Size of the circles indicates the frequency of the haplotype and colors correspond to the discrete biogeographic regions described by Domínguez-Domínguez et al., (2006) were the samples belong as show. The smallest and black circles correspond to mutation steps.

Figure 5. Demographic history of populations of each recovered clade within lineage I and the lineage II using Bayesian Skyline Plot (BSP) from *cytb* sequences. Dotted lines represent the location of the upper bound (Max), the mean (X) and lower bound (Min) of the HPD = 95%.

Supporting information

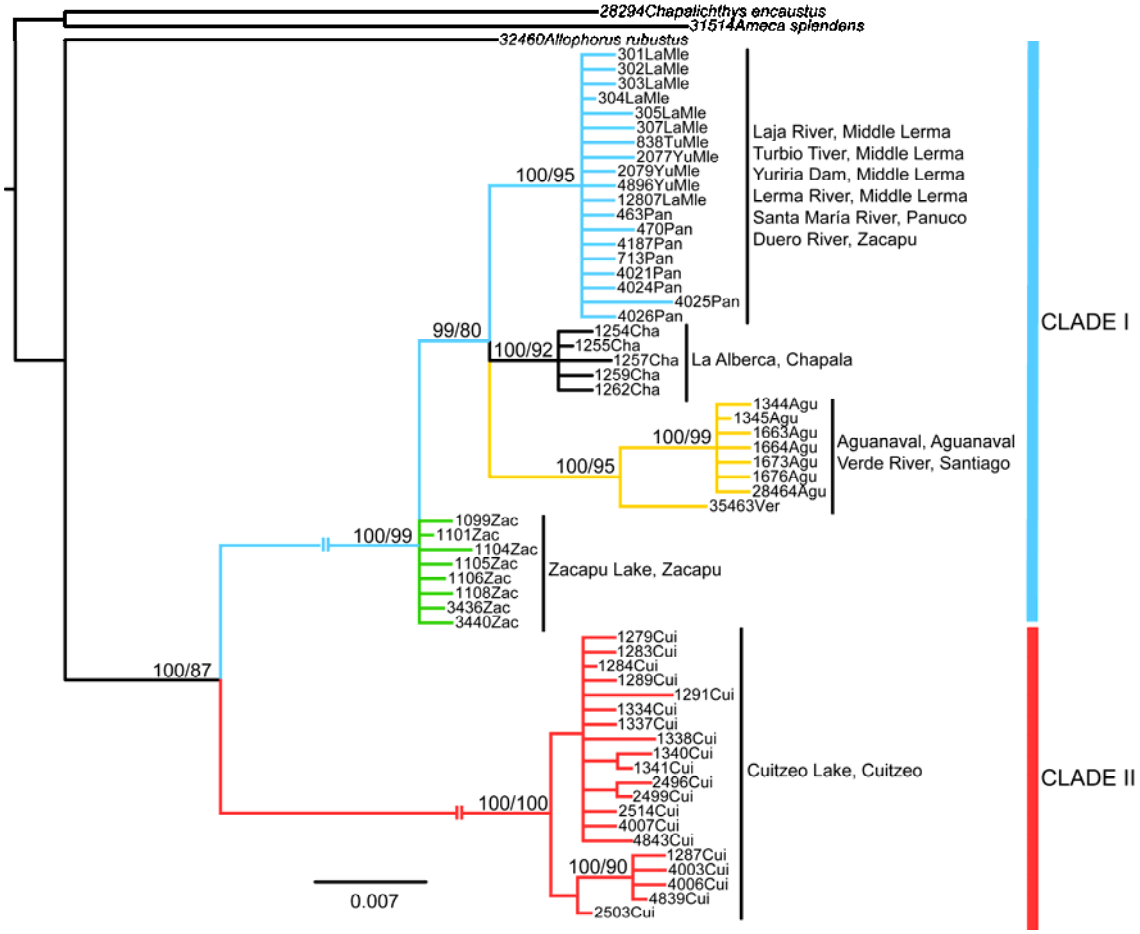


Figure S1. The Bayesian inference tree of *X. Variata* from *cytb* mitochondrial gene (1000 bp). Bayesian posterior probability (>90%; before the diagonal) and maximum likelihood bootstrap values (>80%; after the diagonal) are indicated.

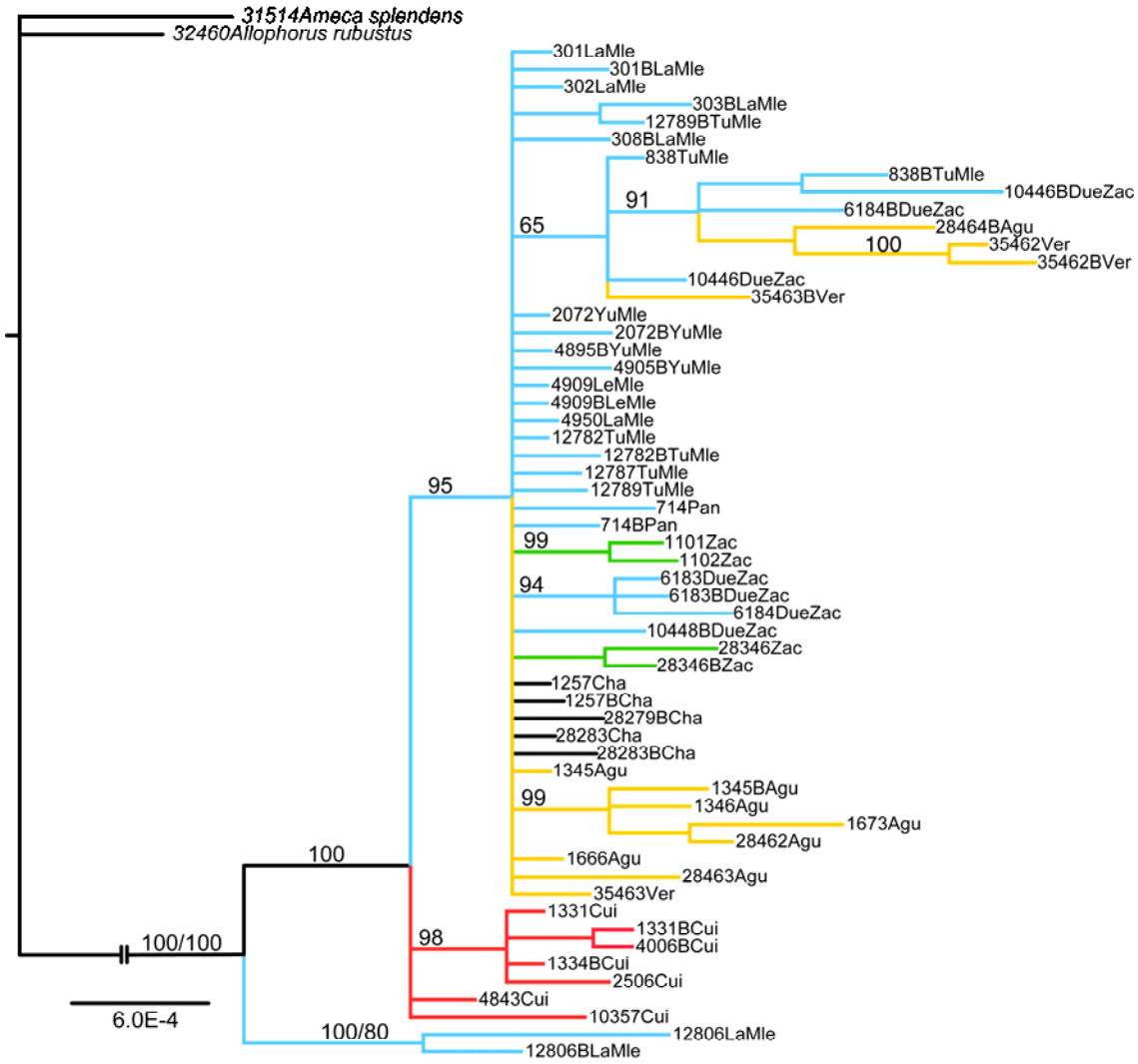


Figure S2. The Bayesian inference tree of X. Variata from concatenated sequences of three nuclear genes (S7, RAG1 and RHO: 3,034 bp). Bayesian posterior probability (>90%; above the branches) and máximum likelihhod bootstrap values (>80%) are indicated.

Table S1. Primers, PCR conditions, and References.

	cytb	*S7*	*RHO*	*RAG1*
Primers	Glu-F	S7RPEX1F	RH193F	RAGF
	Thr-R	S7RPEX3R	RH1073R	RAG9R
Size (bp)	1000	722	847	1465
Reference	Doadrio and Domínguez (2004)	Chow y Hazama, (1998)	Chen, Bonillo & Lecointre (2003)	Quenouille, Bermingham & Planes (2004)
Denaturing (step 1)		94ºC, 3 min.	94ºC, 3 min.	94ºC, 3 min.
Cycles (step 2)		35	35	35
Denaturing		94ºC, 45 s.	94ºC, 45 s.	94ºC, 45 s.
Annealing		56.5ºC, 46 s.	50.5ºC, 45 s.	59.5ºC 45 s.
Extension		72ºC, 60 s.	72ºC, 60 s.	72ºC, 60 s.
Final Extension (step 3)		72ºC, 7 min.	72ºC, 7 min.	72ºC, 7 min.

Table S2. Tissue voucher number, and access number of GenBank

Site	Locality	Biogeographic region	Voucher number	Access number of GenBank			
				cytb	*S7*	*RAG1*	*RHO*
1	Tierra Quemada, Santa María River	Panuco	461	YES	NO	NO	NO
1			462	YES	YES	YES	YES
1			463	YES	NO	NO	NO
1			464	YES	YES	YES	YES
1			465	YES	NO	NO	NO
1			469	YES	NO	NO	NO
1			470	YES	NO	NO	NO
2	Jesús María, Santa María River	Panuco River	709	YES	NO	NO	NO
2			711	YES	NO	NO	NO
2			712	YES	NO	NO	NO
2			713	YES	NO	NO	NO
2			714	YES	YES	YES	YES
2			715	YES	NO	NO	NO
2			716	YES	NO	NO	NO
2			717	YES	NO	NO	NO
2			4021	YES	NO	NO	NO
2			4024	YES	NO	YES	YES
2			4025	YES	NO	NO	NO
2			4026	YES	NO	NO	NO
3	La Alberca	Chapala Lake	1254	YES	NO	NO	NO
3			1255	YES	NO	NO	NO
3			1256	YES	NO	NO	NO
3			1257	YES	YES	YES	YES
3			1258	YES	NO	NO	NO
3			1259	YES	NO	NO	NO
3			1260	YES	NO	NO	NO
3			1261	YES	NO	NO	NO
3			1262	YES	NO	NO	NO
3			2586	YES	NO	NO	NO
3			2587	YES	NO	NO	NO
3			2588	YES	NO	NO	NO
3			2589	YES	NO	NO	NO
3			2590	YES	NO	NO	NO
3			2591	YES	NO	NO	NO

3			2592	YES	NO	NO	NO
3			2593	YES	NO	NO	NO
3			2594	YES	NO	NO	NO
3			2595	YES	YES	YES	YES
3			28278	YES	YES	YES	YES
3			28279	YES	YES	YES	YES
3			28280	YES	NO	NO	NO
3			28281	YES	NO	NO	NO
3			28282	YES	YES	YES	YES
3			28283	YES	YES	YES	YES
3			28284	YES	NO	NO	NO
4	San Francisco del Rincón, Turbio River	Middle Lerma River	832	YES	NO	NO	NO
4			833	YES	NO	NO	NO
4			838	YES	YES	YES	YES
4			4693	YES	NO	NO	NO
4			4695	YES	NO	NO	NO
4			4696	YES	NO	NO	NO
5	Gallinero Dam, Laja River	Middle Lerma River	307	YES	NO	NO	NO
5			308	YES	YES	YES	YES
5			309	YES	NO	NO	NO
5			310	YES	NO	NO	NO
5			311	YES	YES	YES	YES
6	Yuriria Dam	Middle Lerma River	2071	YES			
6			2072	YES	YES	YES	YES
6			2075	YES			
6			2077	YES			
6			2078	YES			
6			2079	YES			
6			4703	YES			
7	Ignacio Allende Dam, Laja River	Middle Lerma River	301	YES	YES	YES	YES
7			302	YES	YES	YES	YES
7			303	YES	YES	YES	YES
7			304	YES			
7			305	YES			
7			4950	YES	YES	YES	YES
7			4951	YES	YES	YES	YES
7			4952	YES			
7			12806	YES	YES	YES	YES
7			12807	YES			
7			12808	YES		YES	

7			12845	YES	YES	YES	YES
8	Urideo Sprig, Lerma River	Middle Lerma River	4909	YES	YES	YES	YES
8			4910	YES	YES	YES	YES
8			4911	YES			
8			4912	YES			
9	Joya Grande, Yuriria Dam	Midle Lerma River	4895	YES	YES	YES	YES
9			4896	YES	YES	YES	YES
9			4897	YES			
9			4898	YES	YES	YES	YES
9			4899	YES			
9			4900	YES	YES	YES	YES
9			4901	YES			
9			4902	YES			
9			4903	YES			
9			4904	YES			
9			4905	YES	YES	YES	YES
10	San José del Rodeo, Turbio River		12781	YES			
10			12782	YES	YES	YES	YES
10			12783	YES			
10			12784	YES			
10			12785	YES			
10			12786	YES			
10			12787	YES	YES	YES	YES
10			12788	YES			
10			12789	YES	YES	YES	YES
10			12790	YES			
11	Afluente Neutla, Laja River	Middle Lerma River	4955	YES	YES	YES	YES
11			4956	YES	YES	YES	YES
11			4957	YES			
12	Belisario, Cuitzeo Lake	Cuitzeo Lake	1279	YES			
12			1283	YES			
12			1284	YES			
12			1285	YES			
12			1287	YES			
12			1289	YES			
12			1291	YES			
12			1293	YES			
12			1294	YES			
12			1299	YES			
12			1301	YES			

13	San Cristóbal, Cuitzeo Lake	Cuitzeo Lake	2496	YES	YES	YES	YES
13			2499	YES			
13			2501	YES	YES		YES
13			2502			YES	YES
13			2503	YES	YES		YES
13			2504			YES	
13			2506	YES			
13			2510	YES			
13			2512	YES			
13			2513	YES			
13			2514	YES			
13			2515	YES			
13			4003	YES			
13			4005	YES			
13			4006	YES			
13			4007	YES			
			10357	YES	YES	YES	YES
			10358	YES			
14	La Mintzita, Morelia River	Cuitzeo Lake	1331	YES	YES	YES	YES
14			1332	YES			
14			1333	YES			
14			1334	YES		YES	YES
14			1336	YES	YES	YES	YES
14			1337	YES			
14			1338	YES		YES	YES
14			1339	YES			
14			1340	YES			
14			1341	YES			
15	Mariano Escobedo, Cuitzeo Lake	Cuitzeo Lake	4839	YES			
15			4843	YES	YES	YES	YES
15			4845	YES			
15			4846	YES			
15			4847	YES			
15			4848	YES	YES	YES	YES
15			4849	YES			
16	Zacapu, Zacapu Lake	Zacapu Lake	1099	YES			
16			1101	YES	YES	YES	YES
16			1102	YES	YES	YES	YES
16			1104	YES	YES	YES	YES
16			1105	YES	YES	YES	YES

16			1106	YES			
16			1108	YES	YES	YES	YES
16			1128	YES	YES	YES	YES
16			3436	YES			
16			3440	YES			
16			28346	YES	YES	YES	YES
16			28347	YES	YES	YES	YES
16			28348	YES			
16			28349	YES			
16			28350	YES			
16			28351	YES			
17	Melchor Ocampo, Duero River	Zacapu region	6183	YES	YES	YES	YES
17			6184	YES	YES	YES	YES
17			6185	YES			
17			6186	YES	YES	YES	YES
17			6187	YES			
17			10446	YES	YES	YES	YES
17			10447	YES	YES	YES	YES
17			10448	YES	YES	YES	YES
17			10449	YES			
17			10450	YES			
17			10451	YES			
17			10452	YES			
18	La Estancia Stream, Verde River	Verde River	35462	YES	YES	YES	YES
18			35463	YES	YES	YES	YES
18			35464	YES	YES		YES
18			35465	YES			
18			36264	YES		YES	
18			36266	YES	YES	YES	YES
18			36267	YES	YES	YES	
18			36268	YES			
18			36269	YES			
18			36270	YES			
18			36271	YES		YES	YES
18			36272	YES		YES	YES
18			36273	YES			
18			36274	YES			
19	Atotonilco, Aguanaval River	Aguanaval River	1344	YES			
19			1345	YES	YES	YES	YES
19			1346	YES	YES	YES	YES

19			4086	YES			
19			4087	YES			
19			4088	YES			
19			28455	YES			
19			28456	YES	YES	YES	YES
19			28462	YES	YES	YES	YES
19			28463	YES	YES	YES	YES
19			28464	YES	YES	YES	YES
20	Valenciana, Aguanaval River	Aguanaval River	1663	YES	YES		
20			1664	YES			
20			1665	YES			
20			1666	YES	YES	YES	YES
20			1667	YES			
20			1668	YES			
20			1669	YES			
20			1670	YES			
20			1673	YES	YES	YES	YES
20			1674	YES			
20			1675	YES			
20			1676	YES	YES	YES	YES
20			1677	YES			
20			1678	YES			

REFERENCES

Arnauld, C., Carot, P. & Fauvet-Berthelot, M. F. (2014). Arqueología de las Lomas en la cuenca lacustre de Zacapu, Michoacán, México. Centro de estudios mexicanos y centroamericanos, México. 231 pp. DOI: 10.4000/books.cemca.2626

Arnqvist, G., Edvardsson, M., Friberg, U. & Nilsson, T. (2000). Sexual conflict promotes speciation in insects. *Proceedings of the National Academy of Sciences of the United States of America*, 97, 10460-10464.

Avise, J. C. (1998). The history and purview of phylogeography: a personal reflection. *Molecular Ecology*, *7*, 371-379.

Avise, J. C. (2000). Phylogeography: The history and formation of species. Harvard university press. 464 pp.

Beltrán-López, R. G., Domínguez-Domínguez, O., Guerrero, J. A., Corona-Santiago, D. K., Mejía-Mojica, H., & Doadrio, I. (2017). Phylogeny and taxonomy of the genus *Ilyodon* Eigenmann, 1907 (Teleostei: Goodeidae), based on mitochondrial and nuclear DNA sequences. *Journal of Zoological Systematics and Evolutionary Research*, 55(4), 340-355. DOI: 10.1111/jzs.12175

Beltrán-López, R. G., Domínguez-Domínguez, O., Pérez-Rodríguez, R., Piller, K. R., & Doadrio, I. (2018). Evolving in the highlands: The case of the Neotropical Lerma Live-bearing *Poeciliopsis infans* (Woolman, 1894) (Cyprinodontiformes: Poeciliidae) in central México. *BMC Evolutionary Biology*, *18*, 56, https://doi.org/10.1186/s12862-018-1172-7.

Beltrán-López, R. G., Domínguez-Domínguez, O., Piller, K. R., Mar-Silva, A. F. & Doadrio, I. Genetic differentiation among populations of the freshwater fish *Goodea atripinnis* Jordan, 1880 (Cyprinodontiformes: Goodeidae), implications for its evolutionary history and taxonomy. *En preparación*

Bermingham, E. & Martin, A. P (1998). Comparative mtDNA phylogeography of neotropical freshwater fishes: testing shared history to infer the evolutionary landscape of lower Central America. *Molecular Ecology*, 7(4), 499-517.

Betancourt-Resendes, I., Pérez-Rodríguez, R., & Domínguez-Domínguez, O. (2018). Speciation of silverside *Chirostoma attenuatum* (Pisces: Atheriniformes) in Central Mexico. *Journal of Zoological Systematics and Evolutionary Research*, 56(3), 323-334. DOI: 10.1111/jzs.12216

Bloom, D. D., Piller, K. R., Lyons, J., Mercado-Silva, N. & Medina-Nava, M. (2009). Systematics and biogeography of the silverside tribe Menidiini (Teleostomi: Atherinopsidae) based on the mitochondrial ND2 gene. *Copeia*, 2, 408–417. https://doi.org/10.1643/CI-07- 151

Bradley, R. D. & Baker, R. J. (2001). A test of the genetic species concept: Cytochrome-b sequences and mammals. *Journal of Mammalogy*, 82, 960-973.

Bryson, R. W. Jr., García-Vázquez, U. O. & Riddle, B. R. (2011a). Phylogeography of Middle American gopher-snakes: mixed responses to biogeographical barriers across the Mexican Transition Zone. *Journal of Biogeography*, 38, 1570–1584

Burridge, C. P., Craw, D. & Waters, J. M. (2006). River capture, range expansion, and cladogenesis: The genetic signature of freshwater vicariance. *Evolution*, 60, 1038-1049.

Casal-López, M. & Doadrio. (2018). The Messinian imprint of the evolution of freshwater fishes of the genus *Luciobarbus* Heckel, 1843 (Teleostei, Cyprinidae) in the western Mediterranean. *Journal of Biogeography*, *45*, 1593-1603.

Chen, W. J., Bonillo, C. & Lecointre, G. (2003). Repeatability of clades as a criterion of reliability: a case study for molecular phylogeny of Acanthomorpha (Teleostei) with larger number of taxa. *Molecular phylogenetics and Evolution*, *26*, 262-288.

Chow, S. & Hazama, K. (1998). Universal PCR primers for S7 ribosomal protein gene introns in fish. *Molecular Ecology*, *7*, 1247-1263.

Correa-Metrio, A., Bush, M. B., Cabrera, K. R., Sully, S., Brenner, M., Hodell, D. A., Escobar, J. & Guilderson, T. (2012). Rapid climate change and no-analog vegetation in lowland Central America during the last 86,000 years. *Quaternary Science Reviews*, 38, 63-75.

De la Vega-Salazar, M. Y. (2006). Estado de Conservación de los peces de la familia Goodeidae (Cyprinodontiformes) en la mesa central de México. *Revista de Biología Tropical*, 54(1), 163-177.

Doadrio, I., & Domínguez, O. (2004). Phylogenetic relationships within the fish family Goodeidae based on cytochrome b sequence data. *Molecular Phylogenetics and Evolution*, 31, 416–430.

Doadrio, I., Perdices, A. & Machordom, A. (1996). Allozymic variation of the endangered killifish *Aphanius iberus* and its application to conservation. *Environmental Biology of Fishes*, *45*, 259–271.

Domínguez-Domínguez, O., Mercado-Silva, N. & Lyons, J. (2005). Conservation status of Mexican goodeids: problems, perspectives, and solutions. *In*: Uribe MC, Grier HJ, editors. Viviparous Fishes. New life publications. pp. 495–504.

Domínguez-Domínguez, O., Doadrio, I., Martínez-Meyer, E., Zambrano, L., & Pérez-Ponce de León, G. (2006). Using ecological niche modeling as a conservation tool for freshwater species: The live-bearing fishes (Goodeidae) in Central México. *Conservation Biology*, 20, 1730–1739.

Domínguez-Domínguez, O., Boto, L., Alda, F., Pérez-Ponce de León, G. & Doadrio, I. (2007a). Human impacts on drainages of the Mesa Central of Mexico, and its genetic effects on an endangered fish, *Zoogoneticus quitzeoensis*. *Conservation Biology,* 21,168–180.

Domínguez-Domínguez, O., Pompa-Domínguez, A. & Doadrio, I. (2007b). A new species of the Genus *Yuriria* Jordan & Evermann, 1896 (actinopterygii, cyprinidae) from the Ameca Basin if the Central Mexican Plateau. *Graellsia*, *63*, 259-271.

Domínguez-Domínguez, O., Alda, F., Pérez-Ponce de León, G., García-Garitagoitia, J. L. & Doadrio, I. (2008). Evolutionary history of the endangered fish *Zoogoneticus quitzeoensis* (Bean, 1898) (Cyprinodontiformes: Goodeidae) using a sequential approach to phylogeography based on mitochondrial and nuclear DNA data. *BMC Evololutionary Biology, 8*, 161.

Domínguez-Domínguez, O., Pérez-Rodríguez, R., Escalera-Vázquez, L. H. & Doadrio, I. (2009). Two new species of the genus *Notropis* Rafinesque,

1817 (Actinopterygii, Cyprinidae) from Lerma River Basin in the Central Mexican Plateau. *Hidrobiológica*, *19*, 159–172.

Domínguez-Domínguez, O., Pedraza-Lara, C., Gurrola-Sánchez, N., Perea, S., Pérez-Rodríguez, R., Israde-Alcántara, I., ... Brooks, D. R. (2010). Historical biogeography of the goodeinae (Cyprinodontiformes). In M. C. Uribe-Aranzabal, & H. J. Grier (Eds.), *Viviparous fishes II* (pp. 13– 30). Florida: New Life Publications.

Domínguez-Domínguez, O., Bernal-Zuñiga, D. M., & Piller, K. R. (2016). Two new species of the genus *Xenotoca* Hubbs and Turner, 1939 (Teleostei, Goodeidae) from central-western México. *Zootaxa*, 4189, 81-98.

Drummond, A. J., Rambaut, A., Shapiro, B. & Pybus, O. G. (2005). Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and Evolution*, *22*, 1185-1192.

Drummond, A. J., Ho, S. Y. W., Phillips, J. & Rambaut, A. (2006). Relaxed phylogenetics and dating with confidence. *PLoS Biology*, 5:e88.

Drummond, A. J., Suchard, M. A., Xie, D. & Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, *29*, 1969–1973.

Duarte-Sánchez, P. (1981). Contribución al conocimiento de los hábitos alimentarios de 12 especies de peces del Lago de Cuitzeo, Mich. Tesis Profesional, Escuela Nacional de Ciencias Biológicas, IPN. México, 63 pp.

Edwards, S. V., Kingan, S. B., Calkins, J. D., Balakrishnan, C. N., Jennings, B., Swanson, W. J. & Sorenson, M. D. (2005). Speciation in birds: Genes, geography, and sexual selection, *Proceedings of the National Academy of Sciences, 3*, 6550–6557.

Excoffier, L. & Lischer, H. E. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, *10*, 564–567.

Ferrari, L., Garduño, V., Pasqueré, G. & Tibaldi, A. (1994). Volcanic and tectonic evolution of central Mexico: Oligocene to present. *Geofísica Internacional*, *33*, 107–123.

Fitzsimons, J. M. (1972). A Revision of two genera of Goodeid fishes (Cyprinodontiformes, Osteichthyes) from the Mexican Plateau. *Copeia*, 4, 728.

Garduño-Monroy, V., Arreygue-Rocha, E., Israde-Alcántara, I. & Rodríguez-Torres, G. (2001). Efectos de las fallas asociadas a sobreexplotación de acuíferos y la presencia de fallas potencialmente sísmicas en Morelia, Michoacán, México. *Revista Mexicana de Ciencias Geológicas*, 18, 37–54.

González-Peña, D., Gómez-Blanco, D., Reboiro-Jato, M., Fernández-Riverola, F. & Posada, D. (2010). ALTER: program-oriented conversion of DNA and protein alignments. *Nucleic Acids Research*, *38*, 14-18.

Gravilets, S., Arnqvist, G. & Friberg, U. (2001). The evolution of female mate choice by sexual conflict. *Proceedings of the Royal Society of London B: Biological Sciences*, *268*, 531-539.

Hasenaka, T. (1994). Size, distribution, and magma output rate for shield volcanoes of the Michoacan-Guanajuato volcanic field, Central Mexico. *Journal of Volcanology and Geothermal Research*, 63, 13–31. doi:10.1016/0377-0273(94)90016-7.

Hasenaka, T. & Carmichael, I. (1985). The cinder cones at Michoacan–Guanajuato, Central Mexico: their age, volume, and distribution, and magma discharge rate. *Journal of Volcanology and Geothermal Research*, 25, 105–124. doi:10.1016/0377-0273(85) 90007-1.

Heled, J. & Drummond, A. J. (2008). Bayesian inference of population size history from multiple loci. *BMC Evolutionary Biology*, 8:289.

Huidobro, L., Morrone, J. J., Villalobos, J. L. & Álvarez, F. (2006). Distributional patterns of freshwater taxa (fishes, crustaceans and plants) from the Mexican transition zone. *Journal of Biogeography*, *33*, 731-741.

Huson, D. H. & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, 23(2), 254–67.

Israde-Alcántara, I. (1999). Lagos Volcánicos y Tectónicos de Michoacán In: *Carta Geológica de Michoacán Escala 1:250000.* Garduño-Monroy, V. H., Corona-Chávez, P., Israde-Alcántara, I., Mennella, L., Arreygue, E., Bigioggero, B. & Chiesa, S. (Eds.). Universidad Michoacana de San Nicolás de Hidalgo, México. 45-72.

Israde-Alcántara, I. & Garduño-Monroy V, H. (1999). Lacustrine record in a volcanic intra-arc setting: the evolution of Late Neogene Cuitzeo basin

system (central-western Mexico, Michoacan). *Palaeogeography, Palaeoclimatology and Palaeoecology*, *151*, 209-227.

Israde-Alcántara, I., Robles-Camacho, J. & Domínguez, J. M. (2008). Beidellite-nontronite clays in Neogene sediments from Cuitzeo-Charo lacustrine basin Michoacán, México. Geological setting and paleoenvironmental implications. *Boletín de la Sociendad Geológica Mexicana*, *60*, 159-171.

Krebs, S. L. & Brandon, R. A. (1984). A new species of salamander (family Ambystomatidae) from Michoacan, Mexico. *Herpetológica*, 40, 238-245.

Kumar, S., Stecher, G. & Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution*, 33(7), 1879-1874, doi: 10.1093/molbev/msw054.

Lanfear, R., Calcott, B., Ho, S. Y. & Guindon, S. (2012). PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyzes. *Molecular Biology and Evolution*, 29(6),1695–701.

Librado, P. & Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*. 25(11), 1451–1452.

Machordom, A. & Doadrio, I. (2001). Evidence of a cenozoic Betic-Kabilian connection based on freshwater fish phylogeography (*Luciobarbus*, Cyprinidae). *Molecular Phylogenetics and Evolution*, 18(2), 252–263.

Martin, O. Y. & Hosken, D. J. (2003). The evolution of reproductive isolation through sexual conflicto. *Nature*, 423, 979-982.

Mateos, M., Sanjur, O. I. & Vrijenhoek, R. C. (2002). Historical biogeography of the livebearing fish genus *Poeciliopsis* (Poeciliidae: Cyprinodontiformes). *Evolution*, *56*, 972–984.

Mejía-Madrid, H., Vásquez-Domínguez, E. & Pérez Ponce de León, G. (2007). Phylogeography and freshwater basins in central Mexico: recent history as revealed by the fish parasite *Rhabdochona lichtenfelsi* (Nematoda). *Journal of Biogeography*, 34, 787-801.

Miller, R. R., Minckley, W. L. & Norris, S. M. (2005). Freshwater Fishes of México. Chicago, IL: The University of Chicago Press.

Moncayo-Estrada, R., Israde-Alcántara, I. & Garduño-Monroy, V. H. (2001). La cherehuita *Hubbsina turneri* De Buen (1941) (Pisces: Goodeidae). Origen, distribución y su uso en la regionalización de la cuenca del Lerma. *Hidrobiológica*, *11*, 1-18.

Near, T. J. & Benard, M. F. (2004). Rapid allopatric speciation in logperch darters (Percidae: Percina). *Evolution*, *58*, 2798–2808.

Ornelas-García, C. P., Domínguez-Domínguez, O. & Doadrio, I. (2008). Evolutionary history of the fish genus *Astyanax* Baird & Girard (1854) (Actinopterygii, Characidae) in Mesoamerica reveals multiple morphological homoplasies. *BMC Evololutionary Biology*, *8*, 340.

Ornelas-García, C. P., Alda, F., Díaz-Pardo, E., Gutiérrez-Hernández, A. & Doadrio I. (2012). Genetic diversity shaped by historical and recent factors in the live-bearing twoline skiffia *Neotoca bilineata*. *Journal of Fish Biology*, 81, 1963-1984.

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), 353–70.

Pedraza-Lara, C., Doadrio, I., Breinholt, J. W. & Crandall, K. A. (2012). Phylogeny and Evolutionary Patterns in the Dwarf Crayfish Subfamily (Decapoda: Cambarellinae). *Plos One*, *7*, e48233.

Perea, S., Cobo-Simo, M. & Doadrio, I. (2016). Cenozoic tectonic and climatic events in southern Iberian peninsula: implications for the evolutionary history of freshwater fish of the genus *Squalius* (Actinopterygii, Cyprinidae). *Molecular Phylogenetics and Evolution*, *97*, 155–169.

Pérez-Rodríguez, R., Domínguez-Domínguez, O., Pérez-Ponce de León, G. & Doadrio, I. (2009). Phylogenetic relationships and biogeography of the genus *Algansea* Girard 1856 (Cypriniformes:Cyprinidae) from central Mexico inferred from molecular data. *BMC Evolutionary Biology, 9*, 223.

Pérez-Rodríguez, R., Domínguez-Domínguez, O., Doadrio, I., Cuevas- García, E., & Pérez-Ponce de León, G. (2015). Comparative historical biogeography of three groups of Neartic freshwater fishes across central México. *Journal of Fish Biology*, *86*, 993–1015.

Pérez-Rodríguez, R., Domínguez-Domínguez, O., Mar-Silva, F., Doadrio, I. & Pérez-Ponce de León, G. (2016). The historical biogeography of the southern group of the sucker genus *Moxostoma* (Teleostei: Catostomidae) and the colonization of Central México. *Zoological Journal of the Linnean Society*, 177, 633-647.

Piller, K. R., Kenway-Lynch, C. S., Camak, D. T., & Domínguez-Domínguez, O. (2015). Phylogeography and population structure of the imperilled redtail splitfin (Goodeidae: *Xenotoca eiseni*): Implications for Conservation. *Copeia*, 103, 440–454. https://doi.org/10.1643/CI-14- 067.

Quenouille, B., Bermingham, E. & Planes, S. (2004). Molecular systematics of the damselfishes (Teleostei: Pomacentridae): Bayesian phylogenetic analyses of mitochondrial and nuclear DNA sequences. *Molecular Phylogenetics and Evolution*, 31, 66–88. PMID: 15019609.

Rambaut, A. & Drummond, A. (2007). Tracer v1.5. Available: http://tree. bio.ed.ac.uk/software/tracer

Rambaut, A., Suchard, M. A., Xie, D. & Drummond, A. J. (2014) Tracer v1.6. Available from http://beast.bio.ed.ac.uk/Tracer.

Rice, W.R. (1989). Analyzing tables of statistical test. *Evolution*, 43, 223-225.

Ritchie, M. G., Webb, S. A., Graves, J. A., Magurran, A. E. & Macías-García, C. (2005). Patterns of speciation in endemic Mexican Goodeid fish: sexual conflict or early radiation?. *Journal of Evolutionary Biology*, *18*, 922-929.

Ritchie, M. G., Hamill, R. M., Graves, J. A., Magurran, A. E., Webb, S. A. & Macías-García, C. (2007). Sex and differentiation: population genetic divergence and sexual dimorphism in Mexican goodeid fish. *Journal of Evolutionary Biology*, *20*, 2048-2055.

Ronquist, F., Teslenko, M., Van der Mark, P., Ayres, D. L., Darling, A., Höhna, S., . . . Huelsenbeck, J. P. (2012). MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, 61, 539–542.

Sandoval F. de P. (1981). Obras Sucesos y Fantasías en el Lago de Chapala. Gobierno de Jalisco. Secretaría General. Unidad Editorial. Colección Textos Jalisco. Serie: Estudios e Inversión. No. 15. 77 pp.

Schönhuth, S., Doadrio, I., Domínguez-Domínguez, O., Hillis, D. M. & Mayden, R. (2008). Molecular evolution of southern North American Cyprinidae (Actinopterygii), with the description of the new genus *Tampichthys* from central Mexico. *Molecular Phylogenetics and Evolution*, *47*, 729-756.

Silvestro, D. & Michalak, I. (2012). raxmlGUI: A graphical frontend for RAxML. *Organisms Diversity and Evolution*, 12, 335–337. https://doi. org/10.1007/s13127-011-0056-0.

Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30, 1312–1313.

Stephens, M. & Donelly, P. A. (2003). Comparison of Bayesian methods for haplotype reconstruction from population genotype data. *American Journal of Human Genetics*, 73(5), 1162–9.

Strecker, U., Faúndez, V. H., Wilkens H: Phylogeography of surface and cave Astyanax (Teleostei) from Central and North America based on cytochrome b sequence data. *Molecular Phylogenetics & Evolution* 2004, 33:469–481

Villa-Villaseñor, I. M. (2013). Variación intraespecífica en la elección de pareja en Xenotoca variata Bean 1887; (Cyprinodontiformes: Goodeidae) entre dos poblaciones de las cuencas de Cuitzeo y Zacapu. Tesis de Licenciatura. Universidad Michoacana de San Nicolás de Hidalgo, México. 62 pp.

Webb, S. A., Graves, J. A., Macías-García, C., Magurran, A. E., Foighil, D. O., & Ritchie, M.G. (2004). Molecular phylogeny of the livebearing Goodeidae (Cyprinodontiformes). *Molecular Phylogenetics and Evolution*, 30, 527–544.

West, R. C. (1964). Surface configuration and asociated geology of middle America. In: Wanchope, R. & West, R. C. (Eds.). Handbook of middle American Indians, University of Texas Press, Austin. p. 33–82.

Zardoya, R. & Doadrio, I. (1999). Molecular evidence on the evolutionary and biogeographical patterns of European Cyprinids. *Journal of Molecular Evolution*, 49, 227–237.

CAPÍTULO IV

Evolving in the highlands: the case of the Neotropical Lerma live-bearing *Poeciliopsis infans* (Woolman, 1894) (Cyprinodontiformes: Poeciliidae) in Central Mexico

Beltrán-López, R. G., Domínguez-Domínguez, O., Pérez-Rodríguez, R., Piller, K. R., & Doadrio, I. (2018). Evolving in the highlands: The case of the Neotropical Lerma Live-bearing *Poeciliopsis infans* (Woolman, 1894) (Cyprinodontiformes: Poeciliidae) in central México. *BMC Evolutionary Biology*, *18*, 56, https://doi.org/10.1186/s12862-018-1172-7.

ABSTRACT

Background Volcanic and tectonic activities in conjunction with Quaternary climate are the main events that shaped the geographical distribution of genetic variation of many lineages. *Poeciliopsis infans* is the only poeciliid species that was able to colonize the temperate highlands of central Mexico. We inferred the phylogenetic relationships, biogeographic history, and historical demography in the widespread Neotropical species *P. infans* and correlated this with geological events and the Quaternary glacial-interglacial climate in the highlands of central Mexico, using the mitochondrial genes Cytochrome b and Cytochrome oxidase I and two nuclear loci, Rhodopsin and ribosomal protein S7.

Results Populations of *P. infans* were recovered in two well-differentiated clades. The maximum genetic distances between the two clades were 3.3% for *cytb*, and 1.9% for *coxI*. The divergence of the two clades occurred *ca*. 2.83 Myr. Ancestral area reconstruction revealed a complex biogeographical history for *P. infans*. The Bayesian Skyline Plot showed a demographic decline, although more visible for clade A, and more recently showed a population expansion in the last 0.025 Myr. Finally, the habitat suitability modelling showed that during the LIG, clade B had more areas with high probabilities of presence in comparison to clade A, whereas for the LGM, clade A showed more areas with high probabilities of presence in comparisons to clade B.

Conclusions *Poeciliopsis infans* has had a complex evolutionary and biogeographic history, which, as in other co-distributed freshwater fishes, seems to be linked to the volcanic and tectonic activities during the Pliocene or early Pleistocene. Populations of *P. infans* distributed in lowlands showed a higher level of genetic diversity than populations distributed in highlands, which could be linked to more stable and higher temperatures in lowland areas. The fluctuations in population size through time are in agreement with the continuous fluctuations of the climate of central Mexico.

Keywords: Endemic fishes, Tecto-volcanism, Quaternary climate, Pliocene, Biogeographical history

BACKGROUND

Volcanic and tectonic activities since the Miocene have had a substantial influence on the diversification of many New World taxa [1, 2]. Paleoclimatic events since the Pliocene, mainly from the Quaternary to the present; also have influenced the distribution of many organisms by changing the climates in boreal, temperate and tropical zones. Geologic and quaternary climatic events together are the main factors that have shaped the geographical distribution of genetic variation at the species, population, and community levels in several taxa, including fishes [3-11].

The beta diversity of freshwater fishes worldwide, including 80% of all freshwater species described, demonstrates that geographical isolation of drainage basins, combined with Quaternary climate changes, provides a parsimonious explanation for present-day patterns of spatial turnover in the global freshwater fish fauna [12]. Under this context, the geographical location, complex topography, geological dynamism including extensive volcanism since the Miocene and the climatic history of central Mexico, changed during the Quaternary [8, 13], have shaped a biogeographically complex area, characterized by ecological components that have allowed for the coexistence of taxa of Neotropical and Nearctic origins, as well as endemic groups [14, 15].

The biogeographic limits of this area have been largely discussed, and even differ depending on the taxa analyzed [14, 16-18]. In general terms, central Mexico is a high plateau bounded by the Sierra Madre Oriental to the east and by the Sierra Madre Occidental to the west and crossed by the Trans Mexican Volcanic Belt (TMVB) from west to east, with elevations up to 1400-1800 MASL [18], a region also called the Mesa Central by some authors [19]. The region is characterized by a temperate climate, thus allowing for the establishment of fish species, mainly of Neartic origin. The climatic changes during the quaternary have influenced the geographical distribution of genetic lineages of terrestrial organisms in space and time, as is the case of snakes [7], lizards [8, 20, 21], birds [9], small mammals [22-24], and plants [2].

The dynamic geological processes that have occurred since the Miocene have promoted the genesis and destruction of aquatic ecosystems [25] and have been considered as the primary forces that have influenced the biogeography and the complex evolution of several taxa of freshwater organisms [26-32]. However, most of the research has focused on understanding the complex evolutionary history of freshwater fishes of Nearctic origin such as Goodeids [28, 29, 33, 34], Cyprinids [30], Catostomids [35], and a combination of taxa [36]. Only one group of species in the

TMVB that are of Neotropical origin, genus *Poeciliopsis*, has been previously investigated in this manner [26].

The small live-bearing topminnow *Poeciliopsis infans* (Woolman, 1894) is a member of the family Poeciliidae, which has more than 220 species of tropical preferences [37]. *Poeciliopsis infans* is the only Neotropical fish species that has colonized the temperate highlands of the TMVB, including the Lerma-Santiago Basin, headwaters of the Ameca, Armeria, Coahuayana, Balsas and Panuco Basins, as well as endorheic lakes in the region (Fig. 1) [38]. *Poeciliopsis infans* is co-distributed with fishes of Nearctic origin of the families Goodeidae, Catostomidae, Ictaluridae and Cyprinidae [38].

Accordingly, since the members of the Poeciliidae are a group of Neotropical origin, most of them are adapted to tropical habitats. Furthermore, since *P. infans* is the only poeciliid species living in the temperate highlands of central Mexico, an area dominated by fish species of Neartic origin, we expect that the evolutionary history of *P. infans* could be explained by recent volcanic and tectonic activities. However, since *P. infans* is a species evolving in the marginal and temperate areas adjacent to the distribution of all of the other species *Poeciliopsis,* it may not follow the same patterns as other co-distributed species in the region.

We extensively sampled throughout the distribution of *P. infans (*Teleostomi: Poeciliidae*)*, and gathered mtDNA and nDNA sequences to infer phylogeographic variation, historical biogeography, and historical demography of the widespread *P. infans.* These data will allow us to examine the influence of geological history and Quaternary glacial-interglacial climatic events, in space and time, in the evolutionary history of Neotropical species evolving in a predominant Nearctic area.

MATERIALS AND METHODS

Sample Collection

Two hundred fifty-six specimens of *P. infans* from throughout the range were collected (Fig. 1 and Table 1) using electrofishing equipment and trawl nets. Tissue samples (fin clips) were preserved in 95% ethanol for DNA extraction, and a maximum of five specimens per site were preserved in 5% formalin. Despite intensive collection efforts, samples were not obtained from some biogeographic regions. Fish and tissue samples were deposited in the fish collection of the Universidad Michoacana de San Nicolas de Hidalgo, Mexico (SEMARNAT registration number MICH-PEC-227-07-09; for voucher numbers see Additional file 1).

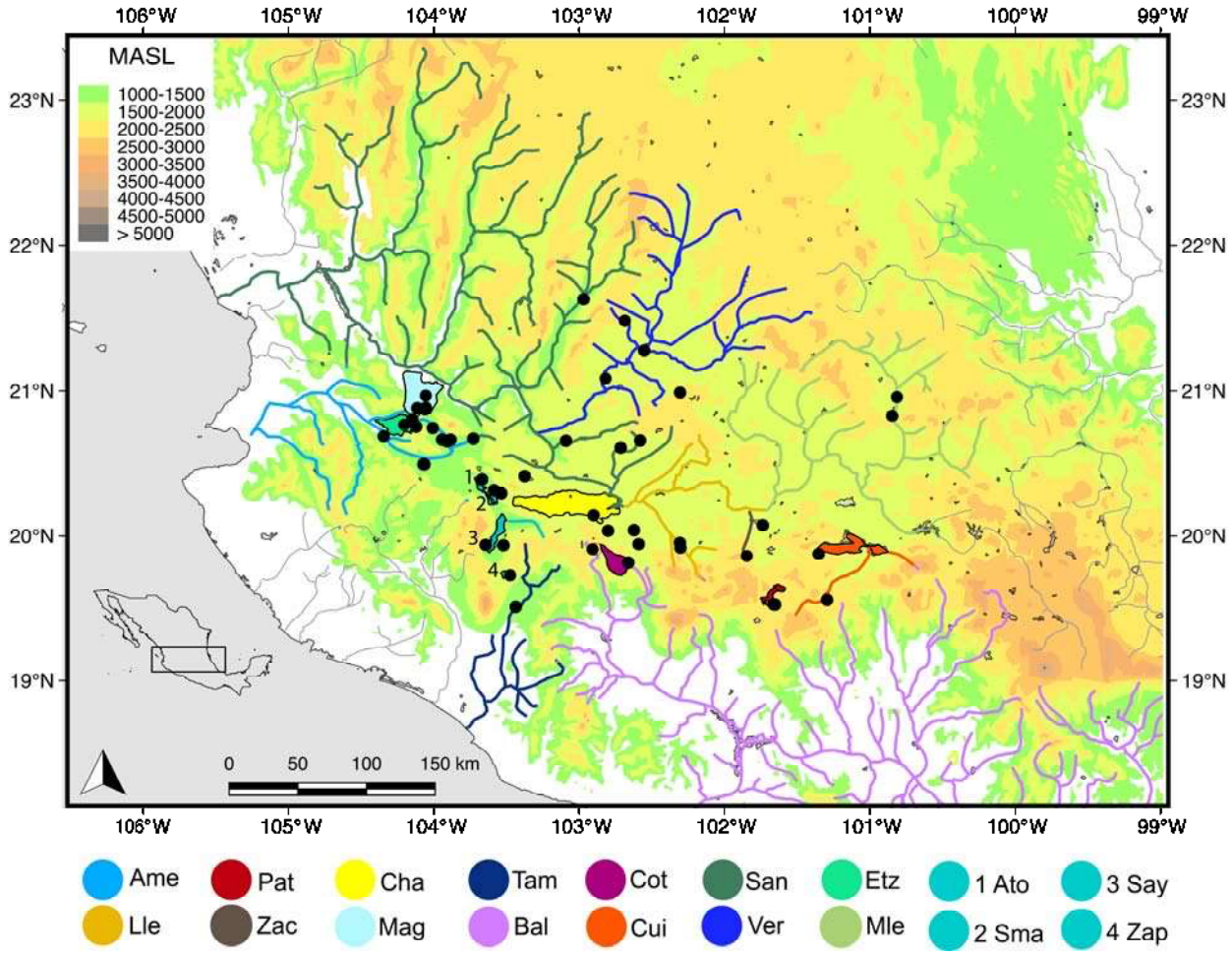


Figure 1 Sampling locations and the biogeographical regions where *Poeciliopsis infans* is distributed. The colors of the biogeographical regions corresponded with the colors used in the phylogenetic analyses. The codes of the biogeographical regions are as follows: (Mag) Magdalena Lake; (Etz) Etzatlan-San Marcos region; (Ver) Verde River; (Ame) Ameca River; (Ato) Atotonilco Lake as the number 1; (Sma) San Marcos Lake as the number 2; (Say) Sayula Lake as the number 3; (Zap) Zapotlan Lake as the number 4, these four lakes belong to Sayula region; (San) Grande de Santiago River; (Tam) Tamazula River; (Pat) Patzcuaro Lake; (Cha) Chapala Lake; (Bal) Balsas River; (Cot) Cotija Lake; (Lle) Lower Lerma Basin; (Mle) Middle Lerma Basin; (Zac) Zacapu region; and, (Cui) Cuitzeo Lake. The meters above see level (MASL) are indicated with level curves.

Table 1 Samples localities and sequence information.

Site	Locality	Basin	Biogeographic region	Sequences number *Cytb*/*coxI*/*S7*/*RHO*	GPS Coordinates
1	Los Venados	Magdalena	Magdalena	5/5/3/4	20° 54´ 5.5´´ N, 104° 4´ 44.8´´ W
2	Laguna	Magdalena	Magdalena	11/8/10/6	20° 54´ 14.2´´ N, 104° 1´ 11.6´´ W
3	Presa San Ignacio	Ameca	Ameca	7/6/3/6	20° 30´ 40.6´´ N, 104° 2´ 12.4´´ W
4	Chapulimita	Ameca	Ameca	2/2/2/2	20° 40´ 48.9´´ N, 103° 54´ 29.3´´ W
5	Salida presa Tecuan	Ameca	Ameca	7/6/5/5	20° 20´ 5.2´´ N, 103° 45´ 20.2´´ W
6	Manantial Los Veneros	Ameca	Ameca	2/2/2/2	20° 40´ 9.7´´ N, 103° 52´ 25.3´´ W
7	Tala, Río Salado	Ameca	Ameca	3/3/2/3	20° 41´ 12.1´´ N, 103° 41´ 36.3´´ W
8	Amatlán de cañas	Ameca	Ameca	6/5/6/5	20° 42´ 13.7´´ N, 104° 18´ 34.4´´ W
9	Teuchitlán	Cocula-La Vega	Ameca	10/6/6/4	20° 40´ 46.8´´ N, 103° 50´ 59.2´´ W
10	San Juanito de Escobedo	Laguna Colorada	Ameca	5/3/5/1	20° 45´ 37.3´´ N, 103° 59´ 39.6´´ W
11	S. M. San Julian	Verde	Verde	3/3/2/2	21° 0´ 31.7´´ N, 102° 17´ 47.9´´ W
12	San Nicolás	Verde	Verde	11/13/5/9	21° 17´ 45.4´´ N, 102° 32´ 59.7´´ W
13	Arroyo La Estancia	Verde	Verde	2/2/1/2	21° 24´ 36.3´´ N, 102° 44´ 15´´ W
14	Río Colorado	Verde	Verde	4/4/0/4	21° 5´ 6.1´´ N, 102° 52´ 10.8´´ W
15	Río Xoconostle-San Juan	Laja	Middle Lerma	3/3/3/3	20° 56´ 31.5´´ N, 100° 58´ 38´´ W
16	Manantial Andrés-Figueroa	San Marcos	Sayula	8/9/7/4	21° 17´ 45.4´´ N, 102° 32´ 59.7´´ W
17	Manantial San Marcos	San Marcos	Sayula	5/5/5/5	20° 20´ 0.4´´ N, 103° 34´ 57.6´´ W
18	Canal Presa Buena Vista	Atotonilco-Sayula	Sayula	6/7/6/3	20° 24´ 28´´ N, 103° 39´ 59.7´´ W
19	Villa corona	Atotonilco-Sayula	Sayula	3/3/3/3	20° 24´ 28.4´´ N, 103° 39´ 59.5´´ W
20	Manantial Cuyacapán	Sayula	Sayula	15/17/9/11	19° 57´ 16.6´´ N, 103° 30´ 51.7´´ W
21	Laguna de Zapotlán	Zapotlan	Sayula	2/2/1/0	19° 44´ 44.9´´ N, 103° 28´ 22.4´´ W
22	Río Las Puentes	Chapala	Chapala	6/7/7/4	20° 3´ 22.4´´ N, 102° 46´ 2.1´´ W
23	Cojumatlán	Chapala	Chapala	3/2/0/1	20° 9´ 45.7´´ N, 102° 52´ 4.6´´ W
24	Los Negritos	Chapala	Chapala	5/3/5/2	20° 3´ 36.3´´ N, 102° 36´ 46.1´´ W
25	Presa Nueva	Chapala	Chapala	5/5/5/4	19° 57´ 46.2´´ N, 102° 34´ 42.4´´ W
26	Manantial La Mintzita	Cuitzeo	Cuitzeo	6/6/5/3	19° 34´ 40´´ N, 101° 16´ 28.7´´ W

27	Ojo de Agua San Cristóbal	Cuitzeo	Cuitzeo	9/10/7/6	19° 53´ 37.4´´ N, 101° 19´ 0.5´´ W
28	Embarcadero Principal	Patzcuaro	Patzcuaro	6/7/5/1	19° 32´ 42.6´´ N, 101° 37´ 2.5´´ W
29	Urandén	Patzcuaro	Patzcuaro	8/7/6/4	19° 32´ 47.4´´ N, 101° 38´ 28.2´´ W
30	Presa Melchor Ocampo	Angulo-Lerma	Zacapu	7/7/6/2	20° 5´ 36.5´´ N, 101° 43´ 57.4´´ W
31	La Zarcita	Angulo-Lerma	Zacapu	9/10/9/4	19° 49´ 19´´ N, 101° 47´ 51´´ W
32	Laguna de Zacapu	Angulo-Lerma	Zacapu	5/5/5/1	19° 49´ 20.9´´ N, 101° 47´ 15.8´´ W
33	Atenquique	Tuxpan	Tamazula	1/1/0/1	19° 31´ 46.3´´ N, 103° 25´ 56.3´´ W
34	Puente en Jacona	Río Duero	Lower Lerma	5/5/5/4	19° 58´ 14´´ N, 102° 17´ 46.2´´ W
35	Presa La Luz	Río Duero	Lower Lerma	8/8/7/5	19° 56´ 13´´ N, 102° 17´ 56.9´´ W
36	Quitupan	Tepalcatepec	Balsas	5/5/5/3	19° 55´ 34.8´´ N, 102° 52´ 54.7´´ W
37	Presa San Juanico	Cotija	Cotija	11/9/11/6	19° 49´ 57.4´´ N, 102° 38´ 25.8´´ W
38	San Sebastián	Etzatlan-San Marcos	Etzatlan-San Marcos	2/2/2/1	20° 49´ 25´´ N, 104° 7´ 10.8´´ W
39	Presa Palo Verde	Etzatlan-San Marcos	Etzatlan-San Marcos	13/12/10/12	20° 46´ 9.6´´ N, 104° 6´ 48.2´´ W
40	Cuescomatitlán	Lago de Cajititlán	Grande de Santiago	5/5/5/2	20° 25´ 48.4´´ N, 103° 21´ 37´´ W
41	Jalpa	Juchipila	Juchipila	2/2/2/1	21° 39´ 6.5´´ N, 102° 57´ 57.8´´ W
42	San Antonio	Santiago-Chapala	Grande de Santiago	2/2/2/2	20° 40´ 27.2´´ N, 102° 33´ 19.4´´ W
43	Presa de Garabato	Santiago-Chapala	Grande de Santiago	6/6/4/5	20° 37´ 28.4´´ N, 102° 41´ 15.6´´ W
44	Río Tinajeros	Santiago-Chapala	Grande de Santiago	7/8/1/1	20° 40´ 20.7´´ N, 103° 4´ 41.9´´ W

DNA Extraction, PCR Amplification and Sequencing

Total genomic DNA was isolated with the Qiagen BioSprint Dneasy Tissue and Blood Kit (Qiagen, Valencia, CA, USA) following the manufacturer's protocol. Two hundred fifty-six individuals were amplified for Cytochrome b (*cytb*: 1,083 bp) and 249 individuals for Cytochrome Oxidase Subunit I (*coxI*: 631 bp), for 1,771 bp of mitochondrial sequence data. The first intron, a fragment of second intron, and the second exon of the gene coding for the S7 ribosomal protein (*S7*: 859 bp), and the gene coding for the Rhodopsin protein (*RHO*: 845 bp) were amplified for 1,704 bp of nuclear sequence data from 201 individuals. For this subset, individuals were chosen in order to represent all biogeographic regions and all the variation shown in mitochondrial dataset. In addition, five sequences of *cytb* were obtained from GenBank with the following accession numbers: AF412134 of Lerma River, AF412135 of Ameca River, AF412136 of Chapala Lake, AF412137 of Santiago River and AF412138 of Panuco River.

Each fragment was individually amplified using the Polymerase Chain Reaction (PCR) in volumes of 12.5 µl, containing 4.25 µl ultrapure water, 0.5 µl of each 0.2 µM primer, 6.25 µl Dream Taq Green PCR Master Mix 2x (Thermo Scientific), and 1 µl (ca. 10-100 ng) of DNA template. The specific PCR protocols of each gene are provided in Additional file 2. The recovered PCR products were purified using ExoSAP-IT (USB Corp.) and submitted to Macrogen Inc. (The Netherlands) for sequencing. Nucleotide sequences were edited and aligned in Mega v.6.06 [39], and the chromatographs were examined by eye. For the nuclear gene (*RHO*), the heterozygous genotypes were phased using DNAsp v.5.10 [40] with the algorithm provided by PHASE v.2.0 [41]. Whenever sequences of *S7* showed heterozygous positions defined by indels, a manual reconstruction of the two-allele phases was performed following the procedure described by [42]. The obtained sequences were deposited in GenBank under the follow access number: for *cytb* (MG028009 to MG028278), for *coxI* (MG028279 to MG028543), for *RHO* (MG100617 to MG100824) and for *S7* (MG366191 to MG366482), (see Additional file 1).

Phylogenetic Analyses and Haplotype Networks

Recombination of the nuclear genes (*RHO*, P = 1.0; *S7*, P = 0.23) was tested using the PHI test in Splitstree v.4.13 [43]. DNA sequences of each of the four genes (*cytb*, *coxI*, *S7* and *RHO*) were collapsed to haplotypes using the web-based program ALTER [44].

The phylogenetic analyses were conducted with each gene separately, with concatenated datasets for mitochondrial genes, for concatenated nuclear genes, and

for the four genes combined. The phylogenetic analyses with concatenated genes were performed with the available sequences of the nuclear genes (201 specimens), the sequences of mitochondrial genes were adjusted to these number, which were representative of all biogeographic regions and account the variability found in mitochondrial genes.

The performance of the phylogenies with the four concatenated genes *vs*. mitochondrial genes concatenated and *vs*. nuclear genes concatenated were assessed using Bayes Factors (BF). Bayes Factors were calculated from the estimated harmonic means of likelihood using the sump command in MrBayes. Decisions were made based on the 2ln BF criterion, with BF > or = 10 considered as strong evidence for rejecting the null hypothesis [45].

Model selection based on Akaike information criterion (AIC) and optimal partition settings were performed using PartitionFinder v.1.1.0 [46], and recovered the best partition by assigning substitution models for each gene. The parameters of each model are provided in Additional file 3.

Gene trees were constructed with Maximum Likelihood implemented in RAxMLGUI v.1.3.1 [47, 48], using the substitution model implemented for each gene. The GTR+G+I [48] substitution model was used for the concatenated genes matrix and 10,000 bootstrap replicates with the algorithm ML + rapid bootstrap. The relative stability of clades was evaluated by 1,000 non-parametric bootstrap replicates [49].

Bayesian Inference was implemented in MrBayes v.3.2.1 [50], with the substitution models for each gene obtained in PartitionFinder. Analyses were run for 10 million generations, with two independent runs implementing four Markov Chain Monte Carlo (MCMC) processes and sampling every 500 generations. We evaluated the chains for convergence with the log-likelihood (-lnL) values of the two independent runs using Tracer v.1.5 [51], and discarded 10% as burn-in to construct the consensus tree. *Poeciliopsis prolifica* was used as outgroup, based on the results of a previous study [26].

In order to determine the geographic distribution of haplotypes for all populations of *P. infans* for nuclear genes, we reconstructed two independent TCS haplotype networks (a phylogenetic network estimation using statistical parsimony) for *RHO* and *S7* sequences using PopArt v.1.7 (http://popart.otago.ac.nz).

Divergence Time Estimation and Genetic Distances

The program BEAST v.1.8.1 [52] was used to estimate the most recent common ancestor for clades within *P. infans.* This analysis was carried out with

a subset of 145 sequences that include all different haplotypes for all genes. Because the lack of fossil data for *Poeciliopsis*, the molecular clock was calibrated using the mutation rate of *cytb* in teleosts of 0.76-2.2%/million years [53, 54, 55]. Since the mutation rate is not available for the other genes, they were included in the analysis without calibration information.

The model parameters were unlinked across *cytb*, *coxI*, *S7* and *RHO* genes and substitution models were set according to the selected model for each gene by PartitionFinder v. 1.1.0 [46]. We applied a lognormal relaxed clock (Uncorrelated) model on branch length [56]. We selected the tree prior Coalescent: Extended Bayesian Skyline Plot [57], and estimated a starting tree using the random method. A MCMC analysis with 50 million of generations was conducted, and sampled every 1,000 generations. We assessed whether parameter values had reached effective sample size and convergence in Tracer v.1.5. [51]. Finally, the maximum clade credibility tree was built, discarding the first 10% of the trees as burn-in, using Tree Annotator v.1.8.1. [52].

Uncorrected genetic distances were calculated among the recovered groups in the phylogenetic trees for each mitochondrial gene (*cytb*, *coxI*), and between all individuals for *S7* and *RHO* in Mega v.6.06 [39]. A bootstrapping process was performed with 1,000 repetitions.

Genetic Diversity and Population Structure

For each gene (*cytb*, *coxI*, *S7* and *RHO*), the number of haplotypes (H), polymorphic sites (S), nucleotide (π) and haplotype (h) diversities were obtained to estimate genetic diversity levels in all populations of *P. infans*. To examine genetic differentiation at different hierarchical levels, as well as geographical patterns of population subdivision, an analysis of molecular variance (AMOVA) was conducted. The AMOVAs were implemented for the four separate genes and groupings according to: 1) inferred clades in phylogenetic analyses, 2) according to the biogeographic regions *sensu* [28] and, a third analysis was performed without *a priori* groupings. The analyses were conducted using 10,000 permutations to assess significance values. All genetic diversity analyses and AMOVAs were performed in Arlequin v.3.5.1 [58].

Ancestral Area Reconstruction

The ancestral area reconstruction for *P. infans* was estimated using the statistical Dispersal-Vicariance (S-DIVA) method [59], and the Dispersal-Extinction-Cladogenesis (DEC) model [60, 61] as implemented in RASP v.3.2 software [62]. These methods use statistical approaches to reconstruct biogeographic history, and allowed us to compare both results. The ultrametric and dichotomous tree obtained for *cytb* in BEAST was used as the tree topology on which ancestral areas were mapped. For both analyses, the maximum number of areas was limited to two. For these analyses, the distributional area of *P. infans* was divided into 15 biogeographic regions: (Mag) Magdalena Lake; (Etz) Etzatlán-San Marcos region; (Ver) Verde River; (Ame) Ameca River; (San) Grande de Santiago River; (Tam) Tamazula River; (Pat) Patzcuaro Lake; (Cha) Chapala Lake; (Bal) Balsas River; (Cot) Cotija Lake; (Lle) Lower Lerma River; (Mle) Middle Lerma River; (Zac) Zacapu Lake and (Cui) Cuitzeo Lake, and, within Sayula region are considered the follow Lakes: (Ato) Atotonilco, (Sma) San Marcos, (Say) Sayula and (Zap) Zapotlan [28, 29].

Historical Demography

The population size fluctuations through time were tested with a Coalescent Bayesian Skyline Plot (BSP) analysis [63] as implemented in BEAST v.1.8.1 [52]. This analysis only was implemented with sequences of *cytb* due the higher number of available sequences. The substitution rate was the same as the divergence time analysis and the substitution model was set according to the select model by PartitionFinder v. 1.1.0 [46].

An uncorrelated relaxed clock model was set *a priori*, and 70 million generations were run, sampled every 1,000 generations. Convergence was assessed with Tracer v.1.5 [51]. The first 10% of the states were discarded as burn-in.

Poeciliopsis infans Distribution Modelling

To evaluate the concordance between the historical demography obtained in BSP analyses and the potential distribution of *P. infans* in the past, we carried out a species distribution modelling analyses at different temporal

scales [64]. The estimations of the current and past population distribution were inferred with MaxEnt v.3.3.1 [65]. Geographical coordinates of 162 sites registered in the database of the Colección de Peces de la Universidad Michoacana de San Nicolás de Hidalgo were used as presence data (see Additional file 4). For the environmental data, we used 19 bioclimatic variables downloaded from WORLDCLIM database [66], http://www.worldclim.org, at a resolution of 30 arc-seconds for Last Inter Glacial (LIG) and 2.5 minutes for Last Glacial Maximum (LGM). The WORLDCLIM variables represent biologically meaningful summaries of precipitation and temperature in the present (1950–2000), and for the past, the Community Climate System Model (CCSM) for the LGM: 0.025 Myr, and the LIG: 0.15–0.10 Myr periods. To construct the models, we employed a logistic output [65], and the default settings. The value of the regularization multiplier was tested for 0.5, 1.0, 1.5 and 2.0, but the highest AUC value was for a regularization multiplier of one, and this value was used in all analyses. The model was run with 100 subsample replicates estimating mean habitat suitability values (S).

To evaluate if the performance of all distribution models was better than one, a random model was assessed using 75% of the presence data to run the model and the remaining 25% for statistical testing. In addition, the area under the receiver operating characteristic curve (AUC) was estimated to assess the accuracy of the models. A jackknife test of variables of importance was conducted to evaluate the relative importance of each climate variable [67]. Variables contributing the least to the model or those highly correlated [68] were removed for each model. The correlation of variables was evaluated through the response curves, which reflect the dependencies induced by correlations between the selected variable and other variables.

RESULTS

Phylogenetic Relationships and Haplotype Networks

The ambiguously aligned positions that showed the sequences of the S7 gene are see in Additional file 5. The BF comparison indicated that the analyses using the four concatenated genes provided a better explanation of the data than the mitochondrial genes concatenated or nuclear genes concatenated (BF

of four genes concatenated *vs* mitochondrial genes concatenated = 1.87 and BF of four concatenated genes *vs* nuclear genes concatenated = 2.2)

The phylogenetic results for both analyses (Maximum likelihood and Bayesian Inference) recovered two well differentiated and well supported clades for each mitochondrial gene (*cytb*: 1,083 bp and *coxI*: 631 bp) (Additional files 6 and 7), for the concatenated mitochondrial genes (*cytb*, *coxI*: 1,771 bp) (Additional file 8), and for all concatenated genes (*cytb*, *coxI*, *S7* and *RHO:* 3,475 bp; Fig. 2).

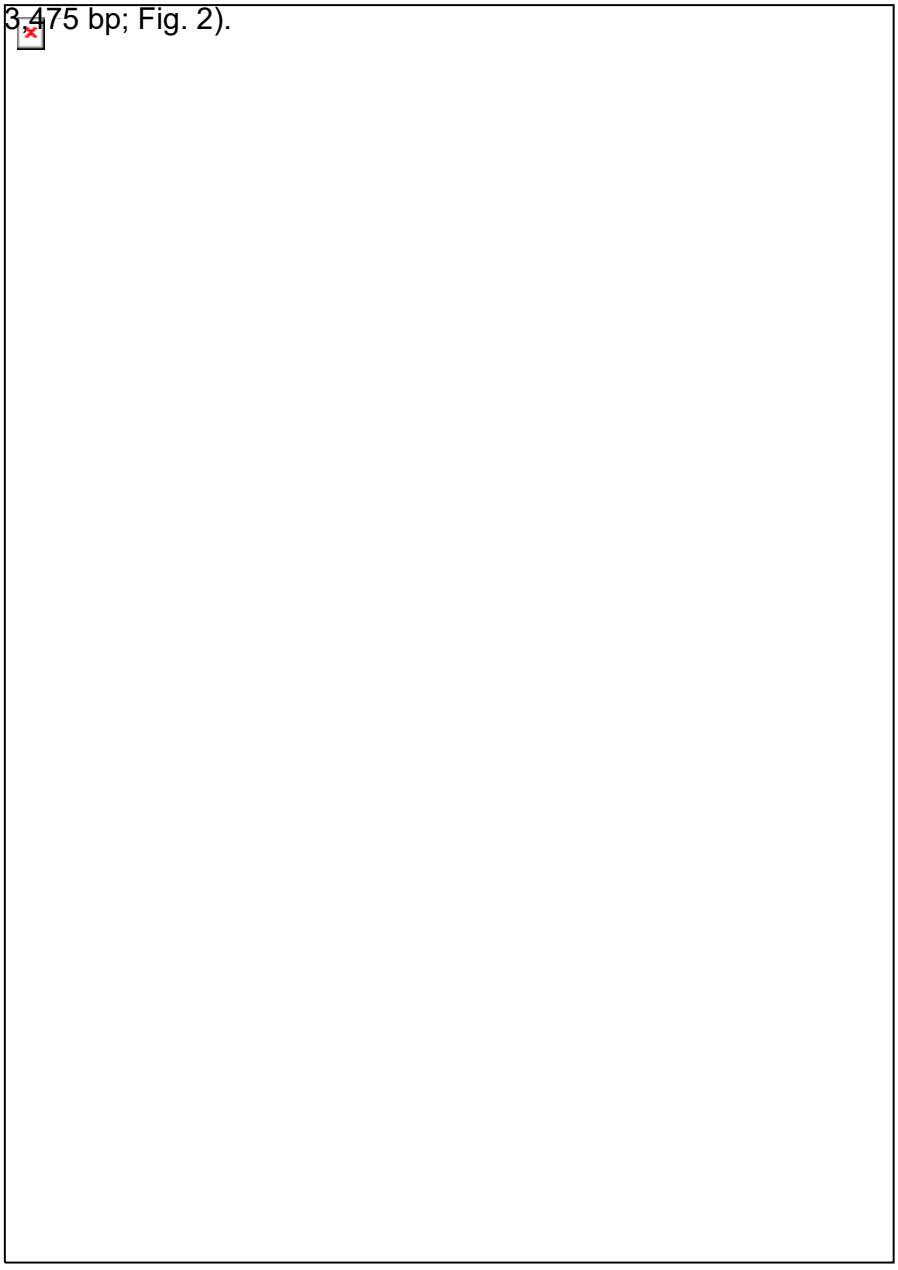


Figure 2 The Bayesian inference tree of *Poeciliopsis infans* from concatenated sequences of two mitochondrial (*cytb*, *coxI*) and two nuclear genes (*S7* and

RHO). Bayesian posterior probability (>0.9; above the branches) and maximum likelihood bootstrap values (>80%; below the branches) are indicated. The divergence time estimations are shown with 95% HPD.

In general terms, clade A clustered populations of the biogeographic regions of lowlands of TMVB, while; clade B clustered populations of the biogeographic regions of the highlands of the TMVB. However, one sample from Ameca, Magdalena and Sayula regions respectively were grouped in clade B. (Figs. 1 and 2).

Clade A clustered individuals of seven biogeographic regions and three sub-clades: sub-clade A1 included individuals from the Etzatlan-San Marcos region, Magdalena Lake and few individuals of Verde River; sub-clade A2 clustered individuals of the Ameca, Verde, and Grande de Santiago Rivers and the Atotonilco Lake of the Sayula region; while the sub-clade A3 clustered individuals of the Sayula, San Marcos and Zapotlan Lakes within the Sayula region, as well as samples from the Ameca River basin and Tuxpan River of the Tamazula biogeographic region. These three sub-clades were well supported in the Bayesian inference analysis, but the bootstrap support for sub-clade A1 was low (87%) (Fig. 2).

The second clade, clade B, clustered individuals of the Middle and Lower Lerma, Grande de Santiago and Balsas Rivers, as well as Cuitzeo, Patzcuaro, Chapala, Cotija and Zacapu Lakes. Also, three samples corresponding to regions that were clustered in clade A were recovered in clade B and corresponded to Magdalena and Atotonilco Lakes, and the Ameca River (Fig. 2). The phylogeny based on each nuclear gene and with both concatenated nuclear genes failed to recover resolved relationships (Additional files 9, 10 and 11).

For the haplotype networks based on the nuclear genes, two haplogroups (A and B) were recovered and these are highly congruent with the concatenated gene phylogeny. There are, however, shared haplotypes between them, with *RHO* network sharing the most haplotypes between groups. The first haplogroup (A), grouped individuals of Ameca, Grande de Santiago, Sayula, Magdalena, Verde and Etzatlan-San Marcos regions, for *S7*, one individual of the Chapala region was found in this haplogroup. The second haplogroup (B)

for both nuclear genes, grouped individuals of the Lower and Middle Lerma, Grande de Santiago, Balsas, Zacapu, Cuitzeo, Patzcuaro, Cotija and Chapala regions. Some individuals of the Sayula region were grouped in the haplogroup B for both genes (see Additional file 12).

Divergence Time Estimation and Genetic Distances

The first isolation event in *P. infans,* separating clades A and B, was estimated to have occurred near the middle Pliocene and middle Pleistocene *ca*. 2.83 Myr (95% HPD: 1.25-4.41 Myr). The first separation event of the three sub-clades within clade A was estimated to have occurred during the Pleistocene period *ca*. 0.95 Myr (95% HPD: 0.39-1.5 Myr), whereas the rest of the isolation events were calculated in less of a million of years (Figs. 2 and 3).

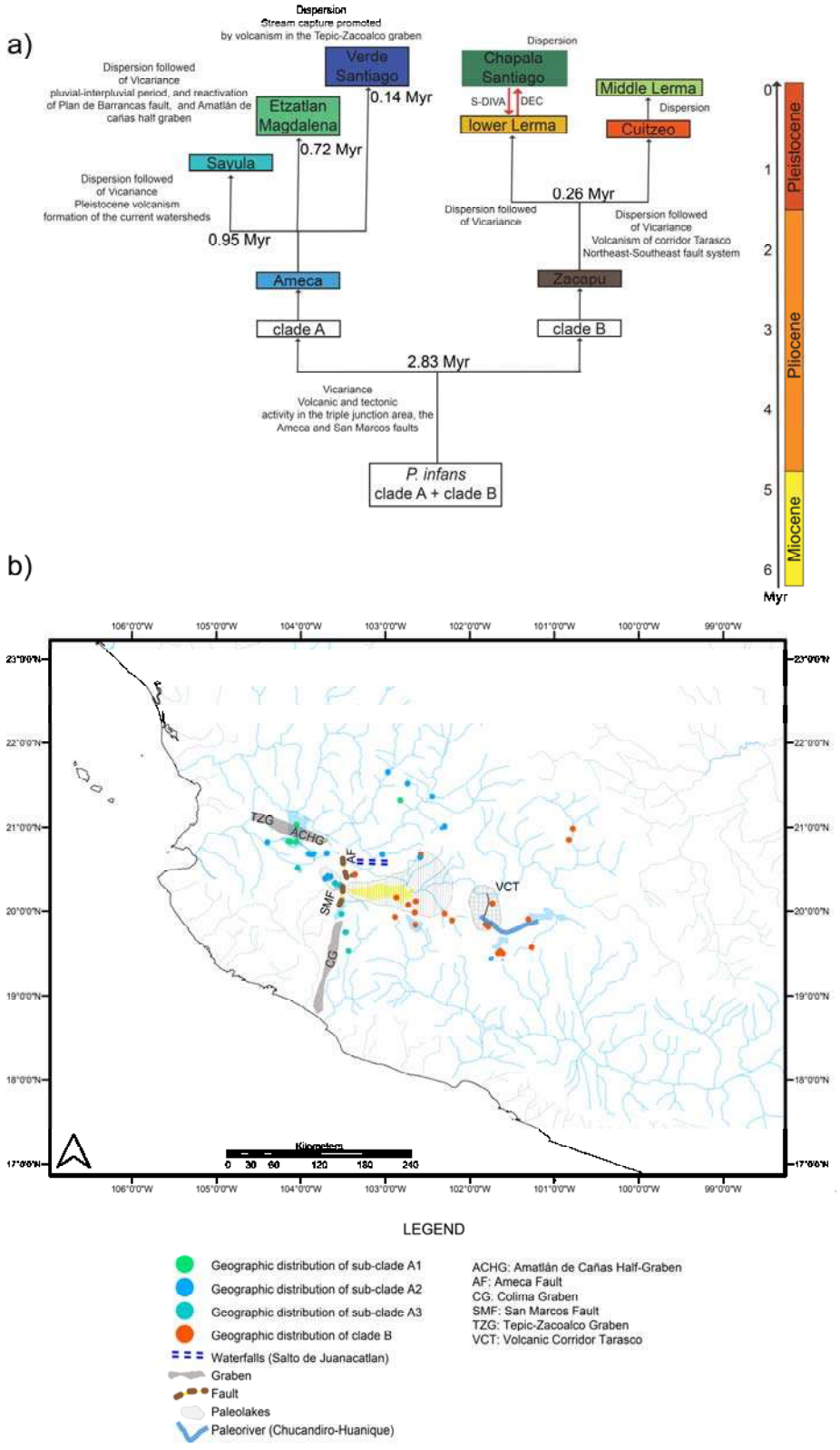


Figure 3 a) Divergence times, biogeographical events in *Poeciliopsis infans*, b) geographic distribution of inferred clades and sub-clades across of TMVB of *P. infans*, and tectonic, volcanic and hydrographical features of the TMVB.

The uncorrected mean genetic distances (*p*-distance) calculated between clade B and sub-clades A1, A2 and A3 for the mitochondrial genes ranged from 0.8-3.3% for *cytb*; and 0.7-1.9% for *coxI* (Table 2). The minimum genetic distances found for *cytb* were between sub-clade A1 and sub-clade A2 (0.8%), and the maximum distances were between sub-clade A3 and clade B (3.3%). Based on *coxI* the minimum genetic distances were observed between sub-clade A2 and sub-clade A3 (0.7%), and the maximum were between sub-clade A1 and clade B (1.9%). Between the three sub-clades of clade A the genetic distances were 0.8-1.1% for *cytb* and 0.7-1.3% for *coxI* (Table 2). For nuclear genes the genetic distances between all individuals included both alleles for each sequence, ranged between 0.0-0.5% for *RHO* and between 0.0-0.6% for S7.

Table 2 Uncorrected genetic distances presented in proportion.

cytb/*coxI*	Clade A			Clade B
	Sub-clade A1	Sub-clade A2	Sub-clade A3	
Sub-clade A1	0.001/0.000	0.008	0.010	0.031
Sub-clade A2	0.013	0.001/0.001	0.011	0.031
Sub-clade A3	0.009	0.007	0.001/0.000	0.033
Clade B	0.019	0.016	0.016	0.002/0.002

Genetic distances within recovered clades and sub-clades of *Poeciliopsis infans* based on *cytb* (to the rigth of the diagonal) and *coxI* (to the left of the diagonal) and between recovered groups in phylogenetic analyses based on *cytb* (above the diagonal) and *coxI* (below the diagonal) genes.

Genetic Diversity and Population Structure

The highest haplotype diversity was found in Cuitzeo (*h*= 0.81) for *cytb* and in the Verde River (*h*=0.62) for *coxI*, followed by the Ameca River Basin for both genes (*h*=0.75 for *cytb* and 0.42 for *coxI*), while, null haplotype diversity were found in the Patzcuaro, Cotija and Balsas regions. Atotonilco Lake, within the Sayula region and the Verde River exhibited the highest nucleotide diversity for *cytb* and *coxI* (π= 0.006, π= 0.004) respectively (Table 3).

Table 3 Genetic diversity for each biogeographic region of *Poeciliopsis infans* based on mitochondrial DNA data (*cytb* and *coxI*).

	cytb – coxI				
Biogeographic region	**N**	**S**	**H**	***π***	***h***
Magdalena	15-12	1-0	2-1	0.000-0.000	0.133 - 0.000
Ameca	37-30	18-5	9-3	0.001-0.001	0.752 - 0.427
Middle Lerma	3-3	3-0	3-1	0.001-0.000	1.000 - 0.000
San Marcos (Sayula)	13-14	3-1	4-2	0.000-0.000	0.525 - 0.362
Atotonilco (Sayula)	9-9	34-0	2-1	0.006-0.000	0.220 - 0.000
Sayula (Sayula)	15-17	4-2	5-3	0.001-0.000	0.695 - 0.227
Cuitzeo	15-16	8-2	6-3	0.001-0.000	0.819 - 0.241
Patzcuaro	14-14	0-0	1-1	0.000-0.000	0.000 - 0.000
Zacapu	21-22	8-2	6-2	0.001-0.000	0.557 - 0.173
Etzatlan-San Marcos	19-16	7-0	7-1	0.000-0.000	0.608 - 0.000
Chapala	19-17	2-0	3-1	0.000-0.000	0.292 - 0.000
Balsas	5-5	0-0	1-1	0.000-0.000	0.000 - 0.000
Cotija	11-9	0-0	1-1	0.000-0.000	0.000 - 0.000
Verde	19-21	9-9	2-3	0.002-0.004	0.280 - 0.620
Grande de Santiago (before SJ)	12-11	3-1	3-2	0.000-0.000	0.547 - 0.388
Grande de Santiago (after SJ)	10-11	2-1	3-2	0.000-0.000	0.375 - 0.250
Lower Lerma	13-13	3-4	2-3	0.000-0.001	0.282 - 0.410

N, sample size, S, polymorphic sites, H, number of haplotypes, *π,* nucleotide diversity *h,* haplotype diversity. SJ = Salto de Juanacatlan.

For the nuclear genes, the highest haplotype diversity was found in the Atotonilco (*h*=0.97), Ameca (*h*=0.9) for *S7*, and the Middle Lerma (*h*=0.73) for *RHO*. Absence of haplotype diversity was found for the Middle Lerma for *S7* and Cuitzeo for *RHO*. The populations of Atotonilco (*π*= 0.01) and the Middle Lerma (*π*= 0.001) exhibited the highest nucleotide diversity for *S7* and *RHO* respectively (Table 4). When the AMOVA was performed without groups *a priori*, the highest variation for all genes were among populations (*cytb*: 91.17%, *coxI*: 96.07%, *S7*: 69.5% and *RHO*: 60.64%) and not within populations. The AMOVA for all genes showed a high percentage of variation when populations were grouped according to the recovered clades and sub-clades from the phylogenetic analyses (*cytb*: 92.24%, *coxI*: 91.03%, *S7*: 45.48% and *RHO*:

53.21%), but not when populations were grouped according to the biogeographic regions [18] (*cytb*: 31.31%, *coxI*: 27.09%, *S7*: -143% and *RHO*: 53.21) (Tables 5 and 6).

Table 4 Genetic diversity for each biogeographic region of *Poeciliopsis infans* based on nuclear DNA data (*S7* and *RHO*).

	S7 – RHO				
Biogeographic region	**N**	**S**	**H**	***π***	***h***
Magdalena	24-18	10-6	6-5	0.002-0.000	0.713 - 0.405
Ameca	50-54	29-2	22-3	0.004-0.000	0.956 - 0.265
Middle Lerma	6-6	0-2	1-3	0.000-0.001	0.000 - 0.733
San Marcos (Sayula)	24-20	1-1	2-2	0.000-0.000	0.159 - 0.505
Atotonilco-Sayula (Sayula)	18-10	39-2	14-2	0.010-0.000	0.973 - 0.200
Sayula (Sayula)	20-22	7-2	6-3	0.002-0.000	0.790 - 0.177
Cuitzeo	24-18	2-0	3-1	0.000-0.000	0.358 - 0.000
Patzcuaro	22-10	1-1	2-2	0.000-0.000	0.173 - 0.466
Zacapu	40-14	4-2	5-2	0.000-0.000	0.315 - 0.142
Etzatlan-San Marcos	32-30	10-3	14-3	0.001-0.000	0.774 - 0.131
Chapala	34-26	8-2	8-3	0.001-0.000	0.711 - 0.280
Balsas	10-6	4-1	4-2	0.001-0.000	0.644 - 0.333
Cotija	22-18	2-1	3-2	0.000-0.000	0.450 - 0.529
Verde	14-32	11-1	6-2	0.004-0.000	0.822 - 0.000
Grande de Santiago	32-26	15-3	16-4	0.002-0.001	0.834 - 0.600
Lower Lerma	26-18	15-2	5-3	0.001-0.000	0.461 - 0.307

N, sample size (included the two alleles of each sequence), S, polymorphic sites, H, number of haplotypes, *π,* nucleotide diversity *h,* haplotype diversity.

Table 5 Analyses of molecular variance (AMOVA) of the mitochondrial data for select groups of *Poeciliopsis infans* at different hierarchical arrangements.

		Cytb		
Testing assumptions	Source of variation	% of variance	Fixation index	*P*- value
Grouped according to recovered clades and sub-clades	Among groups	92.24	Φ_{CT}: 0.92	<0.0001
	Among populations within groups	3.82	Φ_{SC}: 0.49	<0.0001
	Within populations	3.94	Φ_{ST}: 0.96	<0.0001
	Total	100		
Biogeographic regions	Among groups	31.31	Φ_{CT}: 0.31	ns
	Among populations within groups	63.5	Φ_{SC}: 0.94	<0.0001
	Within populations	5.19	Φ_{ST}: 0.94	<0.0001
	Total	100		
Without grouping a priori	Among populations	91.17	Φ_{ST}: 0.91	<0.0001
	Within populations	8.83		
	Total	100		
		CoxI		
Testing assumptions	Source of variation	% of variance	Fixation index	*P*- value
Grouped according to recovered clades and sub-clades	Among groups	91.03	Φ_{CT}: 0.91	<0.0001
	Among populations within groups	6.01	Φ_{SC}: 0.66	<0.0001
	Within populations	2.96	Φ_{ST}: 0.97	<0.0001
	Total	100		
Biogeographic regions	Among groups	27.09	Φ_{CT}: 0.27	ns
	Among populations within groups	69.01	Φ_{SC}: 0.94	<0.0001
	Within populations	3.9	Φ_{ST}: 0.96	<0.0001
	Total	100		
Without grouping a priori	Among populations	96.07	Φ_{ST}: 0.96	<0.0001
	Within populations	3.93		
	Total	100		

ns = not significant.

Table 6 Analyses of molecular variance (AMOVA) of the nuclear data for select groups of *Poeciliopsis infans* at different hierarchical arrangements.

		S7		
Testing assumptions	Source of variation	% of variance	Fixation index	*P*- value
Grouped according to recovered clades and sub-clades	Among groups	45.48	Φ_{CT}: 0.45	<0.0001
	Among populations within groups	28.26	Φ_{SC}: 0.51	<0.0001
	Within populations	26.26	Φ_{ST}: 0.73	<0.0001
	Total	100		
Biogeographic regions	Among groups		Φ_{CT}:	ns
	Among populations within groups		Φ_{SC}:	<0.0001
	Within populations		Φ_{ST}:	<0.0001
	Total	100		
Without grouping a priori	Among populations	69.5	Φ_{ST}: 0.69	<0.0001
	Within populations	30.5		
	Total	100		

		RHO		
Testing assumptions	Source of variation	% of variance	Fixation index	*P*- value
Grouped according to recovered clades and sub-clades	Among groups	53.21	Φ_{CT}: 0.53	<0.0001
	Among populations within groups	12.42	Φ_{SC}: 0.26	<0.0001
	Within populations	34.37	Φ_{ST}: 0.65	<0.0001
	Total	100		
Biogeographic regions	Among groups	15.86	Φ_{CT}: 0.15	ns
	Among populations within groups	44.92	Φ_{SC}: 0.53	<0.0001
	Within populations	39.22	Φ_{ST}: 0.60	<0.0001
	Total	100		
Without grouping a priori	Among populations	60.64	Φ_{ST}: 0.60	<0.0001
	Within populations	39.36		
	Total	100		

Ancestral Area Reconstruction

Ancestral area reconstruction using DEC and S-DIVA revealed a complex biogeographical history for *P. infans*, with several events of dispersal and vicariance (Figs. 3 and 4). Vicariance was most common in ancient events

in comparison to dispersal events. For both analyses, the ancestral areas estimated for *P. infans* were Zacapu Lake and Ameca River, but with low probabilities (6.5% for DEC; 16.7% for S-DIVA). The same explanation for both analyses was found for the biogeographic history of the northwest populations of *P. infans* clustered in phylogenetic clade A. The biogeographical route for this clade included a dispersal event from the Ameca River toward the Sayula and Etzatlan-San Marcos regions, followed by a vicariant event that separated the populations of Ameca River with respect to Etzatlan-San Marcos and Sayula regions. A more recent dispersal event from the Ameca to the Verde River was recovered (S-DIVA >50%, DEC >27% of probabilities).

The biogeographical history of the southeastern clade B differs slightly in both biogeographical methods. For both methods, one dispersal event occurred from Zacapu toward Cuitzeo, followed by a vicariant event that separated these two regions. For DEC, dispersal events occurred from Zacapu toward the Lower Lerma and from here to Chapala, while there was also a recent dispersal event from Cuitzeo to the Middle Lerma. The S-DIVA differed with the DEC in that dispersal events from Zacapu were toward Chapala and once both regions were separated, a second dispersal event occurred from Chapala to the Lower Lerma (S-DIVA 100% and DEC >60%) (Figs. 3 and 4).

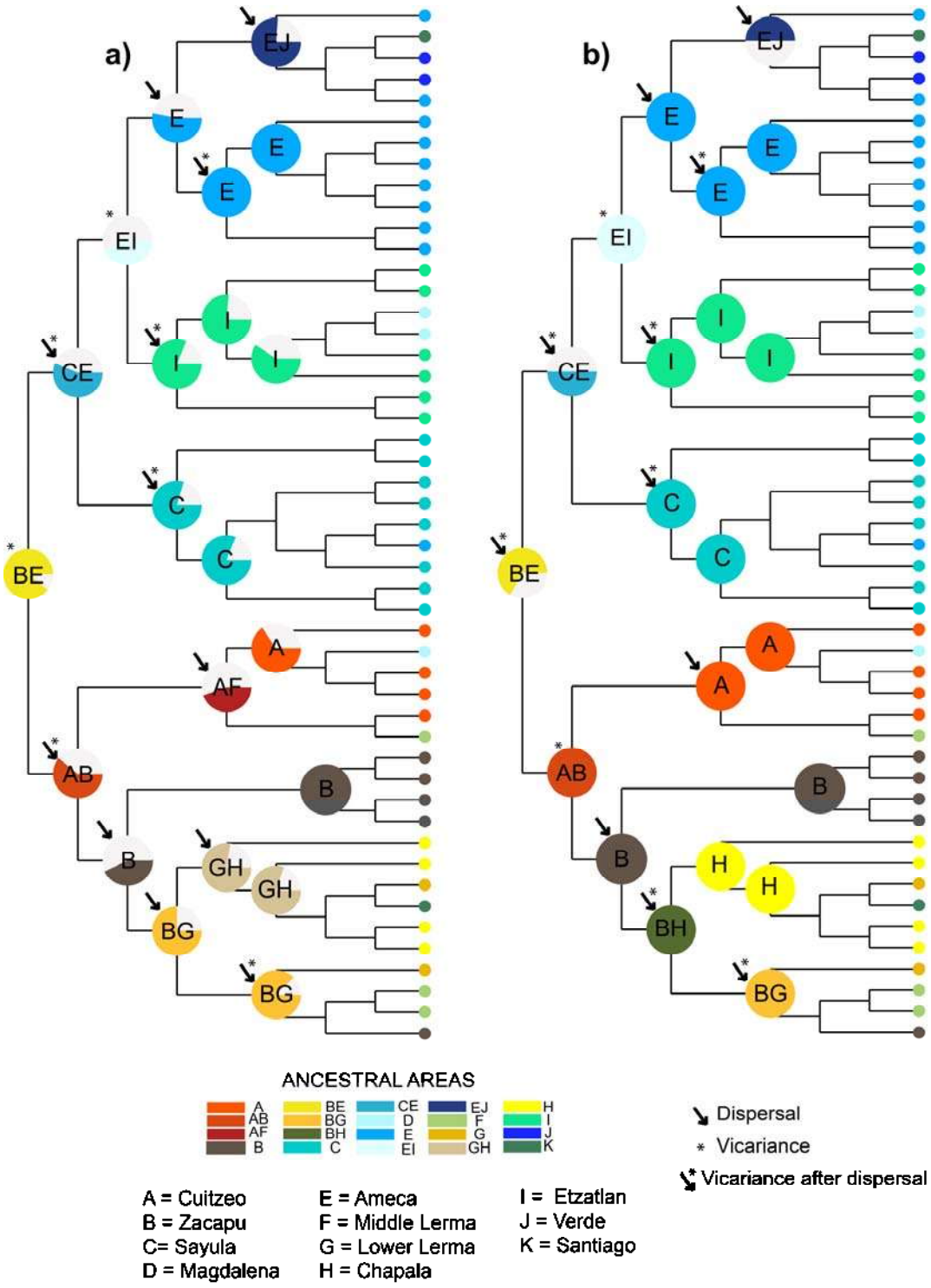


Figure 4 Ancestral area reconstruction with a) DEC and b) S-DIVA base in BEAST tree using the biogeographical regions proposed by Domínguez-Domínguez et al., (2006).

Historical Demography

The BSP analyses for *cytb* for populations of clade A showed a demographic decline for Magdalena Lake, San Marcos, Sayula and Atotonilco Lakes (belonging to Sayula region), Ameca, Grande de Santiago and Etzatlan-San Marcos regions between 0.15 and 0.1 Myrs. More recently, after a demographic decline, a population expansion was detected in the last 0.025 Myr. For the Verde River, a population reduction was detected following a more recent population expansion. For this clade, the regions Zapotlan and Tamazula were not included due to the low number of individuals (Fig. 5).

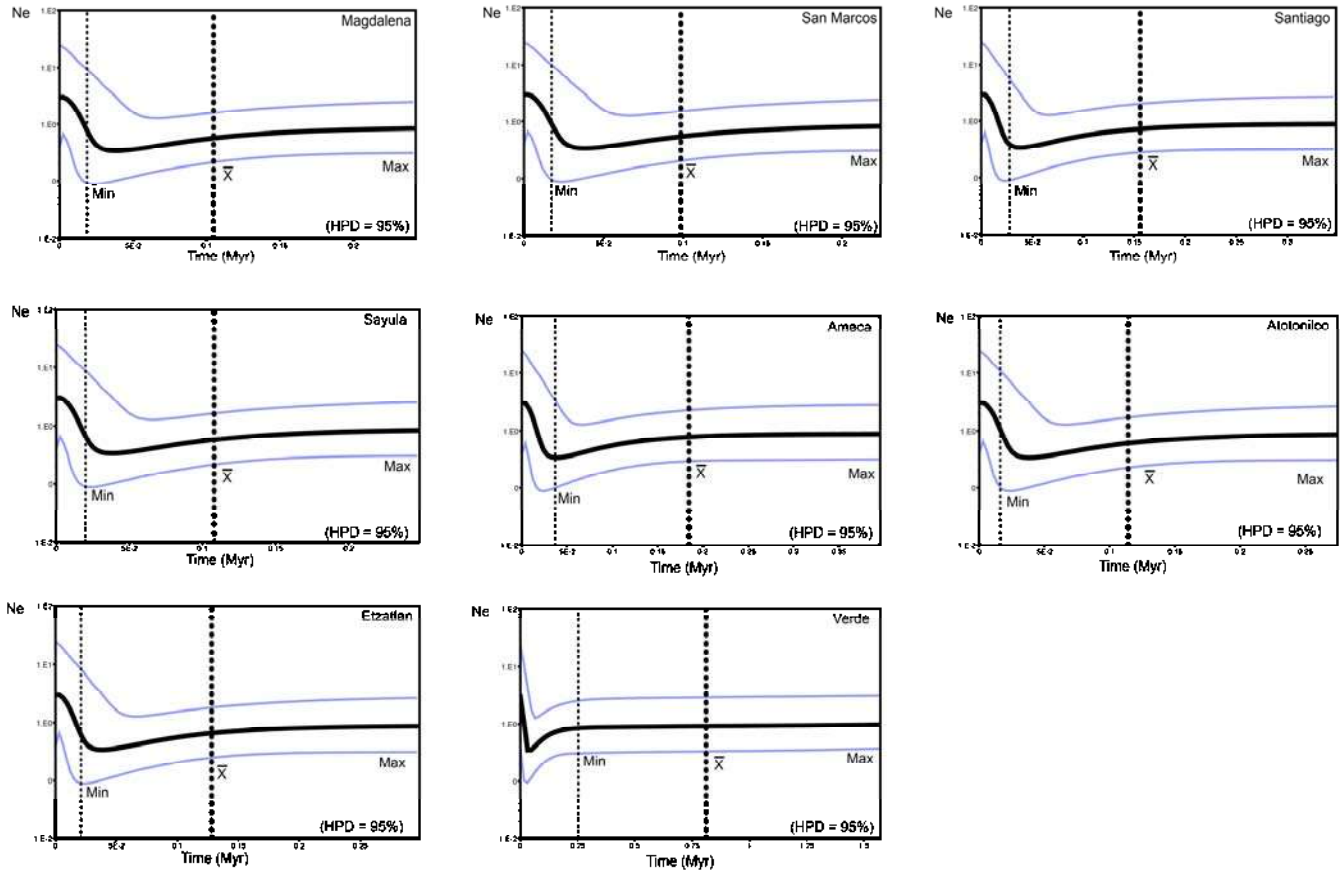


Figure 5 Demographic history of populations of each biogeographical region grouped in clade A using Bayesian Skyline Plots (BSP) from *cytb* sequences. Dotted lines represent the location of the upper bound (Max), the mean (X) and lower bound (Min) of the HPD = 95%.

For clade B, all analyzed groups revealed a demographic decline in the last 0.15-0.1 Myr, followed by population expansion around ≤0.18 Myr, as found for clade A. For this clade the Balsas, Cotija and Middle Lerma basins were not included due the low number of samples (Fig. 6).

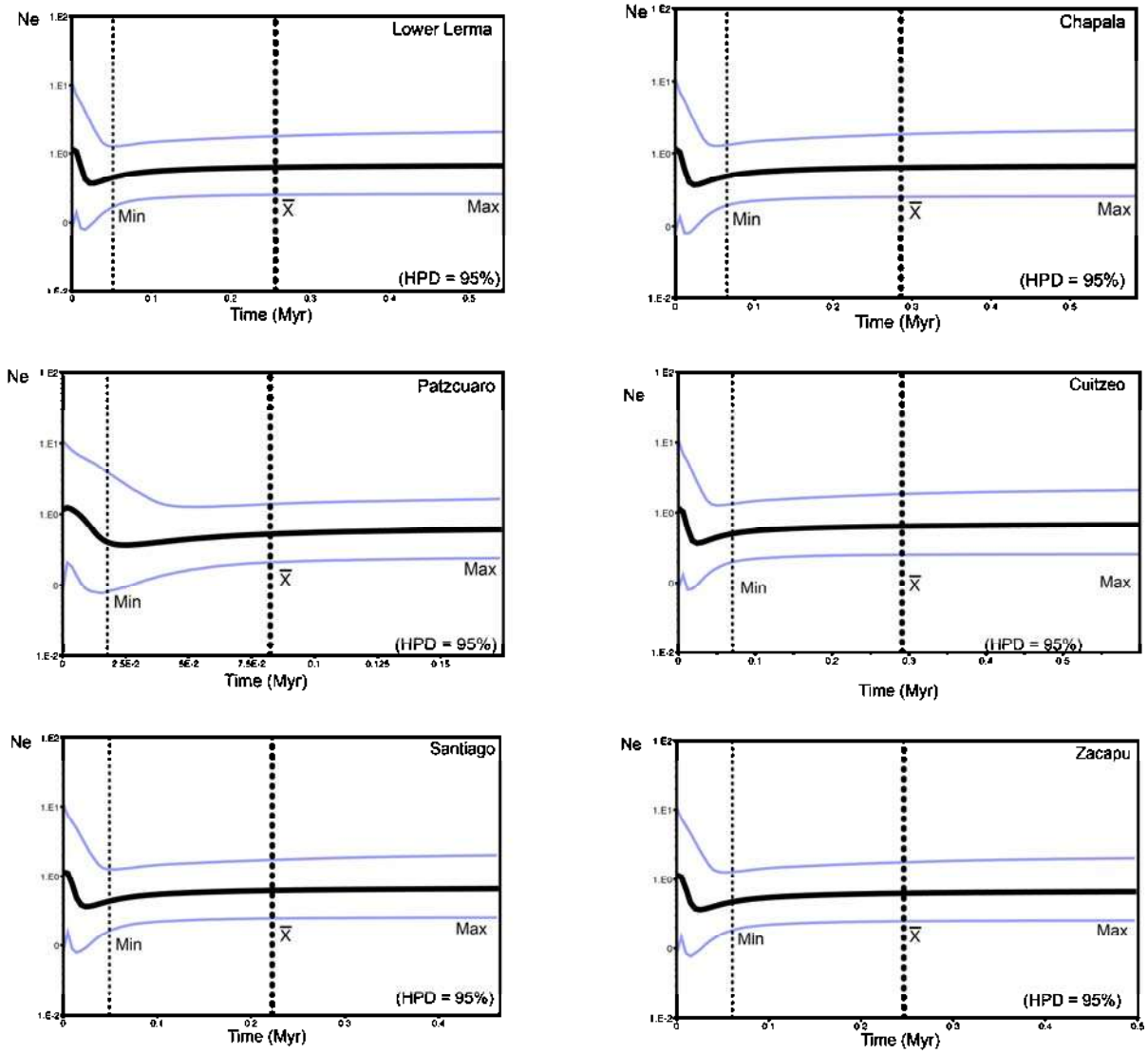


Figure 6 Demographic history of populations of each biogeographical region grouped in clade B using Bayesian Skyline Plots (BSP) from *cytb* sequences. Dotted lines represent the location of the upper bound (Max), the mean (X) and lower bound (Min) of the HPD = 95%.

Poeciliopsis infans Distribution Modelling

The habitat suitability modelling for populations of *P. infans* estimated for current (1965-1978) and past (LIG: 0.15-0.10 Myr, and LGM: 0.025-0.020 Myr) time periods, showed high precision and acceptable predictive power with all models (AUC >0.96) [69]. For the current conditions, the two variables with the highest gain were BIO3 Isothermality (BIO2/BIO7)*100 and BIO6 Min temperature of coldest month. In the LIG model, the two variables with the highest gain were BIO1 Annual Mean Temperature and BIO16 Precipitation of Wettest Quarter. For the LGM model, the two variables with the highest gain were BIO4 Temperature seasonality (standard deviation*100), and BIO6 Min Temperature of coldest month. The modelling of habitat suitability showed that in the LGM, the habitat suitability for *P. infans* was better in the lowlands, whereas for the LIG, the habitat suitability increased in the highlands (Fig. 7).

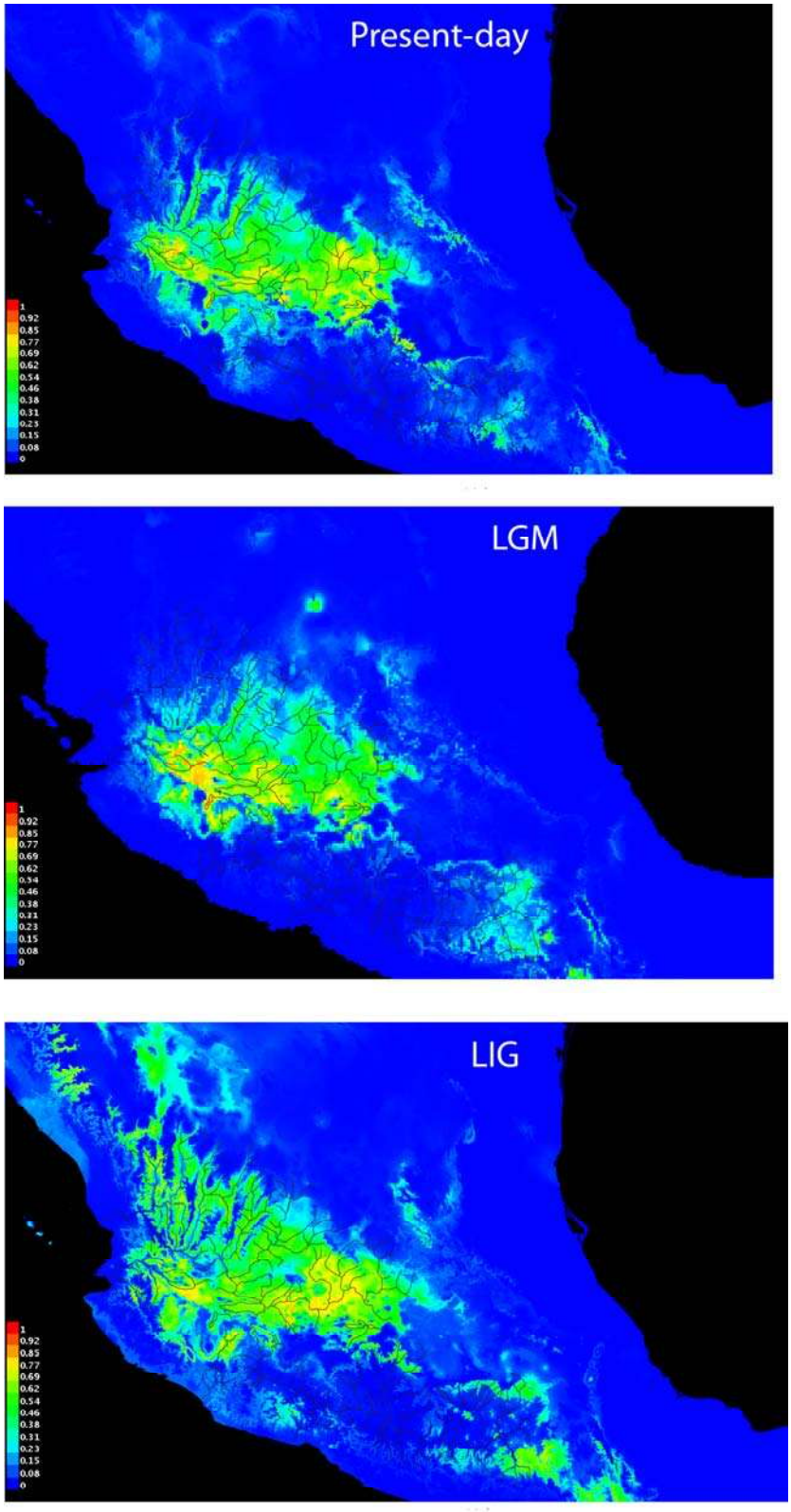


Figure 7 Species distribution modelling for: Current period, Last Inter Glacial (LIG: 0.15-0.10 Myr) and Last Glacial Maximum (LGM: 0.025-0.020 Myr), the probability of presence of the species is in scale of colors.

DISCUSSION

Biogeographic and Evolutionary History of P. infans

The recovery of two well-differentiated clades within *P. infans* (Fig 2 and Additional file 8) indicates a long history of isolation, with subsequent genetic differentiation, which seems to be linked to the intense volcanic and tectonic activity in central Mexico during the Pliocene (Figs. 2 and 3). This pattern previously has been reported, in which the genetic differentiation has been linked with the past configuration of the rivers more than the current hydrology of this region of central Mexico, as is the case for the goodeid, *Zoogoneticus quitzeoensis* (Bean, 1898), and for species of the cyprinid genus *Algansea* [30, 33].

Populations of *P. infans* are largely differentiated by the isolation of watersheds as have occurred in other freshwater fishes with low dispersal ability; however, some geographical features within a river or basin have been shown to be sufficient to differentiate populations of *P. infans* [26].

We recovered the samples from the Rio Grande de Santiago Basin grouped in the two main clades for phylogenetic analyses and in the two haplogroups for haplotype networks (A, B). This pattern, in which the samples from the Santiago River were grouped in two clades, previously was found in another study [26]. Specimens from the Rio Grande de Santiago Basin sampled upstream of the falls, el Salto de Juanacatlan, belong to clade B, whereas populations sampled downstream of the falls were grouped within clade A, suggesting that the Salto de Juanacatlan formation, a waterfall 20m high, has been an important and ancient barrier promoting differentiation between populations. This geologic feature previously has been shown to represent a barrier for fish faunal interchange between the Santiago and Chapala Lake [26, 38].

Ancestral area reconstructions recovered very low marginal probabilities for biogeographical routes from the ancestral areas of all populations of *P. infans* (DEC = 0.01; S-DIVA = 0.08; Figs. 3 and 4). Therefore, both ancestral area reconstructions were unable to resolve the state of the node separating clade A from clade B, but were able to reconstruct the ancestral areas and biogeographical routes within each clade.

Despite these low marginal probabilities, the most plausible event that separated the Ameca River and Lake Zacapu of its ancestral area of distribution was a vicariant event. Our date estimation for the diversification of the two main clades (A and B) was *ca*. 2.83 Myr (1.25-4.41 Myr), between the middle Pliocene and the early Pleistocene periods (Figs. 3 and 4).

The geological activity during the middle Pliocene and early Pleistocene in central Mexico promoted the cladogenesis of clade A from clade B. This deep divergence of the two clades is coincident with the interruption of the ancient connection of the upper Ameca River with drainages in central Mexico by tectonic and volcanic activity at *ca*. 3-1 Myr [70], (Fig. 3). As a result, this dispersal route could have been blocked at the end of the Pliocene and during the Pleistocene. This occurred when the hydrological systems that shaped the complex Chapala-Lerma Paleosystem (Ameca River, Magdalena, Chapala, Lerma River and the lakes distributed along the Colima graben) became isolated due to volcanic and tectonic activity in the triple junction area, the Ameca and San Marcos faults at *ca*. 3.5-1.5 Myr (Fig. 3) [71].

The isolation of the region where clade A is distributed during the Pliocene, has also been reported in other freshwater species including the cyprinids *Yuriria amatlana* Domínguez-Domínguez et al., 2007 and *Algansea amecae* Pérez-Rodríguez et al., 2009, as well as the goodeid *Allotoca goslinei* Smith & Miller, 1987 [29, 30, 72, 73]. In addition, other freshwater organisms shown a similar pattern, as *Cambarellus chapalanus* (Faxon, 1898), which has two divergent genetic groups, one distributed in Chapala Lake and the other in the Ameca River Basin separated at *ca*. 2.6 Myr [32]. A similar pattern was mentioned before for *Poeciliopsis* [26], who suggested that the ancestors of the strictly northern clade of *Poeciliopsis* must have been distributed across the region presently occupied by our clade B, which includes Zacapu.

The discrepancies between the mitochondrial and nuclear genes are related to the mixture of some individuals of haplogroup A with haplogroup B. In this case, considering the high genetic distances with mitochondrial genes, the divergence times, and the results from the AMOVAs, we suggested that it could be the result of retention of ancestral polymorphisms, as has been shown for other freshwater fishes of central Mexico [74].

Biogeography within Clade A

The biogeographic analyses showed a dispersal event between Sayula and the Ameca basin, suggesting an early connection more than a million years ago. This connection was previously found for *P. infans* between the Ameca River and Atotonilco Lake of the Sayula region [26].

The isolation of the Ameca, Etzatlan-San Marcos, Magdalena and Lakes of the Sayula region, could be due to the formation of the current watersheds during the Pleistocene epoch *ca*. 0.95 Myr (95% HPD: 0.39-1.5 Myr; Figs. 2 and 3), when the connections of the Ameca River and Atotonilco-Sayula Lakes were disrupted by Pleistocene volcanism and the intense tectonic activity of the so called triple junction [71, 75]. This is also congruent with the presence of *Ameca splendens* Miller & Fitzsimons, 1971, in the Ameca River and Sayula regions, with a divergence time between the two populations calculated in less than a million of years [29].

Other dispersal events from the Ameca to Etzatlan-San Marcos and Magdalena regions were recovered. The climatic changes during this pluvial-interpluvial period, beginning *ca*. 0.90 Myr [4, 76], could have promoted this dispersal event when Magdalena Lake was considerably larger and extended to Etzatlan-San Marcos area [77]. After that time, the Ameca was isolated from the Etzatlan-San Marcos and Magdalena regions by a vicariant event *ca*. 0.72 Myr (DEC = 0.209; S-DIVA = 1.0), which could be associated with the reactivation of the Plan de Barrancas fault during the Quaternary (*ca*. 1.0 Myr), and with the Amatlán de Cañas half graben (Fig. 3) that was formed *ca.* 3.4 Myr [13, 70, 78].

Finally, a dispersal event from the Ameca River to the Verde and Santiago Rivers also was found (*ca*. 0.14 Myr), and this is supported by previous findings suggesting that these populations have been connected until very recently through stream capture of the Ameca and Verde Rivers, which was facilitated by the volcanism in the Tepic-Zacoalco graben [26, 79], (Fig. 3).

The biogeographic events that isolated the three recovered sub-clades are supported by the AMOVA analysis, maximizing the Φ_{CT} when samples were grouped into four groups, included the three sub-clades within clade A (Tables 5 and 6), but not when they were grouped according to biogeographic regions as have been proposed for other freshwater fishes, including goodeids and cyprinids [33, 36].

Since goodeids, cyprinids, and *P. infans* have evolved in spatiotemporal congruence, the differences found in the evolutionary history of *P. infans* could be related to the biogeographic origin of each group. *Poeciliopsis infans* has a Neotropical origin, whereas goodeids and cyprinids are of Neartic origin [18]. As a result, we expect that Quaternary climate changes have influenced genetic variation and the distributional patterns of *P. infans.* Moreover, the genetic diversity was higher for regions clustered in the clade A (lowlands) than for regions grouped in clade B (highlands). This could also be linked to more stable high temperatures in lowland areas, which could also promote the diversification of populations found for this lowland clade (clade A). This pattern has been reported in plants that shown that the habitat and environment changes affect the genetic diversity [80-83], as is the case of *Caragana microphylla*, a species distributed in two different habitats that shown that populations from the high temperature region had lower genetic diversity than those from medium and low temperature regions [81]. The only exception of this pattern is the Magdalena Lake population, which has the lowest genetic diversity within this clade, a pattern that is explained due to the history of instability of this hydrological basin, with extreme and intermittent periods of flooding and drying [84], promoting recurrent events of bottlenecks and loss of genetic diversity.

Also, other factors, in addition to the environment, such as life-history traits, breeding systems, dispersal mechanism, geographic variation, range and life span and the histories of populations have affected genetic variation between populations [81-83].

Biogeography within Clade B

Our results are in accordance with the proposed connections and isolation between the Cuitzeo and Zacapu regions, which has occurred several times during the Pleistocene [85]. The connection between both regions has been postulated to occur through the Chucandiro-Huaniqueo paleo-river, a connection disrupted less than 1 Myr, due to the volcanism of the Tarasco corridor and the activity of the Northeast-Southeast fault system *ca.* 0.7 to 0.5 Myr [86], (Fig. 3). This disruption has been proposed as the cause of the isolation of different fish species between Cuitzeo and Zacapu including the goodeids *Skiffia lermae* (Meek, 1902), *Goodea atripinnis* (Jordan, 1880),

Alloophorus robustus (Bean, 1892) and *Hubbsina turneri* (de Buen, 1940) [29, 85], and the cyprinids *Algansea tincella* (Valenciennes, 1844) and *Yuriria alta* (Jordan, 1880) [30, 72].

After the isolation of Zacapu and Cuitzeo, the DEC and S-DIVA showed a dispersal event from Cuitzeo towards the middle Lerma River, an event that is in accordance with recent isolation of this hydrological basin with respect to Zacapu and Cuitzeo lakes. This event was previously proposed for the goodeids *Hubbsina turneri* [85], and *Zoogoneticus quitzeoensis* (Bean, 1898), that are distributed in Cuitzeo, Zacapu, and the middle Lerma [33], as well as *Neotoca bilineata* (Bean, 1887) for which a low level of genetic differentiation (mtDNA) was found for populations from Cuitzeo and the middle Lerma [87].

Regarding the second biogeographic route for this clade, which is different for the DEC and S-DIVA analyses, the most plausible route is a dispersal event from Zacapu toward the lower Lerma Basin, followed by a dispersal event toward Chapala Lake and the Santiago River respectively, as is found in DEC. Zacapu is currently connected with the lower Lerma through the Angulo River, for which *P. infans* could have dispersed toward Chapala Lake and the Santiago River via the lower Lerma. This information is congruent with the presence of *Notropis calientis* (Jordan & Snyder, 1899), *Yuriria alta*, *Algansea tincella*, *Xenotoca variata* (Bean, 1887), *Chapalichthys encaustus* (Jordan & Snyder, 1899), *Allotoca dugesii* (Bean, 1887) and *Moxostoma austrinum* Bean, 1880, in the Middle Lerma, Lower Lerma and Grande de Santiago basins, as well as in Chapala Lake region [35, 36].

Human Mediated Dispersion

In general, we found biogeographic correspondence in the distribution of clades, but incongruences were also found for some populations and areas, as is the case of Balsas, Cotija and Patzcuaro, populations that shown a null genetic and haplotype diversities, with all sampled individuals belong to the most common haplotype of clade B.

These results have two possible explanations: 1) a secondary dispersal and colonization event, which is unlikely due to the historical isolation of all of those drainages with respect to contiguous drainages [29, 88], or 2) a founder effect due a dispersal event mediated by humans. Human-based introductions

represent the most probable explanation according to the geographic distribution of related haplotypes and the null genetic diversity of Balsas, Cotija and Patzcuaro regions, since for Patzcuaro this species has been reported as a human-mediate introduction [89, 90] (Tables 3 and 4).

Several species in the family Poeciliidae have been introduced for mosquito control worldwide and have spread successfully to over 40 countries [91]. Other possible ways of introduction of poeciliids is the release of the organisms by aquarists, through the use of this fish as food source for commercial introduced fish, or accidentally transported with commercially important fishes stocked into water bodies, such a species of tilapia (*Oreochromis* and *Tilapia*), which have been widely introduced throughout Mexico [89, 92].

Historical Demography and Distribution Modeling

Fluctuations in population size shown by BSP in populations of both clades of *P. infans* agree with the continuous fluctuations of the climate and water levels of hydrological systems in central Mexico due to glacial and interglacial cycles [93].

For clade A, the analysis of historical demography showed a demographic decline at *ca* 0.15-0.1 Myr, followed by a recent population expansion, estimated to start around 0.025 Myr in most of the populations analyzed (Fig. 5). These genetic based analyses are congruent with the distribution modeling results, in which drainages where clade A is distributed, show restricted areas with high probabilities (≤0.77) to support populations of *P. infans* during the LIG (0.15-0.10 Myr), localized mainly to a small area within the Ameca region (Fig. 7). Whereas for the LGM (0.025-0.020 Myr), an increase in the areas with high probabilities (≤0.77) to support populations of *P. infans* is observed, covering most of the present day distribution of clade A populations. This recent population expansion for almost all populations of clade A could explain why the genetic diversity is highest in this clade rather than in the clade B. It is well known that stable populations that persisted from the LGM to the present harbor disproportionately large amounts of unique genetic diversity [94].

The decline in the clade B population seems to starts after 0.075 Myr, followed by a population expansion at *ca.* ≤0.018 Myr in all the populations

analyzed, however, we take with caution the population expansion of some biogeographic regions for both clades, because the size increase is out of the HPD limits (Fig. 6).

These results are also congruent with the distribution modeling results, since the distribution modeling during LIG (0.15-0.10 Myr) also showed a high proportion of areas where clade B is distributed with high probabilities (≤ 0.77) to support populations of *P. infans*. Whereas for the LGM (0.025-0.020 Myr), a decrease in the area with high probability of presence (≥ 0.77) is observed for areas were clade B is distributed. These results are expected for a species with tropical preferences inhabiting highlands, where temperatures declines of 8.5ºC during LGM have been postulated [95], followed by a expansion of the distribution range when the last ice age ended and an increase of the temperature and water level of hydrological systems have been recorded [93]. It has been shown that the climatic fluctuations were accompanied by a loss of genetic diversity and even extinction of populations that were unable to adapt to these changes and find suitable conditions [96], as could be the case for *P. infans* that is restricted to basin drainages. This is congruent with the distribution modeling of clade B distributed in highlands of central Mexico, and explains the low genetic diversity found in almost all populations within this clade.

Finally, in the present day modeling that represents an Inter-glacial period, it showed an extended distributional area with a probability of presence ≥ 0.77 in the upper parts of the distributional range for both clades A and B. These results of population expansion in lower areas during Glacial Maximum and in upper areas during Inter Glacial periods are also congruent with the Neotropical origin of *P. infans*, suggesting that areas that maintain high temperature are more suitable for *P. infans*. This is also congruent with reports suggesting the displacement of plant communities as a response to the climate cooling, resulting in the migration to low altitudes [97], for which the climate change in glacial cycles could be caused by modifications in the composition of the communities favoring more resistant species to the new environmental condition [93, 97].

Conservation Implications

The regions occupied by *P. infans* have been heavily impacted by habitat loss due to overexploitation, pollution, habitat degradation, and the introduction of non-native species [98-100]. The negative impacts of these anthropogenic activities on native freshwater fishes of central Mexico are well known [33, 101-103]. These activities can exacerbate the loss of genetic diversity, which is especially harmful for a species with a long history of instability due to the natural climatic fluctuations, as is the case of *P. infans*. This is especially true for those populations distributed in clade B (Tables 3 and 4). For this reason, the establishment of each recovered clade and sub-clade as Operational Conservation Unit (OCU) [104] is necessary in order to conserve the unique genetic pool that each clade represents.

CONCLUSIONS

The results of this study indicate that *P. infans* has had a long history of isolation and subsequent genetic differentiation, which appears to be linked to the intense volcanic and tectonic activity in central Mexico. The separation of the two recovered clades appears to have been promoted by the geological activity during the middle Pliocene, and early Pleistocene in central Mexico, *ca* 2.83 Myr. *Poeciliopsis infans* is a co-distributed species with other group of fishes as goodeids and cyprinids, evolving in spatiotemporal congruence. The differences in the recovered biogeographic patterns of *P. infans* is likely related to the biogeographic origin of each group, since *P. infans* has Neotropical origin, and goodeids and cyprinids are of Neartic origin. Populations of *P. infans* distributed in lowlands showed a higher level of genetic diversity than populations distributed in highlands, which could be linked to more stable and higher temperatures in lowland areas. Finally, fluctuations in population size are supported by the continuous fluctuations of the climate and water levels of hydrological systems in central Mexico due to glacial and interglacial cycles.

Additional files

Additional file 1: Tissue voucher number, and access number of GenBank.

Site	Locality	Biogeographic region	Voucher number	Access number of GenBank			
				Cytb	*coxI*	*RHO*	*S7*
1	Los Venados	Magdalena	9429	MG028024	MG028296	MG100624	MG366198
1			9429B	---	---	NO	MG366199
1			9431	MG028025	MG028297	MG100625	MG366200
1			9431B	---	---	MG100626	MG366201
1			9432	MG028026	MG028298	NO	NO
1			9434	MG028027	MG028299	MG100627	NO
1			9434B	---	---	MG100628	NO
1			9435	MG028028	MG028300	MG100629	MG366202
2	Laguna	Magdalena	4600	MG028023	MG028295	MG100622	MG366196
2			4600B	---	---	MG100623	MG366197
2			9959	MG028029	MG028301	MG100630	MG366203
2			11695	MG028030	MG028302	NO	MG366204
2			11696	MG028031	NO	NO	NO
2			11697	MG028032	MG028303	NO	MG366205
2			11697B	---	---	NO	MG366206
2			11698	MG028033	MG028304	MG100631	MG366207
2			11699	MG028034	MG028305	MG100632	MG366208
2			11699B	---	----	NO	MG366209
2			11700	MG028035	MG028306	MG100633	MG366210
2			28074	MG028036	NO	NO	MG366211
2			28074B	---	---	NO	MG366212
2			28075	MG028037	MG028307	NO	MG366213
2			28076	MG028038	NO	MG100634	MG366214
3	Presa San Ignacio	Ameca	4632	MG028039	MG028308	MG100635	MG366215
3			4632B	---	---	NO	MG366216
3			42907	MG028070	MG028334	MG100663	NO
3			42908	MG028071	MG028335	NO	NO
3			42909	MG028072	MG028336	MG100664	NO
3			42910	MG028073	MG028337	MG100665	NO
3			42911	MG028074	NO	MG100666	MG366254
3			42927	MG028075	MG028338	MG100667	MG366255

4	Chapulimita	Ameca	4654	MG028040	MG028309	MG100636	MG366217
4			4654B	---	---	MG100637	MG366218
4			4655	MG028041	MG028310	MG100638	MG366219
4			4655B	---	---	NO	MG366220
5	Salida presa Tecuan	Ameca	5049	MG028042	MG028311	MG100639	MG366221
5			5049B	---	---	NO	MG366222
5			5050	MG028043	NO	MG100640	MG366223
5			5051	MG028044	MG028312	MG100641	MG366224
5			5051B	---	---	NO	MG366225
5			5052	MG028045	MG028313	MG100642	MG366226
5			5052B	---	---	NO	MG366227
5			5053	MG028046	MG028314	NO	NO
5			5054	MG028047	MG028315	MG100643	MG366228
5			5054B	---	---	NO	MG366229
5			5059	MG028048	MG028317	NO	NO
6	Manantial Los Veneros	Ameca	9920	MG028049	MG028318	MG100644	MG366230
6			9920B	---	---	NO	MG366231
6			9921	MG028050	MG028319	MG100645	MG366232
7	Tala, Río Salado	Ameca	11841	MG028051	MG028320	MG100646	MG366233
7			11842	MG028052	MG028321	MG100647	NO
7			11843	MG028053	MG028322	MG100648	NO
8	Amatlán de cañas	Ameca	11979	MG028054	MG028323	MG100649	MG366234
8			11979B	---	---	NO	MG366235
8			11980	MG028055	MG028324	MG100650	MG366236
8			11980B	---	---	MG100651	MG366237
8			11981	MG028056	NO	MG100652	MG366238
8			11981B	---	---	MG100653	MG366239
8			11982	MG028057	MG028325	NO	MG366240
8			11982B	---	---	NO	MG366241
8			11984	MG028058	MG028326	MG100654	MG366242
8			11985	MG028059	MG028327	MG100655	MG366243
8			11985B	---	---	NO	MG366244
9	Teuchitlán	Ameca	28085	MG028060	NO	NO	NO
9			28086	MG028061	NO	NO	NO
9			28087	MG028062	NO	NO	MG366245
9			28088	MG028063	NO	NO	NO
9			28089	MG028064	MG028328	MG100656	MG366246
9			28089B	---	---	NO	MG366247
9			31540	MG028065	MG028329	MG100657	MG366248
9			31540B	---	---	MG100658	NO
9			31541	MG028066	MG028330	MG100659	MG366249

9			31541B	---	---	MG100660	MG366250
9			31542	MG028067	MG028331	MG100661	MG366251
9			31543	MG028068	MG028332	MG100662	MG366252
9			31543B	---	---	NO	MG366253
9			31544	MG028069	MG028333	NO	NO
10	S. M. San Julian	Verde	4672	MG028224	MG028487	NO	MG366429
10			4673	MG028225	MG028488	MG100780	NO
10			4674	MG028226	MG028489	MG100781	MG366430
10			4674B	---	---	MG100782	NO
11	San Nicolás	Verde	35449	NO	MG028490	NO	NO
11			35450	MG028227	MG028491	MG100783	NO
11			35451	MG028228	MG028492	MG100784	NO
11			35452	NO	MG028493	NO	NO
11			35453	MG028229	MG028494	MG100785	MG366431
11			35454	MG028230	MG028495	MG100786	NO
11			35455	MG028231	MG028496	MG100787	NO
11			35456	MG028232	MG028497	MG100788	MG366432
11			35456B	---	---	NO	MG366433
11			35457	MG028233	MG028498	MG100789	MG366434
11			35457B	---	---	NO	MG366435
11			35458	MG028234	MG028499	NO	NO
11			35459	MG028235	MG028500	NO	NO
11			36260	MG028242	MG028507	MG100796	MG366437
11			36260B	---	---	NO	MG366438
11			36261	MG028243	MG028508	MG100797	MG366439
11			36261B	---	---	NO	MG366440
12	Arroyo La Estancia	Verde	35460	MG028236	MG028501	MG100790	MG366436
12			35461	MG028237	MG028502	MG100791	NO
13	Río Colorado	Verde	36256	MG028238	MG028503	MG100792	NO
13			36257	---	---	MG100793	NO
13			36258	MG028240	MG028505	MG100794	NO
13			36259	MG028241	MG028506	MG100795	NO
14	Río Xoconostle-San Juan	Medio Lerma	4948	MG028076	MG028339	MG100668	YES
14			4948B	---	---	MG100669	MG366256
14			4949	MG028077	MG028340	MG100670	MG366257
14			12810	MG028078	MG028341	MG100671	MG366258
15	Manantial Andrés-Figueroa	Sayula	5023	MG028079	MG028342	NO	MG366259
15			5024	NO	MG028343	NO	MG366260
15			5026	MG028080	MG028344	MG100672	MG366261
15			5026B	---	---	MG100673	NO
15			5027	MG028081	MG028345	NO	MG366262

15			5044	MG028082	MG028346	MG100674	NO
15			5044B	---	---	MG100675	NO
15			5047	MG028083	MG028347	MG100676	MG366263
15			18962	MG028104	MG028369	MG100692	MG366292
15			18963	MG028105	MG028370	NO	NO
15			18964	MG028106	MG028371	MG100693	MG366293
15			18964B	---	---	MG100694	NO
16	Manantial San Marcos	Sayula	28115	MG028109	MG028375	MG100699	MG366300
16			28116	MG028110	MG028376	MG100700	MG366301
16			28116B	---	---	MG100701	NO
16			28117	MG028111	MG028377	MG100702	MG366302
16			28118	MG028112	MG028378	MG100703	MG366303
16			28119	MG028113	MG028379	MG100704	MG366304
17	Canal Presa Buena Vista	Sayula	8360	MG028085	MG028349	MG100678	MG366266
17			8360B	---	---	NO	MG366267
17			8361	NO	MG028350	NO	NO
17			8362	MG028086	MG028351	NO	MG366268
17			8363	MG028087	MG028352	NO	MG366269
17			8363B	---	---	NO	MG366270
17			8364	MG028088	MG028353	MG100679	MG366271
17			8364B	---	---	NO	MG366272
17			8365	MG028089	MG028354	NO	MG366273
17			8366	MG028090	MG028355	MG100680	MG366274
17			8366B	---	---	NO	MG366275
18	Villa corona	Sayula	5070	MG028084	MG028348	MG100677	MG366264
18			5070B	---	---	NO	MG366265
18			28109	MG028107	MG028373	MG100695	MG366296
18			28109B	---	---	MG100696	MG366297
18			28110	MG028108	MG028374	MG100697	MG366298
18			28110B	---	---	MG100698	MG366299
19	Manantial Cuyacapán	Sayula	10565	MG028091	MG028356	MG100681	MG366276
19			10565B	---	---	NO	MG366277
19			10566	MG028092	MG028357	MG100682	MG366278
19			10566B	---	---	NO	MG366279
19			10567	MG028093	MG028358	NO	MG366280
19			10567B	---	---	NO	MG366281
19			10568	MG028094	MG028359	MG100683	MG366282
19			10568B	---	---	NO	MG366283
19			10570	MG028095	MG028360	MG100684	MG366284
19			10571	MG028096	MG028361	MG100685	MG366285
19			10571B	---	---	NO	MG366286

19			10572	MG028097	MG028362	MG100686	NO
19			10573	MG028098	MG028363	MG100687	MG366287
19			10573B	---	---	MG100688	MG366288
19			10574	MG028099	MG028364	NO	NO
19			10575	MG028100	MG028365	MG100689	NO
19			10577	MG028101	MG028366	NO	NO
19			10578	MG028102	MG028367	MG100690	MG366289
19			10578B	NO	NO	MG100691	MG366290
19			42833	MG028115	MG028381	NO	NO
19			42834	NO	MG028382	NO	NO
19			42835	MG028116	MG028383	MG100705	NO
19			42836	NO	MG028384	NO	NO
19			42932	MG028117	MG028385	MG100706	MG366305
20	Laguna de Zapotlán	Sayula	14095	MG028103	MG028368	NO	MG366291
20			42814	MG028114	MG028380	NO	NO
21	Río Las Puentes	Chapala	17979	NO	MG028456	MG100749	MG366382
21			17979B	---	---	MG100750	NO
21			17980	NO	MG028457	NO	NO
21			17981	MG028189	MG028458	NO	MG366383
21			17981B	---	---	NO	MG366384
21			17982	MG028190	MG028459	MG100751	MG366385
21			17982B	---	---	NO	MG366386
21			17983	MG028191	NO	MG100752	MG366387
21			17984	MG028192	MG028460	MG100753	MG366388
21			17993	MG028193	MG028461	MG100754	MG366389
21			17993B	---	---	NO	MG366390
21			17995	MG028194	MG028462	MG100755	MG366391
21			17995B	---	---	NO	MG366392
22	Cojumatlán	Chapala	28130	MG028195	NO	NO	NO
22			28131	MG028196	MG028463	NO	NO
22			28132	MG028197	MG028464	MG100756	NO
22			28132B	---	---	MG100757	
23	Los Negritos	Chapala	28186	MG028198	NO	NO	MG366393
23			28189	MG028199	NO	NO	MG366394
23			28189B	---	---	NO	MG366395
23			28191	MG028200	MG028465	NO	MG366396
23			28192	MG028201	MG028466	MG100758	MG366397
23			28193	MG028202	MG028467	MG100759	MG366398
23			28193B	---	---	NO	MG366399
24	Presa Nueva	Chapala	28264	MG028203	MG028468	MG100760	MG366400
24			28264B	---	---	NO	MG366401

24			28265	MG028204	MG028469	NO	MG366402
24			28265B	---	---	NO	MG366403
24			28266	MG028205	MG028470	MG100761	MG366404
24			28266	---	---	NO	MG366405
24			28267	MG028206	MG028471	MG100762	MG366406
24			28267B	---	---	NO	MG366407
24			28268	MG028207	MG028472	MG100763	MG366408
24			28268B	---	---	NO	MG366409
25	Manantial La Mintzita	Cuitzeo	10153	MG028118	MG028386	MG100707	MG366306
25			10154	MG028119	MG028387	MG100708	MG366307
25			10154B	---	----	NO	MG366308
25			10155	MG028120	MG028388	NO	MG366309
25			10155B	---	---	NO	MG366310
25			10156	MG028121	MG028389	NO	NO
25			10157	MG028122	MG028390	NO	MG366311
25			10164	MG028123	MG028391	MG100709	MG366312
25			10164B	---	---	NO	MG366313
26	Ojo de Agua San Cristóbal	Cuitzeo	10359	MG028124	MG028392	NO	MG366314
26			10359B	---	---	NO	MG366315
26			10360	MG028125	MG028393	MG100710	NO
26			10361	MG028126	MG028394	MG100711	NO
26			10362	MG028127	MG028395	MG100712	NO
26			10364	MG028128	MG028396	MG100713	MG366316
26			10365	MG028129	MG028397	MG100714	MG366317
26			10366	NO	MG028398	NO	MG366318
26			10369	MG028130	MG028399	NO	MG366319
26			10369B	---	---	NO	MG366320
26			10370	MG028131	MG028400	MG100715	MG366321
26			10372	MG028132	MG028401	NO	MG366322
27	Embarcadero Principal	Patzcuaro	10265	MG028133	MG028402	NO	MG366323
27			10266	MG028134	NO	NO	MG366324
27			10267	NO	MG028403	NO	NO
27			10286	MG028135	MG028404	NO	MG366325
27			10287	MG028136	MG028405	MG100716	MG366326
2			10288	MG028137	MG028406	NO	NO
27			10289	NO	MG028407	NO	NO
27			10290	MG028138	MG028408	NO	MG366327
28	Urandén	Patzcuaro	23906	MG028139	MG028409	MG100717	MG366328
28			23906B	---	---	MG100718	
28			26806	MG028140	MG028410	MG100719	MG366329
28			26807	MG028141	MG028411	NO	NO

28			26808	MG028142	MG028412	MG100720	MG366330
28			26809	MG028143	NO	NO	MG366331
28			26810	MG028144	MG028413	NO	MG366332
28			26811	MG028145	MG028414	NO	NO
28			26812	MG028146	MG028415	MG100721	MG366333
29	Presa Melchor Ocampo	Zacapu	10431	MG028147	MG028416	MG100722	MG366334
29			10431B	---	---	MG100723	NO
29			10437	MG028148	MG028417	NO	NO
29			10439	MG028149	MG028418	MG100724	MG366335
29			10440	MG028150	MG028419	NO	MG366336
29			10441	MG028151	MG028420	NO	MG366337
29			10443	MG028152	NO	NO	MG366338
29			10444	MG028153	MG028421	NO	MG366339
29			10445	NO	MG028422	NO	NO
30	La Zarcita	Zacapu	25106	MG028155	MG028424	MG100725	MG366342
30			25106B	---	---	NO	MG366343
30			25107	MG028156	MG028425	MG100726	NO
30			25109	MG028157	MG028426	MG100727	MG366344
30			25110	MG028158	MG028427	MG100728	MG366345
30			25111	NO	MG028428	NO	MG366346
30			25112	MG028159	MG028429	NO	MG366347
30			25113	MG028160	MG028430	NO	MG366348
30			25114	MG028161	MG028431	NO	MG366349
30			25115	MG028162	MG028432	NO	MG366350
30			25116	MG028163	MG028433	NO	MG366351
31	Laguna de Zacapu	Zacapu	25102	MG028154	MG028423	NO	MG366340
31			25102B	---	---	NO	MG366341
31			26801	MG028164	MG028434	MG100729	MG366352
31			26802	MG028165	MG028435	NO	MG366353
31			26803	MG028166	MG028436	NO	MG366354
31			26803B	---	---	NO	MG366355
31			26804	MG028167	MG028437	NO	MG366356
31			26804B	---	---	NO	MG366357
32	Atenquique	Tamazula	12012	MG028188	MG028455	MG100748	NO
33	Puente en Jacona	Bajo Lerma	17964	MG028266	MG028531	MG100815	NO
33			17965	MG028267	MG028532	MG100816	MG366468
33			17975	MG028268	MG028533	NO	MG366469
33			17975B	---	---	NO	MG366470

33			17976	MG028269	MG028534	MG100817	MG366471
33			17976B	---	---	NO	MG366472
33			17977	MG028270	MG028535	MG100818	MG366473
34	Presa La Luz	Bajo Lerma	39556	MG028273	MG028536	MG100819	MG366475
34			39557	MG028271	MG028537	MG100820	MG366476
34			39558	MG028272	MG028538	MG100821	MG366477
34			39559	NO	MG028539	NO	MG366478
34			39560	MG028276	MG028540	NO	NO
34			39561	MG028275	MG028541	NO	NO
34			39562	MG028274	MG028542	NO	MG366479
34			39563	MG028277	MG028543	MG100822	MG366480
34			39563B	---	---	NO	MG366481
34			39564	MG028278	NO	MG100823	MG366482
34			39564B	---	---	MG100824	NO
35	Quitupan	Balsas	18739	MG028208	MG028473	MG100764	MG366410
35			18740	MG028209	MG028474	NO	MG366411
35			18740B	---	---	NO	MG366412
35			18745	MG028210	MG028475	NO	MG366413
35			18746	MG028211	MG028476	MG100765	MG366414
35			18747	MG028212	MG028477	MG100766	MG366415
35			18747B	---	---	MG100767	NO
36	Presa San Juanico	Cotija	28133	MG028213	MG028478	NO	MG366416
36			28134	MG028214	MG028479	MG100768	MG366417
36			28134B	---	---	MG100769	NO
36			28135	MG028215	MG028480	MG100770	MG366418
36			28136	MG028216	NO	NO	MG366419
36			28136B	---	---	NO	MG366420
36			28137	MG028217	NO	MG100771	MG366421
36			28137B	---	---	NO	MG366422
36			28138	MG028218	MG028481	MG100772	MG366423
36			28139	MG028219	MG028482	MG100773	MG366424
36			28139B	---	---	MG100774	NO
36			28140	MG028220	MG028483	MG100775	MG366425
36			28164	MG028221	MG028484	MG100776	MG366426
36			28164B	---	---	MG100777	NO
36			28165	MG028222	MG028485	MG100778	MG366427
36			28166	MG028223	MG028486	MG100779	MG366428
37	San Sebastián	Etzatlan-San Marcos	11920	MG028168	MG028438	MG100730	MG366358
37			11920B	---	---	MG100731	MG366359
37			11921	MG028169	MG028439	NO	MG366360
37			11921B	---	---	NO	MG366361

38	Presa Palo Verde	Etzatlan-San Marcos	32485	MG028175	MG028443	MG100735	MG366371
38			32485B	---	---	NO	MG366372
38			32486	MG028176	MG028444	MG100736	MG366373
38			32487	MG028177	MG028445	MG100737	NO
38			32488	MG028178	MG028446	MG100738	MG366374
38			32490	MG028179	MG028447	MG100739	MG366375
38			32491	MG028180	MG028448	MG100740	MG366376
38			32492	MG028181	MG028449	MG100741	MG366377
38			32493	MG028182	MG028450	MG100742	MG366378
38			32494	MG028183	NO	NO	NO
38			32495	MG028184	MG028451	MG100743	MG366379
38			32495B	---	---	MG100744	NO
38			32577	MG028185	MG028452	MG100745	MG366380
38			32579	MG028186	MG028453	MG100746	NO
38			32580	MG028187	MG028454	MG100747	MG366381
39	San Juanito de Escobedo	Etzatlan-San Marcos	28171	MG028170	MG028440	NO	MG366362
39			28171B	---	---	NO	MG366363
39			28172	MG028171	NO	MG100732	MG366364
39			28173	MG028172	MG028441	MG100733	MG366365
39			28173B	---	---	NO	MG366366
39			28174	MG028173	NO	NO	MG366367
39			28174B	---	---	NO	MG366368
39			28175	MG028174	MG028442	MG100734	MG366369
39			28175B	---	---	NO	MG366370
40	Cuescomatitlán	Grande de Santiago	5094	MG028244	MG028509	MG100798	MG366441
40			5094B	---	---	MG100799	MG366442
40			5095	MG028245	MG028510	NO	MG366443
40			5096	MG028246	MG028511	NO	MG366444
40			5097	MG028247	MG028512	MG100800	MG366445
40			5097B	---	---	NO	MG366446
40			5098	MG028248	MG028513	NO	MG366447
40			5098B	---	---	NO	MG366448
41	Jalpa	Grande de Santiago	5769	MG028249	MG028514	NO	MG366449
41			20875	MG028250	MG028372	MG100801	MG366450
41			20875B	---	---	MG100802	NO
42	San Antonio	Grande de Santiago	36290	MG028251	MG028515	MG100803	MG366451
42			36290B	---	---	MG100804	MG366452
42			36291	MG028252	MG028516	MG100805	MG366453
42			36291B	---	---	MG100806	MG366454
43	Presa de Garabato	Grande de Santiago	36314	MG028253	MG028517	MG100807	MG366455
43			36314B	---	---	NO	MG366456

43			36315	MG028254	MG028518	NO	MG366457
43			36315B	---	---	NO	MG366458
43			36319	MG028255	MG028519	MG100808	MG366459
43			36319B	---	---	NO	MG366460
43			36320	MG028256	MG028521	MG100809	MG366461
43			36320B	---	---	MG100810	NO
43			36322	MG028257	MG028520	MG100811	MG366462
43			36322B	---	---	NO	MG366463
43			36323	MG028258	MG028522	MG100812	MG366464
43			36323B	---	---	NO	MG366465
44	Río Tinajeros	Grande de Santiago	36324	MG028259	MG028523	MG100813	MG366466
44			36324B	---	---	MG100814	MG366467
44			36326	MG028260	MG028524	NO	NO
44			36327	MG028261	MG028525	NO	NO
44			36328	MG028262	MG028526	NO	NO
44			36329	MG028263	MG028527	NO	NO
44			36330	MG028264	MG028528	NO	NO
44			36331	MG028265	MG028529	NO	NO
44			36332	NO	MG028530	NO	NO
	Poeciliopsis prolifica		42668	MG028009	MG028279	MG100617	MG366191
	Poeciliopsis prolifica		42668B	---	---	NO	MG366192
	Poeciliopsis prolifica		42669	MG028010	MG028280	MG100618	MG366193
	Poeciliopsis prolifica		42670	MG028011	MG028281	MG100619	NO
	Poeciliopsis prolifica		42670B	---	---	MG100620	NO
	Poeciliopsis prolifica		42671	MG028012	MG028282	MG100621	NO
	Poeciliopsis prolifica		42812	MG028013	MG028283	NO	NO
	Poeciliopsis prolifica		42815	MG028014	MG028284	NO	NO
	Poeciliopsis prolifica		42816	MG028015	MG028285	NO	NO
	Poeciliopsis prolifica		42817	MG028016	MG028286	NO	NO
	Poeciliopsis prolifica		42818	MG028017	MG028287	NO	NO
	Poeciliopsis prolifica		42820	MG028018	MG028289	NO	NO
	Poeciliopsis prolifica		42821	MG028019	MG028290	NO	NO
	Poeciliopsis prolifica		42822	MG028020	MG028291	NO	MG366194
	Poeciliopsis prolifica		42822B	---	---	NO	MG366195
	Poeciliopsis prolifica		42824	MG028021	MG028292	NO	NO
	Poeciliopsis prolifica		42825	MG028022	MG028293	NO	NO

Additional file 2: Primers, PCR conditions, and References.

	Cytb	*coxI*	*S7*	*RHO*
Primers	Glu-F Thr-R	Fish-F1 Fish-R1	S7RPEX1F S7RPEX3R	RH193F RH1073R
Size (pb)	533	626	859	845
Reference	Doadrio and Domínguez, (2004).	Ward *et al*. (2005).	Chow & Hazama, (1998)	Chen, Bonillo & Lecointre (2003)
Denaturing (step 1)	94ºC, 2 min.	94ºC, 2 min.	94ºC, 3 min.	94ºC, 3 min.
Cycles (step 2)	35	35	32	32
Denaturing	94ºC, 45 s.	94ºC, 30 s.	94ºC, 45 s.	94ºC, 45 s.
Annealing Extension	47ºC, 1 min. 72ºC, 90 s.	50ºC, 30 s. 72ºC, 1 min.	58.5ºC, 30 s. 72ºC, 90 seg.	55ºC, 45 s. 72ºC, 90 s.
Final Extension (step 3)	72ºC, 5 min.	72ºC, 10 min.	72ºC, 7 min.	72ºC, 7 min.

Additional file 3: Models selected with Akaike information criterion and the parameters of each gene.

	Cytb	*coxI*	*S7*	*RHO*
Model	GTR+I+G	GTR+I	TVM+I+G	TrN+I+G
Frec. A	0.2457	0.2488	0.2802	0.1890
Frec. C	0.3046	0.2696	0.1863	0.2900
Frec. G	0.1469	0.1718	0.2347	0.2276
Frec. T	0.3029	0.3098	0.2988	0.2935
P-inv	0.0000	0.8300	0.1600	0.0000
Gamma shape	0.1220	N/A	0.0250	0.9983

Additional file 4: Geographical coordinates of the 162 sites registers in Colección de Peces de la Universidad Michoacana de San Nicolás de Hidalgo used as presence data for species distribution modelling.

Locality	Latitude	Longitude
4 mi S San Miguel de Allende	20.850226	-100.793422
below laguna de Naranja	19.7813659	-101.762748
1 km downstream Platanal	19.9346	-102.252
1 mi ESE of Teuchitlan	20.67878	-103.838685
1 mi WNW of Cuyutlan	20.424082	-103.375023
2 km S of Magdalena	20.89278	-104.031065
2.5 mi S of San Marcos	20.3	-103.533333
200 m above main bridge into Compostela	21.231934	-104.900894
25 m offshore of Manantial Ojo de la Liebre	19.8222613	-101.788056
25 mi E of Ameca	20.687398	-103.693749
3.7 mi E Atequiza	20.403876	-103.070958
5 km W of Teuchitlan	20.69267	-103.91669
5 mi N of Yahualica	21.2299	-102.851
5 mi NE Piedad	20.407505	-101.956027
5 mi W of Ameca	20.481476	-103.975073
6 de Enero	21.52464	-104.80384
6 mi N Colimilla	20.691996	-103.228296
7 mi E of Penjamo	20.4532116	-101.607094
Above & below the presa at Copalita	20.838393	-103.419
Achacales	19.70577	-104.14599
Andocutin	19.9384243	-100.854608
Angamacutiro	20.1446232	-101.704673
Apaseo El Grande	20.5396064	-100.696937
Araro,manantiales al E Cuitzeo	19.9071176	-100.831779
Arroyo en carr. A la Estancia Km 1	21.4100833	-102.7375
Arroyo en Quitupan	19.9243333	-102.871972
Arroyo en San Carlos	20.7852222	-102.766306
Arroyo San Andres	20.780819	-104.181026
Arroyo San Gabriel	21.128684	-100.861203
Atotonilco	21.0021667	-100.799472
Balneario at Tocumbo	19.703312	-102.514432
Balneario Atotonilco	20.5613871	-102.510781
Balneario Chorros de Tala	20.69547	-103.67935
Balneario El Rincon	20.689558	-103.841509
below dam for Presa Ignacio Allende	20.840874	-100.828553
between Etzatlan and San Marcos	20.7776	-104.164
between Ocotlan and Chapala	20.387814	-102.824525
Bordo cerca de Chimaliquin	21.3566667	-102.801889
bridge at Estanzuela	20.5243611	-104.336667

bridge between Zamora and Jacona	19.972381	-102.295393
Canal at Tarecuato	19.8436	-102.479261
Canal de Querendaro 2	19.8682102	-100.972431
Canal en Rancho Nuevo	20.7008056	-102.939583
Capacho	19.9613278	-101.224772
Carretera 304 San Julian-San Miguel el Alto	21.0088056	-102.296639
Chiquimitio	19.799	-101.246083
Cienega en La Purisima	19.5221667	-103.342472
Cieneguilla	20.9522222	-100.795194
Copandaro	19.899473	-101.216961
Copandaro de Jimenez	19.8898489	-101.668758
Desembocadura del Rio Querendaro	19.9009528	-100.976
ditch 2 mi N Guadalajara	20.75	-103.366667
ditch 2.5 N Etzatlan	20.797845	-104.094427
ditch 8 mi by hwy 70 east of Ameca	20.546389	-103.907617
ditch between Ameca & Hwy 15	20.591616	-103.828087
Dren la Cinta	20.0860194	-101.154617
E edge of Belen del Refugio	21.530129	-102.432634
El Borbollon, La Maiza	19.5026833	-101.384592
El Chacalito	19.8129333	-104.240933
El Nacimiento	20.5403286	-100.613519
El Palo Verde reservoir	20.769055	-104.11573
El Parian	19.6892194	-101.26838
Estacion Querendaro	19.8823167	-100.931714
exhacienda de Guadalupe	19.6589726	-101.273074
Ferrocarrileros	20.8126389	-100.818611
Granja Sanhuarripa	20.7770249	-104.163603
Hacienda San Sebastian	20.822912	-104.119572
Hwy 54 bridge in Apozol	21.476142	-103.088792
Iramuco	19.9577778	-100.923056
Irrigation canal 3.7 mi NE Alvaro Obregon	19.806554	-101.063069
La Canal, on E of Tocumbo	19.7033	-102.519
La Mintzita	19.644973	-101.274336
La Vega	20.5833333	-103.85
Lago de Atotonilco	20.3603	-103.654
Lago de Camecuaro	19.902411	-102.209455
Lago de Chapala	20.279166	-103.1875
Lago de Cuitzeo	19.9628975	-101.059816
Lago de Magdalena	20.902806	-104.017245
Lago de Zapotlan	19.748632	-103.469951
Laguna Colorada	20.756772	-103.989213
Laguna Corralejo	20.5127318	-101.607914
Laguna de Cajititlan	20.415926	-103.329852
Laguna de Sayula	20.075605	-103.509025
Laguna de Yuriria	20.241507	-101.203946
Laguna de Zacapu	19.823495	-101.787291

Laguna San Marcos	20.29057	-103.551612
Laguna Zacoalco	20.243351	-103.589273
Las Adjuntas	20.670525	-101.859447
Los Cipreses	19.827634	-101.787137
Los Lavaderos	19.8835722	-100.447078
Los Negritos	20.061113	-102.609368
Manantial Bellas Fuentes	19.8215836	-101.68013
Manantial en San Jose de Gracia	20.6742222	-102.555389
Manantial in Quinceo	19.7339767	-101.222332
Manantial La Luz	19.9370475	-102.299697
Manantial Orandino	19.956206	-102.325994
Manantial San Cristobal	19.9610861	-101.315353
Manantial San Francisco del Rincon	21.0512833	-101.844344
Molino viejo cerca de la Cofradia	20.3898611	-103.755028
N Tanhuato	20.291134	-102.329762
near Balneario El Cortijo	20.992746	-100.797794
near Cuyacapan	19.95954	-103.51704
near Estancia de Ayones	20.901222	-104.079639
Ojo de Agua de Santiaguito	21.046216	-101.835825
Palo Blanco at bridge crossing	19.7417	-104.178
Parque La Angostura	19.8277229	-101.787093
Peninsula de San Agustin del Pulque	19.9533389	-101.109139
Petatan, Lago de Chapala	20.1626944	-102.867944
Pond at end of Lago Union de Tula	19.940254	-104.257415
Presa Aristeo Mercado	19.929525	-101.669319
Presa Buenavista	20.3349722	-103.755611
Presa de Cointzio	19.616234	-101.280209
Presa de Garabato	20.6245556	-102.687667
Presa de Huapango	19.9246389	-99.8017778
Presa El Alamo	20.1986413	-99.7867317
Presa el Pajonal	19.539375	-101.417772
Presa Ignacio Allende	20.8507052	-100.823794
Presa Juriquilla	20.7039699	-100.462281
Presa Melchor Ocampo	20.0945908	-101.738532
Presa Palote	21.18	-101.68944
Presa San Antonio de Huaracha	19.962975	-102.578083
Presa San Ignacio	20.6357222	-103.931733
Presa San Juanico	19.847777	-102.685833
Presa Teclan	20.3205	-103.734861
Presa Valle de Juarez	19.9385556	-102.949719
Pueblo Rio Laja	21.20625	-100.922333
Puente Chapulimita	20.68025	-103.908139
Ribera del Lago de Cuitzeo, 2.5 km al norte del Salitre	19.9167444	-101.299483
Rinconcillo	20.7889444	-100.806944
Rio Ameca at Ameca	20.543587	-104.043992
Rio Ameca en San Blasito	20.7038139	-104.309581

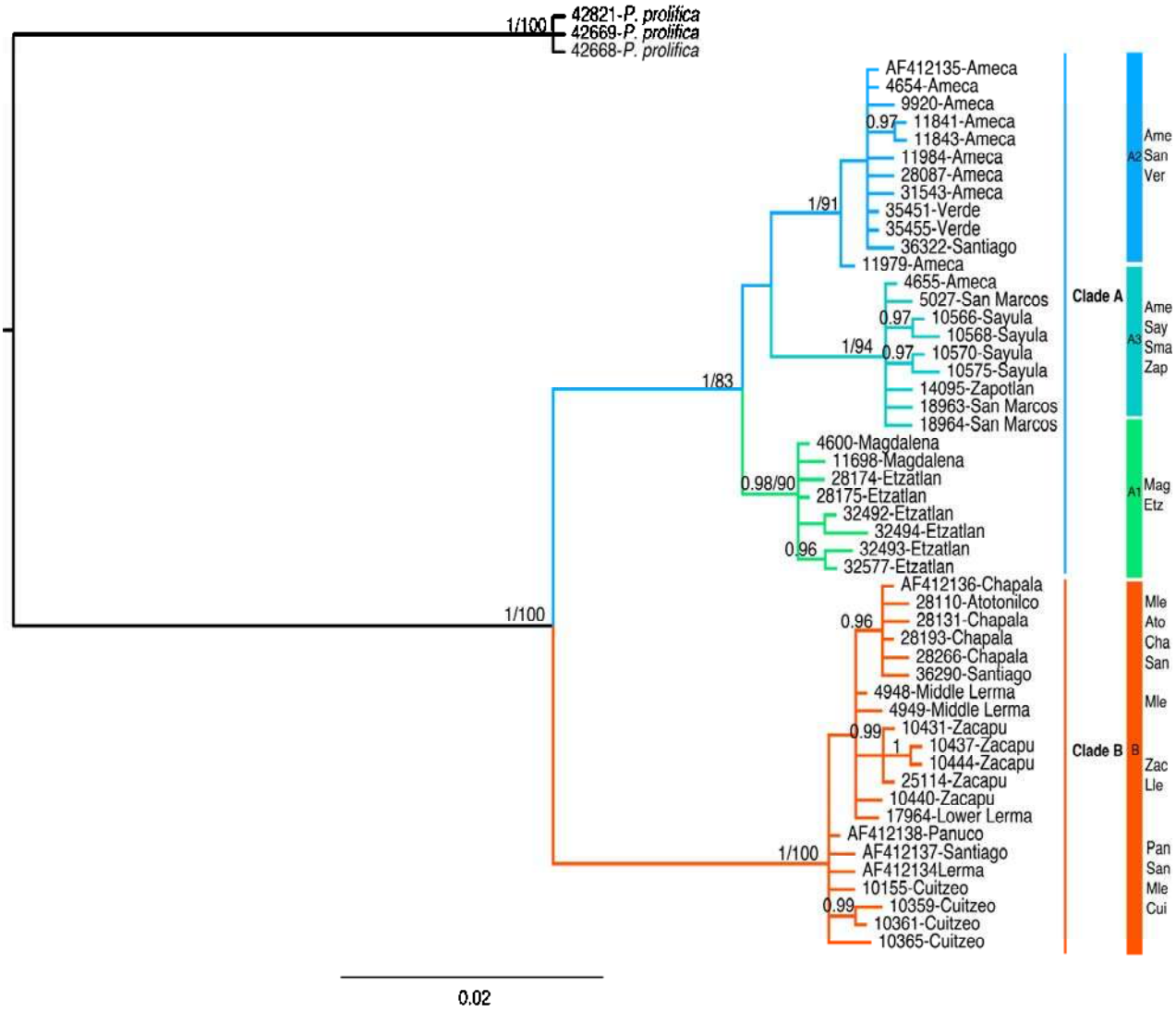
Rio Calvillo	21.847038	-102.739145
Rio Celio	19.947655	-102.304475
Rio chiquito de Amatlan de Cañas	20.8026417	-104.417678
Rio de la Laja below bridge	20.670345	-100.751515
Rio Duero, 9.2 mi W Chilchota	19.912489	-102.205813
Rio Grande, north of Zacapu	19.824226	-101.77389
Rio juchipila en el poblado de jalpa	21.6518	-102.966
Rio Juchipila N Santa Rosa	21.602895	-102.94872
Rio la Patera	19.9211278	-101.724569
Rio Salado	20.68675	-103.693361
Rio Santiago at Poncitlan	20.38502	-102.923955
Rio Teuchitlan, E edge of Teuchitlan	20.682196	-103.843366
Rio Tizapan	20.0431667	-103.803722
Rio Turbio	20.75	-101.833333
Rio Turbio, 8 mi E Penjamo	20.4532	-101.607
Rio Tuxpan, 1 km N of Atenquique	19.530125	-103.430328
Rio Xoconoxtle San Juan	20.9420833	-100.977222
Rio Zula bajo el puente	20.4131911	-102.724989
Road by La Quemada	21.3271944	-101.095806
S Empalme Escobedo	20.597196	-100.746606
San Nicolas, Rio Verde	21.29595	-102.549917
SE corner, in Yahualica	21.167672	-102.881113
Spring-fed pond N Jaripo	19.942262	-102.591945
Stream E El Refugio de los Orendain	20.725792	-103.64856
stream flowing into a presa near La Quemada	20.9631056	-104.052878
Stream near Santa Anna	20.539257	-103.434305
Stream near Santa Cruz de las Flores	20.469271	-103.506936
Tarejero Spring	19.8215407	-101.717526
Texcalme	20.4580944	-104.070367
Tierra colorada	19.7998806	-101.247481
Trib to Ameca W of Ameca	20.535054	-104.057926

Additional file 5: Ambiguously Aligned Regions for *S7.*

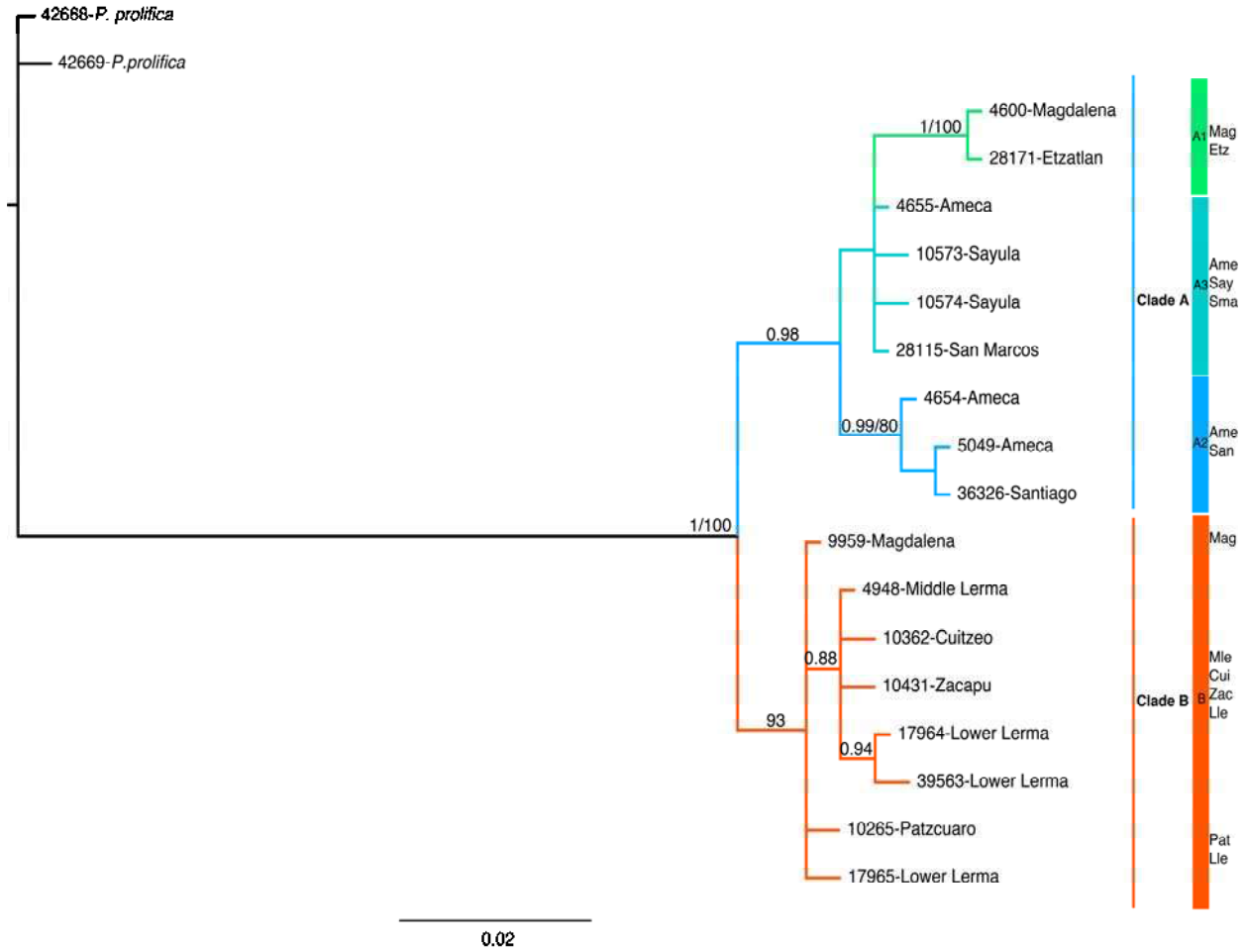
	BP																			
Clade	244	392	402	403	404	410	411	417	418	419	420	431	456	457	458	471	472	504	505	511
A1	C	-/T	A	A	-/A	-/T	A	T	-/T	A	A	T	A	T	G	G	C	A	G	C
A2	C	-/T	A	A	-/A	-/T	-/A	T	-/T/G	A	A	-//T	-/A	-/T	-/G	-/G	-/C	-/A	-/G	-/C
A3	C	-/T	-/A	-/A	-/A	-/T	A	-/T	-/T/G	-/A	-/A	T	A	T	G	G	C	A	G	C
B	-	-/T	-	-	-	-	A	-/T	-/G	-/A	-/A	T	A	T	G	G	C	A	G	C

	BP																			
Clade	512	605	606	607	608	609	610	613	614	615	616	617	618	619	620	621	622	623	624	625
A1	A	-/A	-/T	-/T	-/A	-/G	-/C	T	A	G	C	T	T	A	A	G	C	T	A	G
A2	-/A/T	A	T	T	A	G	C	-/T	-/A	-/G	-/C	-/T	-/T	A	A	G	C	T	A	G
A3	T/A	A	T	T	A	G	C	-/T	A	-/G	-/C	-/T	-/T	-/A	-/A	-/G	-/C	-/T	-/A	-/G
B	A	A	T	T	A	G	C	T	A	G	C	T	T	A	A	G	C	T	A	G

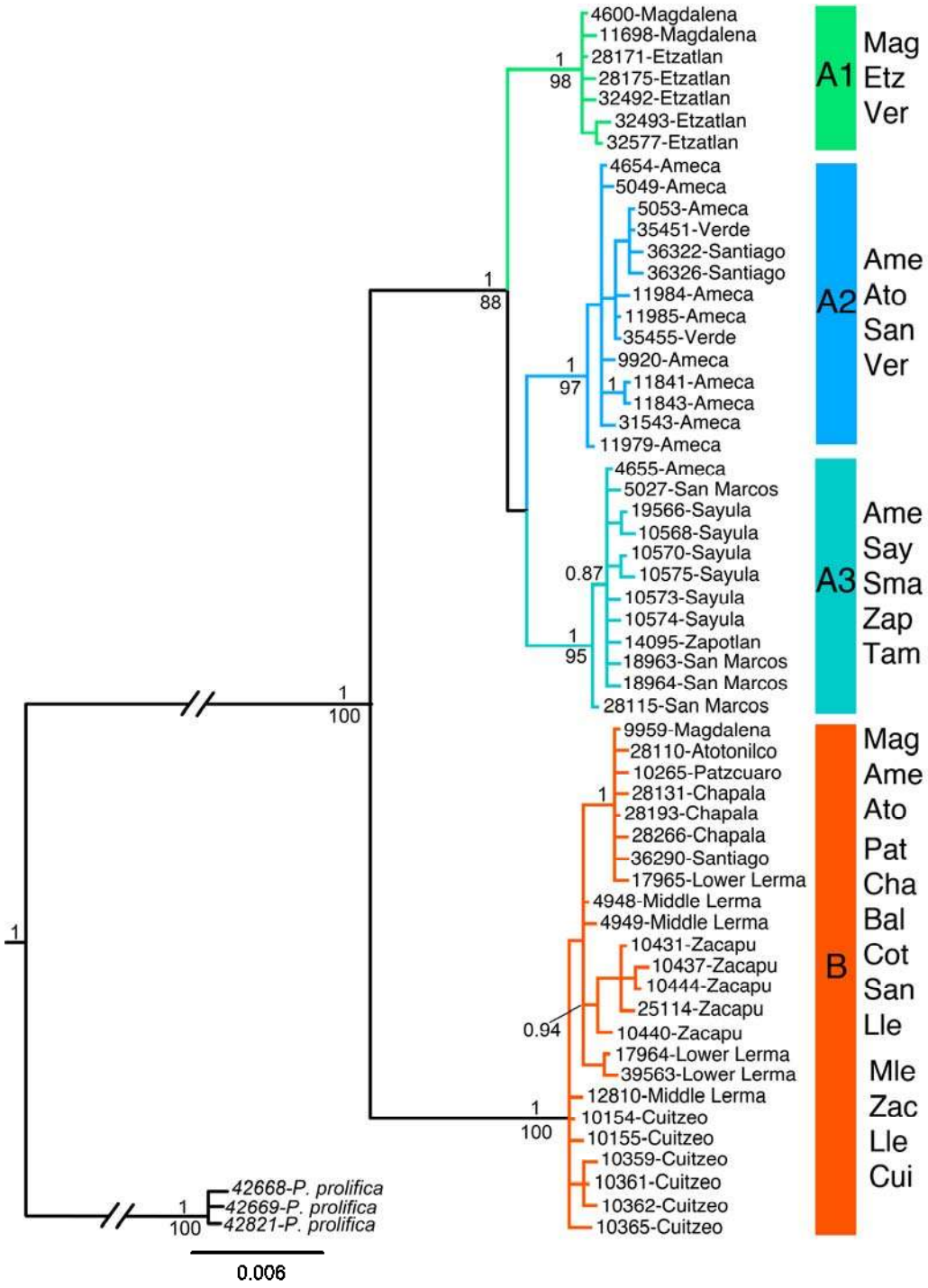
	BP												
Clade	626	627	628	629	644	708	709	710	711	781	782	785	786
A1	A	C	G	A	A	A	C	A	T	C	A	G	C
A2	-/A	-/C	-/G	-/A	-/A	A	C	A	T	C	A	G	C
A3	-/A	-/C	-/G	A	A	A	C	A	T	C	A	G	C
B	A	C	G	A	A	-/A	-/C	-/A	-/T	-/C	-/A	-/G	-/C



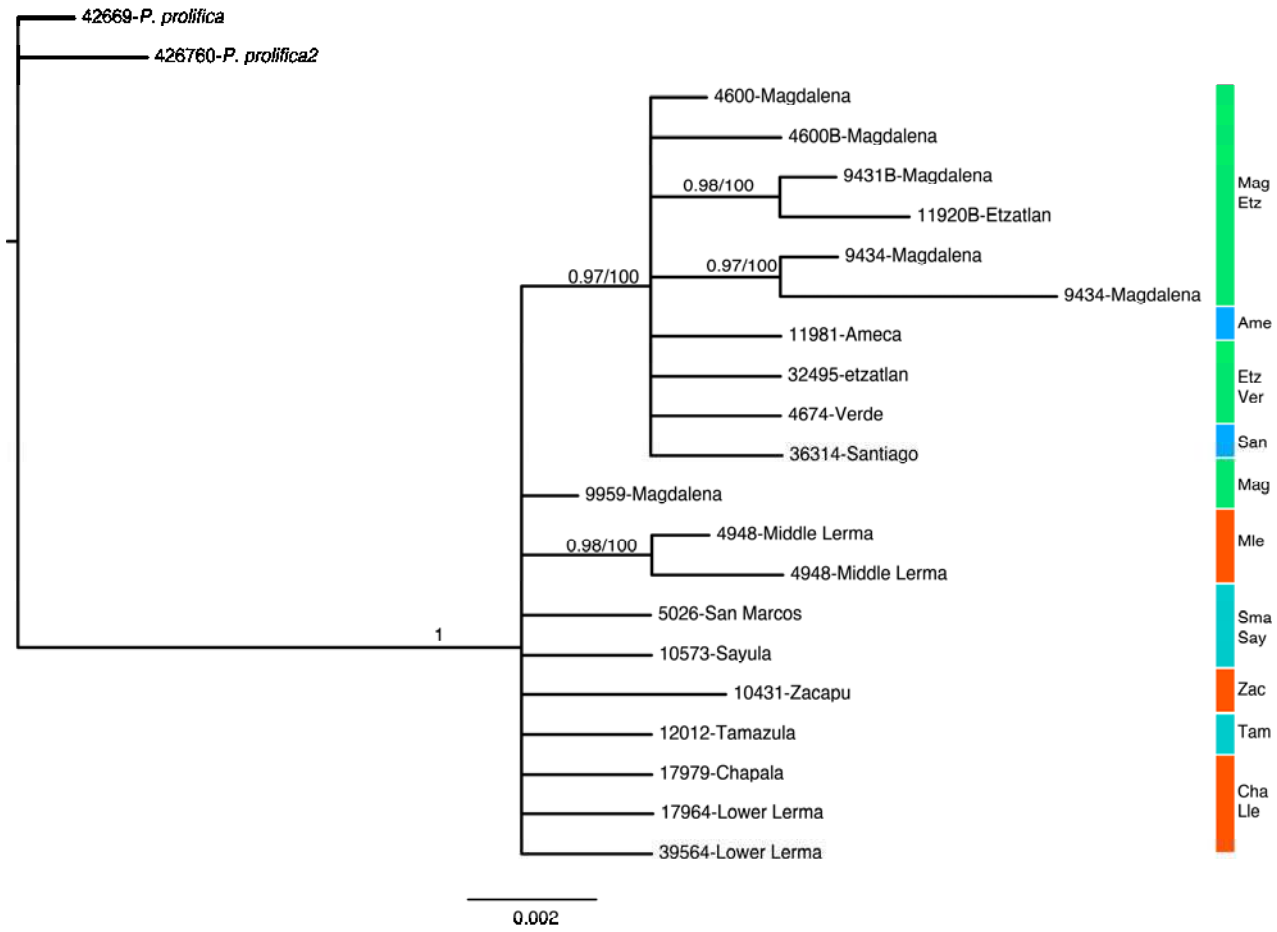
Additional file 6: The Bayesian inference tree of *P. infans* from *cytb* mitochondrial gene (1,083 bp). Bayesian posterior probability (>0.9; above the branches) and maximum likelihood bootstrap values (>80%; below the branches) are indicated.



Additional file 7: The Bayesian inference tree of *P. infans* from *coxI* mitochondrial gene (631 bp). Bayesian posterior probability (>0.9; above the branches) and maximum likelihood bootstrap values (>80%; below the branches) are indicated.

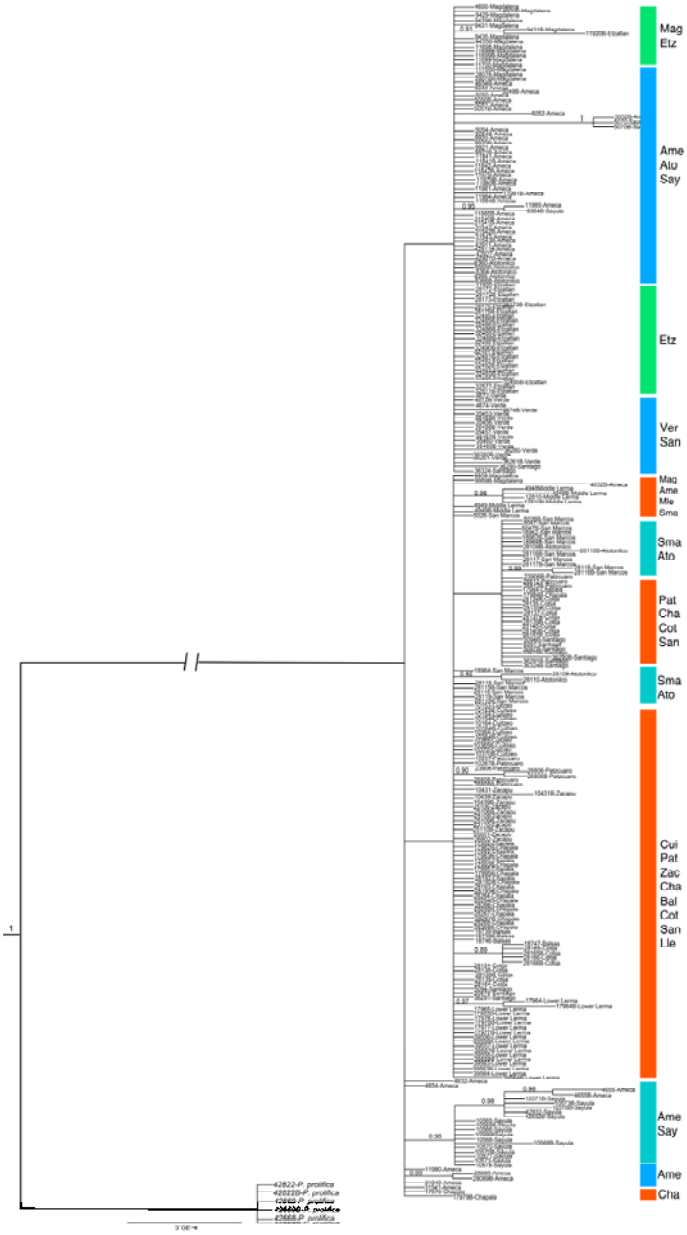


Additional file 8: The Bayesian inference tree of *P. infans* from concatenated sequences of two mitochondrial genes (*cytb*, *coxI;* 1,771 bp). Bayesian posterior probability (>0.9; above the branches) and maximum likelihood bootstrap values (>80%; below the branches) are indicated.

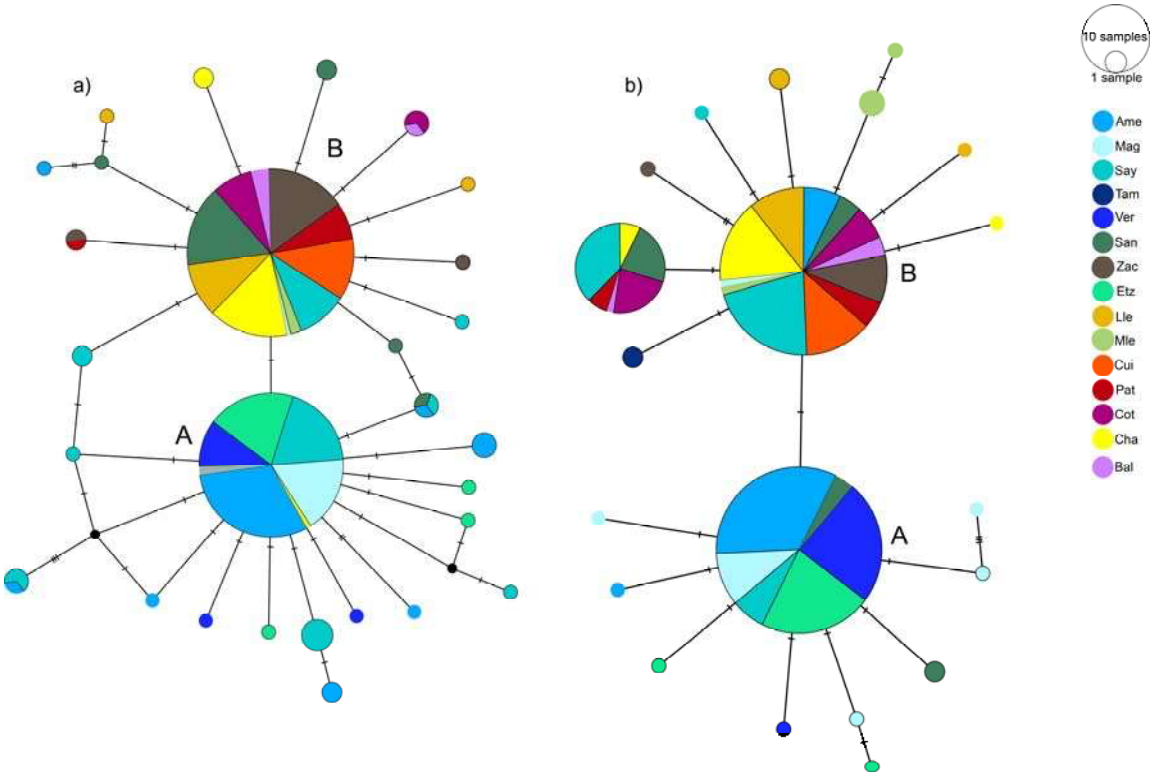


Additional file 9: The Bayesian inference tree of *P. infans* from *RHO* nuclear gene (845 bp). Bayesian posterior probability (>0.9; above the branches) and maximum likelihood bootstrap values (>80%; below the branches) are indicated.

Additional file 10: The Bayesian inference tree of *P. infans* from *S7* nuclear gene (859 bp). Bayesian posterior probability (>0.9; above the branches) and maximum likelihood bootstrap values (>80%; below the branches) are indicated.



Additional file 11. The Bayesian inference tree of P. infans from concatenated sequences of two nuclear genes (S7 and RHO: 1,704 bp). Bayesian posterior probability (>0.9) and maximum likelihood bootstrap values (>80%) are indicated.



Additional file 12. Haplotype networks for nuclear genes, a) *S7* gene, b) *RHO* gene. The two recovered haplogroups are show with labels A and B.

Abbreviations

mtDNA: Mitochondrial Deoxyribonucleic Acid;
nDNA: Nuclear Deoxyribonucleic Acid;
SEMARNAT: Secretaria de Medio Ambiente y Recursos Naturales

Declarations

Ethics approval and consent to participate

The research has been carried out within an appropriate ethical framework. All procedures performed, including field sampling techniques and laboratory protocols, as anesthetised and euthanasia techniques used in this study were reviewed and approved by a committee of Mexican Ministry of Environmental and Natural Resources (SEMARNAT), under collection permit number PPF/DGOPA-262/17. This procedures were also assessed and approved by the

Institutional Biosecurity and Bioethics committee of the Institute of Chemical and Biological Research, Universidad Michoacana de San Nicolás de Hidalgo, México. The specimens were anesthetised using tricaine mesylate (MS-222) to anesthetize the fishes, according with the Official Mexican Norm NOM-051-ZOO-1995 and NOM-033-SAG/ZOO-2014 for humanitarian treatment in the mobilization of animals. The study organism is neither protected nor endangered.

Consent for publication

Not applicable

Availability of data and material

All data generated or analyzed during this study are included in this published article [and its additional files]. The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Competing interest

The authors declare that they have no competing interests.

Funding

Funding for this study was provided by the Ministerio de Economía y Competitividad y FEDER, Spain (CGL2016-75262-P) to ID and PRODEM, CIC-UMSNH, Chester Zoo garden and CONACYT sabbatical grant to ODD, as well as funds from the U.S. National Science Foundation (DEB 1354930) to KRP. These funding body do not had a specific role in the design of the study and collection, analyses, and interpretation of data in writing the manuscript.

Authors' contributions

RGBL, OD, conceived the ideas; RGBL, OD, RPR, conducted the fieldwork and collected the specimens; RGBL, OD, RPR, and ID, analyzed the data and RGBL, OD, RPR, KP, and ID write the manuscript. All authors have read and approved the manuscript.

Acknowledgements

This work was conducted as part of the doctoral studies of R. G. Beltrán-López, and she was supported with a scholarship from the Consejo Nacional de Ciencia y Tecnología (CONACYT). We thank Isaí Betancourt for his help with sampling, and to Berenice García Andrade and Silvia Perea for their help in laboratory and for their comments to improve this work. We thank three anonymous reviewers for their comments that improved this manuscript.

REFERENCES

1. Castoe TA, Daza JM, Smith EN, Sasa MM, Kuch U, Campbell JA, et al. Comparative phylogeography of pitvipers suggests a consensus of ancient Middle American highland biogeography. J. Biogeogr. 2009; 36(1): 88–103.
2. Ruíz-Sánchez E, Specht ChD. Influence of the geological history of the Trans-Mexican Volcanic Belt on the diversification of *Nolina parviflora* (Asparagaceae: Nolinoideae). J. Biogeogr. 2013; 40(7): 1336-1347.
3. Hewitt GM. Some genetic consequences of ice ages, and their role in divergence and speciation. Biol. J. Linn. Soc. 1996; 58(3): 247–276.
4. Hewitt GM. The genetic legacy of the Quaternary ice ages. Nature. 2000; 405(6789): 907–913.
5. Hewitt GM. Genetic consequences of climatic oscillations in the Quaternary. Philos. Trans. R. Soc. B. 2004; 359(1442): 183-195.
6. Bermingham E, Martin AP. Comparative mtDNA phylogeography of neotropical freshwater fishes: testing shared history to infer the evolutionary landscape of lower Central America. Mol. Ecol. 1998; 7(4): 499-517.
7. Bryson RW, García-Vázquez UO, Riddle BR. Phylogeography of Middle American Gophersnakes: Mixed responses to biogeographical barriers across the Mexican Transition Zone. J. Biogeogr. 2011; 38: 1570-1584.
8. Bryson RW, García-Vázquez UO, Riddle BR. Diversification in the Mexican Horned Lizard *Phrynosoma orbiculare* across a dynamic landscape. Mol. Phylogenet. Evol. 2012a; 62(1): 87-96.

9. Bryson RW, Chaves J, Smith BT, Miller MJ, Winker K, Pérez-Emán JL, et al. Diversification across the New World within the 'blue' Cardinalids (Aves: Cardinalidae). J. Biogeogr. 2014; 41(3): 587-599.
10. Ghanavi HR, González EG, Doadrio I. Phylogenetic relationships of freshwater fishes of the genus *Capoeta* (Actinopterygii, Cyprinidae) in Iran. Ecol. Evol. 2016; 6(22): 8205-8222.
11. Perea S, Cobo-Simo M, Doadrio I. Cenozoic tectonic and climatic events in southern Iberian Peninsula: Implications for the evolutionary history of freshwater fish of the genus *Squalius* (Actinopterygii, Cyprinidae). Mol. Phylogenet. Evol. 2016; 97: 155-169.
12. Leprieur F, Tedesco PA, Hugueny B, Beauchard O, Dürr HH, Brosse S, et al. Partitioning global patterns of freshwater fish beta diversity reveals contrasting signatures of past climate changes. Ecol. Lett. 2011; 14(4): 325-334.
13. Ferrari L, Orozco-Esquivel T, Manea V, Manea M. The dynamic history of the Trans-Mexican Volcanic Belt and the Mexico subduction zone. Tectonophysics. 2012; 522-523: 122-149.
14. Halffter G. Biogeography of the montane entomofauna of Mexico and Central America. Annu. Rev. Entomol. 1987; 32: 95-114.
15. Morrone JJ. Halffter´s Mexican Transition zone (1962-2014), cenocrons and evolutionary biogeography. J. Zoolog. Syst. Evol. Res. 2015; 53(3): 249-257.
16. Marshall CJ, Liebherr JK. Cladistic biogeography of the Mexican transition zone. J. Biogeogr. 2000; 27(1): 203-216.
17. Halas D, Zamparo D, Brooks DR. A protocol for studying biotic diversification by taxon pulses. J. Biogeogr. 2005; 32(2): 249-260.
18. Domínguez-Domínguez O, Pérez-Ponce de León G. ¿La mesa central de México es una provincia biogeográfica? Análisis descriptivo basado en componentes bióticos dulceacuícolas. Rev. Mex. Biodivers. 2009; 80(3): 835-852.
19. West RC. Surface configuration and asociated geology of Middle America. In: Wanchope R, West RC, editors. Handbook of Middle American Indians, University of Texas Press, Austin. 1964. p. 33-82.

20. Zaldívar-Riverón A, Nieto-Montes de Oca A, Laclette JP. Phylogeny and evolution of dorsal pattern in the Mexican endemic lizard genus *Barisia* (Anguidae: Gerrhonotinae). J. Zoolog. Syst. Evol. Res. 2005; 43(3): 243–257.
21. Bryson RW, García-Vázquez UO, Riddle BR. Relative roles of Neogene vicariance and Quaternary climate change on the historical diversification of Bunchgrass Lizards (*Sceloporus scalaris* group) in Mexico. Mol. Phylogenet. Evol. 2012b; 62(1): 447-457.
22. Arellano E, González-Cozátl FX, Rogers DS. Molecular systematics of Middle American harvest mice *Reithrodontomys* (Muridae), estimated from mitochondrial cytochrome b gene sequences. Mol. Phylogenet. Evol. 2005; 37(2): 529–540.
23. Puebla-Olivares F, Bonaccorso E, Espinosa de los Monteros A, Omland KE, Llorente-Bousquets JE, Peterson AT, et al. Speciation in the Emerald Toucanet (*Aulacorhynchus prasinus*) complex. Auk. 2008; 125(1): 39–50.
24. Castañeda-Rico S, León-Paniagua L, Vázquez-Domínguez E, Navarro-Sigüenza G. Evolutionary diversification and speciation in rodents of the Mexican lowlands: The *Peromyscus melanophrys* species group. Mol. Phylogenet. Evol. 2014; 70(1): 454-463.
25. Miller RR, Smith ML. Origin and geography of the fishes of Central Mexico. In: Hocutt CH, Wiley EO, editors. The Zoogeography of North American freshwater fishes. Wiley-Intersciences publication. 1986. p. 487-519.
26. Mateos M, Sanjur OI, Vrijenhoek RC. Historical biogeography of the livebearing fish genus *Poeciliopsis* (Poeciliidae: Cyprinodontiformes). Evolution. 2002; 56(5): 972-984.
27. Huidobro L, Morrone JJ, Villalobos JL, Álvarez F. Distributional patterns of freshwater taxa (fishes, crustaceans and plants) from the Mexican transition zone. J. Biogeogr. 2006; 33(4): 731-741.
28. Domínguez-Domínguez O, Doadrio I, Pérez-Ponce de León G. Historical biogeography of some river basins in Central Mexico evidenced by their goodeine freshwater fishes: A preliminary hypothesis using secondary Brooks Parsimony Analysis (PBA). J. Biogeogr. 2006; 33(8): 1437–1447.
29. Domínguez-Domínguez O, Pedraza-Lara C, Gurrola-Sánchez N, Perea S, Pérez-Rodríguez R, Israde-Alcántara I, et al. Historical biogeography of

the Goodeinae (Cyprinodontiforms). In: Uribe MC, Grier HJ, editors. *Viviparous Fishes II.* New Life Publications. 2010. p. 34-69.

30. Pérez-Rodríguez R, Domínguez-Domínguez O, Pérez-Ponce de León G, Doadrio I. Phylogenetic relationships and biogeography of the genus *Algansea* Girard (Cypriniformes: Cyprinidae) of central Mexico inferred from molecular data. BMC Evol. Biol. 2009; 9: 223.
31. Parra-Olea G, Windfield JC, Velo-Antón G, Zamudio KR. Isolation in habitat refugia promotes rapid diversification in a montane tropical Salamander. J. Biogeogr. 2012; 39(2): 353-370.
32. Pedraza-Lara C, Doadrio I, Breinholt JW, Crandall KA. Phylogeny and Evolutionary Patterns in the Dwarf Crayfish Subfamily (Decapoda: Cambarellinae). Plos One. 2012; 7(11): e48233.
33. Domínguez-Domínguez O, Alda F, Pérez-Ponce de León G, García-Garitagoitia JL, Doadrio I. Evolutionary history of the endangered fish *Zoogoneticus quitzeoensis* (Bean, 1898) (Cyprinodontiformes: Goodeidae) using a sequential approach to phylogeography base don mitochondrial and nuclear DNA data. BMC Evol. Biol. 2008a; 8: 161.
34. Domínguez-Domínguez O, Pérez-Rodríguez R, Doadrio I. Morphological and genetic comparative analyses of populations of *Zoogoneticus quitzeoensis* (Cyprinodontiformes: Goodeidae) from Central Mexico, with description of a new species. Rev. Mex. Biodivers. 2008b; 79(2): 373-383.
35. Pérez-Rodríguez R, Domínguez-Domínguez O, Mar-Silva F, Doadrio I, Pérez-Ponce de León G. The historical biogeography of the southern group of the sucker genus *Moxostoma* (Teleostei: Catostomidae) and the colonization of Central Mexico. Zool. J. Linn. Soc. 2016; 177(3): 633-647.
36. Pérez-Rodríguez R, Domínguez-Domínguez O. Doadrio I, Cuevas-García E, Pérez-Ponce de León G. Comparative historical biogeography of three groups of Neartic freshwater fishes across central Mexico. J. Fish Biol. 2015; 86(3): 993–1015.
37. Lucinda PHF. Family Poeciliidae. In: Reis RE, Kullander SO, Ferraris CJJr, editors. Check list of the freshwater fishes of South and Central America. Porto Alegre: EDIPUCRS. 2003. p. 555-581.
38. Miller R, Minckley W, Mark, S. Freshwater Fishes of México. University of Chicago. Estados Unidos de América. 2005.

39. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Mol. Biol. Evol. 2013; 30(12): 2725–2729.
40. Librado P, Rozas J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 2009; 25(11): 1451–1452.
41. Stephens M, Donelly P. A comparison of Bayesian methods for haplotype reconstruction from population genotype data. Am. J. Hum. Genet. 2003; 73(5): 1162–1169.
42. Sousa-Santos C, Robalo JI, Collares-Pereira MJ, Almada VC. Heterozygous indels as useful tools in the reconstruction of DNA sequences and in the assessment of ploidy level and genomic constitution of hybrid organisms. DNA Seq. 2005; 16(6): 462–467.
43. Huson DH, Bryant D. Application of Phylogenetic Networks in Evolutionary Studies. Mol. Biol. Evol. 2006; 23(2): 254–267.
44. González-Peña D, Gómez-Blanco D, Reboiro-Jato M, Fernández-Riverola F, Posada D. ALTER: program-oriented conversion of DNA and protein alignments. Nucleic Acids Res. 2010; 38(2): 14-18.
45. Kass RE, Raftery, AE. Bayes factors. J. Amer. Statist. Assoc. 1995; 90(430): 773–795.
46. Lanfear R, Calcott B, Ho SY, Guindon S. PartitionFinder: Combined selection of partitioning schemes and substitution models for phylogenetic analyzes. Mol. Biol. Evol. 2012; 29(6): 1695–1701.
47. Silvestro D, Michalak I. RAxMLGUI: a graphical front-end for RAxML. Org. Divers. Evol. 2012; 12(4): 335–337.
48. Stamatakis A. RAxML Version 8: A tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies. Bioinformatics. 2014; 30(9): 1312-1313.
49. Alfaro ME, Zooler S. Lutzoni F. Bayes or bootstrap? A simulation study comparing the performance of Bayesian Markov chain Monte Carlo sampling and bootstrapping in assessing phylogenetic confidence. Mol. Biol. Evol. 2003; 20(2): 255–266.
50. Ronquist F, Teslenko M, Van der Mark P, Ayres DL, Darling A, Höhna S, et al. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 2012; 61(3): 539–542.

51. Rambaut A, Drummond A. Tracer v1.5. 2007. Available: http://tree.bio.ed.ac.uk/software/tracer.

52. Drummond AJ, Suchard MA, Xie D, Rambaut A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. 2012; 29(8): 1969–1973.

53. Zardoya R, Doadrio I. Molecular evidence on the evolutionary and biogeographical patterns of European Cyprinids. J. Mol. Evol. 1999; 49(2): 227–237.

54. Machordom A, Doadrio I. Evidence of a cenozoic Betic-Kabilian connection based on freshwater fish phylogeography (*Luciobarbus*, Cyprinidae). Mol. Phylogenet. Evol. 2001; 18(2): 252–263.

55. Near TJ, Benard MF. Rapid allopatric speciation in logperch darters (Percidae: *Percina*). Evolution. 2004; 58(12): 2798–2808.

56. Drummond AJ, Ho SYW, Phillips J, Rambaut A. Relaxed phylogenetics and dating with confidence. Plos Biol. 2006; 5, e88.

57. Heled J, Drummond AJ. Bayesian inference of population size history from multiple loci. BMC Evol. Bio. 2008; 8: 289.

58. Excoffier L, Lischer HE. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol. Ecol. Res. 2010; 10(3): 564-567.

59. Nylander JAA, Olsson U, Alström P, Sanmartín I, Baker A. Accounting for phylogenetic uncertainty in biogeography: a Bayesian approach to Dispersal–Vicariance Analysis of the thrushes (Aves: *Turdus*). Syst. Biol. 2008; 57(2): 257–268.

60. Ree RH, Moore BR, Webb CO, Donoghue, MJ. A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. Evolution. 2005; 59(11): 2299-2311.

61. Ree RH, Smith SA. Maximum-Likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. Syst. Biol. 2008; 57(1): 4-14.

62. Yu Y, Harris AJ, Blair C, He J. RASP (Reconstruct Ancestral State in Phylogenies): A tool for historical biogeography. Mol. Phylogenet. Evol. 2015; 87: 46-49.

63. Drummond AJ, Rambaut A, Shapiro B, Pybus OG. Bayesian coalescent inference of past population dynamics from molecular sequences. Mol. Biol. Evol. 2005; 22(5): 1185-1192.
64. Elith J, Graham CH, Anderson RP, Dudík M, Ferrier S, Guisan A, et al. Novel methods improve prediction of species' distributions from occurrence data. Ecography. 2006; 29(2): 129–151.
65. Phillips SJ, Anderson RP, Schapire RE. Maximum entropy modeling of species geographic distributions. Ecol. Modell. 2006; 190(3-4): 231–259.
66. Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. Very high resolution interpolated climate surfaces for the global land areas. Int. J. Climatol. 2005; 25(15): 1965–1978.
67. Pearson RG, Raxworthy CJ, Nakamura M, Peterson AT. Predicting species distributions from small numbers of occurrence records: A test case using cryptic geckos in Madagascar. J. Biogeogr. 2007; 34(1): 102–117.
68. Graham MH. Confronting multicollinearity in ecological multiple regression. Ecology. 2003; 84(11): 2809–2815.
69. Araujo MB, Pearson RG, Thuiller W, Erhard M. Validation of species-climate impact models under climate change. Glob. Change Biol. 2005; 11(9): 1504–1513.
70. Ferrari L, Rosas-Elguera J. Late Miocene to Quaternary extension and the northern boundary of the Jalisco block, western Mexico. The Tepic-Zacoalco rift revised. Geol. Soc. Am. Bull. Special Paper. 1999; 334: 1-23.
71. Rosas-Elguera J, Urrutia-Fucugauchi J. Tectonic control on the volcano-sedimentary sequence of the Chapala graben, western Mexico. Int. Geol. Rev. 1998; 40(4): 350–362.
72. Domínguez-Domínguez O, Pompa-Domínguez A, Doadrio I. A new species of the genus *Yuriria* Jordan & Evermann, 1896 (Actinopterygii, Cyprinidae) from the Ameca basin of the Central Mexican Plateau. Graellsia. 2007; 63(2): 259-271.
73. Pérez-Rodríguez R, Pérez Ponce de León G, Domínguez-Domínguez O, Doadrio I. A new species of *Algansea* (Actinopterygii: Cyprinidae) from the Ameca River basin, in Central Mexico. Rev. Mex. Biodivers. 2009b; 80(2): 483–490.

74. Alda F, Reina RG, Doadrio I, Bermingham E. Phylogeny and biogeography of the *Poecilia sphenops* species complex (Actinopterygii, Poeciliidae) in Central America. Mol. Phylogenet. Evol. 2013; 66: 1011-1026.

75. Garduño V, Tibaldi A. Kinematic evolution of the continental active triple junction of the western Mexican Volcanic Belt. C. R. Acad. Sci. 1991; 2: 135–142.

76. Webb T, Bartlein PJ. Global changes during the last 3 million years: climatic controls and biotic responses. Annu. Rev. Ecol. Evol. Syst. 1992; 23: 141–173.

77. Conant R. Observations on Garter Snakes of the *Thamnophis eques* Complex in the Lakes of Mexico´s Transvolcanic Belt, with Descriptions of New Taxa. American Museum Novitates. Bull. Am. Mus. Nat. Hist. 2003; 3406: 64.

78. Pacheco J, Mortera-Gutiérrez C, Delgado-Granados H, Singh S, Valenzuela R, Shapiro N, et al. Tectonic significance of an earthquake sequence in the Zacoalco half-graben, Jalisco, Mexico. J. South Am. Earth Sci. 1999; 12(6): 557–565.

79. Nelson SA, Carmichael ISE. Pleistocene to recent alkalic volcanism in the región of Sanganguey volcano, Nayarit, Mexico. Contrib. Mineral. Petrol. 1984; 85(4): 321-335.

80. Huang WD, Zhao XY, Zhao X, Zhao HL, Wang SK, Lian J. A combined approach using ISSR and ITS analysis for the characterization of *Artemisia halodendron* from Horqin sandy land, northern China. Biochem. Syst. Ecol. 2011; 39(4-6): 346-351.

81. Huang WD, Zhao XY, Zhao X, Li Y, Lian J. Effects of environmental factors on genetic diversity of *Caragana microphylla* in Horqin Sandy Land, northeast China. Ecol. Evol. 2016; 6(22): 8256-8266.

82. Su ZH, Zhang ML. A range wide geographic pattern of genetic diversity and population structure of *Hexinia polydichotoma* (Asteraceae) in Tarim Basin and adjacent areas. Biochem. Syst. Ecol. 2014; 56: 49-59.

83. Faye A, Deblauwe V, Mariac C, Damien R,Sonké B, Vigouroux Y, Couvreur TLP. Phylogeography of the genus *Podococcus* (Palmae/Arecaceae) in Central African rain forests: climate stability predicts unique genetic diversity. Mol. Phylogenet. Evol. 2016; 105: 126-138.

84. Perea S, Doadrio I. Phylogeography, historical demography and hábitat suitability modelling of freshwater fishes inhabiting seasonally fluctuating Mediterranean river systems: a case study using the Iberian cyprinid *Squalis valentinus*. Mol. Ecol. 2015; 24(14): 3706-3722.

85. Moncayo-Estrada R, Israde-Alcántara I, Garduño-Monroy VH. La cherehuita *Hubbsina turneri* De Buen (1941) (Pisces, Goodeidae). Origen, distribución y su uso en la regionalización de la cuenca del Lerma. Hidrobiologica. 2001; 11(1): 1-13.

86. Israde-Alcántara I. Lagos Volcánicos y Tectónicos de Michoacán. In: Garduno-Monroy VH, Corona-Chavez P, Israde-Alcántara I, MennellaL, Arreygue E, Bigioggero B, Chiesa, S, editors. Carta Geológica de Michoacán Escala 1:250000. Universidad Michoacana de San Nicolás de Hidalgo. 1999. p. 45-72.

87. Ornelas-García CP, Alda F, Díaz-Pardo E, Gutiérrez-Hernández A, Doadrio I. Genetic diversity shaped by historical and recent factors in the live-bearing twoline skiffia *Neotoca bilineata*. J. Fish Biol. 2012; 81(6): 1963-1984.

88. Corona-Santiago DK, Doadrio I, Domínguez-Domínguez O. Evolutionary History of the Live-bearing endemic *Allotoca diazi* species complex (Actinopterygii, Goodeinae): evidence of founder effect events in the Mexican Prehispanic period. Plos One. 2015; 10(5): e0124138.

89. Galindo-Villegas J, Sosa-Lima F. Gonopodial system review and a new fish record of *Poeciliopsis infans* (Cyprinodontiformes: Poeciliidae) for Lake Patzcuaro, Michoacan, central Mexico. Rev. Biol. Trop. 2002; 50(3-4): 1151-1157.

90. Ramírez-Herrejón JP, Castañeda-Sam LS, Moncayo-Estrada R, Caraveo-Patiño J, Balart EF. Trophic ecology of the exotic Lerma livebearer *Poeciliopsis infans* (Cyprinodontiformes: Poeciliidae) in the Lago de Pátzcuaro, Central Mexico. Rev. Biol. Trop. 2013; 61(3): 1289-1300.

91. Welcomme R. A history of international introductions of inland aquatic species. ICES Marine Science Symposia. 1992; 194: 3–14.

92. Contreras-MacBeath T, Mejía-Mojica H, Carrillo-Wilson R. Negative impact on the aquatic ecosystems of the state of Morelos, Mexico from introduced

aquarium and other commercial fish. Aquarium Sci. Conserv. 1998; 2(2): 67-78.

93. Caballero M, Lozano-García MS, Vázquez-Selem L, Ortega B. Evidencias de cambio climático y ambiental en las cuencas de altura del centro de México durante el último máximo glacial. Bol. Soc. Geol. Mex. 2010; 62(3): 359–377.
94. Hampe A, Petit R. Conserving biodiversity under climate change: the rear edge matters. Ecol. Lett. 2005; 8(5): 461-467.
95. Lachniet MS, Vázquez-Selem L. Last Glacial Maximum equilibrium line altitudes in the circum-Caribbean (Mexico, Guatemala, Costa Rica, Colombia, and Venezuela): Quat. Int. 2005; 138-139: 129-144.
96. Hofreiter M, Stewart JR. Ecological change, range fluctuations and population dynamics during the Pleistocene. Curr. Biol. 2009; 19(14): R584-R594.
97. Lozano-García S, Torres-Rodríguez E, Ortega B, Vázquez G, Caballero M. Ecosystem responses to climate and disturbances in western central Mexico during the late Pleistocene and Holocene. Paleogeogr. Palaeoclimatol. Palaeoecol. 2013; 370: 184-195.
98. Alfaro R, Martínez V, Segovia N, Peña P, López M, Armienta M, et al. Radon behavior in springs and Wells around Cuitzeo Lake, Lerma River Basin, Mexico. Geofís. Int. 2002; 41(4): 439-446.
99. De la Vega-Salazar MY, Avila-Luna E, Macías-García C. Ecological evaluation of local extinction: the case of two genera of endemic Mexican fish, *Zoogoneticus* and *Skiffia*. Biodivers. Conserv. 2003; 12(10): 2043-2056.
100. Domínguez-Domínguez O, Mercado-Silva N, Lyons J. Conservation status of Mexican goodeids: problems, perspectives, and solutions. In: Uribe MC, Grier HJ, editors. *Viviparous Fishes.* New Life Publications. 2005. p. 495-504.
101. Mercado-Silva N, Lyons JD, Maldonado GS, Nava MM. Validation of a fish-based index of biotic integrity for streams and rivers of Central Mexico. Rev. Fish Biol. Fish. 2002; 12(2): 179-191.

102. Domínguez-Domínguez O, Pérez Ponce de León G. Los goodeidos, peces endémicos del centro de México. CONABIO. Biodiversitas. 2007; 75: 12-15.
103. Domínguez-Domínguez O, Boto L, Alda F, Pérez-Ponce de León G, Doadrio I. Human impacts on drainages of the Mesa Central, Mexico, and its genetic effects on an endangered fish, *Zoogoneticus quitzeoensis*. Conserv. Biol. (2007b); 21(1): 168-180.
104. Doadrio I, Perdices A, Machordom A. Allozymic variation of the endangered killifish *Aphanius iberus* and its application to conservation. Env. Biol. Fish. 1996; 45(3): 259–271.

V. DISCUSIÓN GENERAL

Los procesos evolutivos que han moldeado la distribución, la estructura y la diversidad genética actual de los organismos en el Centro de México, han sido ampliamente estudiados en los últimos años, siendo el foco de interés de biólogos evolutivos y biogeógrafos (Bryson et al., 2011a,b,c,d; Domínguez-Domínguez et al., 2008a; 2010; Huidobro et al., 2006; Ornelas-García et al., 2012; Pedraza-Lara et al., 2012; Pérez-Rodríguez et al., 2009; 2015).

Los resultados obtenidos en esta investigación, contribuyen a incrementar el conocimiento sobre los patrones filogeográficos de cuatro grupos de especies de peces vivíparos, endémicos del Centro de México, tres pertenecientes a la subfamilia Goodeinae y uno a la Poeciliinae. Vale la pena notar como resultado de esta investigación los patrones contrastantes entre las especies estudiadas, con unas especies presentando fuerte estructura genética, mientras que otras con una baja o nula diferenciación genética entre regiones biogeográficas.

Por su parte, las especies del género *Ilyodon* (Goodeinae), representan de los pocos géneros de este grupo distribuidos en las cuencas hidrológicas de la costa del Pacífico, mostrando ser un grupo fuertemente estructurado, incluso dentro de cuencas hidrológicas. La alta plasticidad fenotípica reconocida para el género, ha generado controversias importantes en el número de especies que lo componen (Grudzien y Turner, 1984a,b; Turner y Grosse, 1980).

Poeciliopsis infans (Poeciliinae), *Goodea* spp. y *Xenotoca variata* (Goodeinae) representan las especies de estas subfamilias con el más amplio rango de distribución, encontrándolas tanto en cuencas que drenan al Pacífico, como cuencas que drenan al Atlántico, así como cuencas endorreicas (Miller et al., 2005). Los tres grupos de especies comparten su distribución en cuencas del Centro de México y presentan diferencias biológicas y ecológicas importantes. Mientras que *P. infans* y *X. variata* presenta marcado dimorfismo sexual, *G. atripinnis* es considerada una especie monomórfica (Ritchie et al., 2007), *G. atripinnis* y *X. variata* son especies de origen neártico, mientras que *P. infans* tiene un origen neotropical.

Dentro de los resultados obtenidos en las especies estudiadas, se demuestra que los procesos evolutivos encontrados en general son complejos,

como consecuencia de eventos paleogeográficos y paleoclimáticos ocurridos durante el Plioceno y el Pleistoceno y que persisten aún hasta nuestros días; lo que ha dado lugar a los patrones filogenéticos, filogeográficos, biogeográficos, así como de estructuración y diversidad genética en las especies *X. variata* (Beltrán-López et al., *en preparación*), el género *Ilyodon* (Beltrán-López et al., 2017) y en *P. infans* (Beltrán-López et al., 2018). Mientras que *Goodea* spp., que representa el grupo de peces más ampliamente distribuido en el Centro de México, mostró baja diferenciación genética y divergencia reciente entre las poblaciones, la cual no corresponde con la historia geológica y climática del Centro de México (Beltrán-López et al., *en preparación*).

V. I Patrones filogenéticos y filogeográficos

En las especies estudiadas, encontramos patrones filogenéticos y filogeográficos contrastantes. Por una parte para las especies del género *Ilyodon*, se encontró una politomía basal, por lo cual la historia evolutiva del género es dificíl de interpretar (Beltrán-López et al., 2017). A pesar de la politomía, se observan seis grupos genéticos bien diferenciados, de acuerdo a los clados y sub-clados obtenidos en los análisis filogenéticos, sin embargo, estos resultados no corresponden con la taxonomía reconocida del género.

De manera similar se observa una historia evolutiva compleja, para *Poeciliopsis infans* y *Xenotoca variata*, se encontraron dos linajes genéticos divergentes para cada una de las especies. En ambas especies, se reconoce que los dos clados bien diferenciados, indican una larga historia de aislamiento, con subsecuente diferenciación genética, la cual parece estar ligada a la intensa actividad volcánica y tectónica en el Centro de México, durante el Plioceno y el Pleistoceno, hace aproximadamente 2.83 millones de años (1.25-4.41 millones de años) para el caso de *P. infans* y 2.87 millones de años (1.44-4.31 millones de años) para el caso de *X. variata*. Aunque los tiempos de divergencia para ambas especies son similares, el evento cladogenético que separó a los dos linajes de cada una de las especies ocurrió en regiones biogeográficas diferentes.

Finalmente, un patrón contrastante fue reconocido en las especies del género *Goodea*, para el cual se encontraron diferencias genéticas mínimas

dentro y entre regiones biogeográficas a lo largo de su amplio rango de distribución, en la red de haplotipos con *cytb* se encontraron dos haplogrupos separados por dos pasos mutacionales y tiempo de divergencia entre estos dos haplogrupos de aproximadamente 0.81 millones de años (0.33-1.3 millones de años).

Para *Ilyodon* spp., debido a la politomía basal encontrada, no se llevó a cabo una filogenia fechada, sin embargo, la zona de distribución del género esta localizada en una zona que tuvo alta actividad volcánica durante el Plioceno y el Pleistoceno, hace 3.5 y 1.5 millones de años, especifícamente, la actividad de la triple unión (grabens de Tepic-Zacoalco, Chapala y Colima), el sistema de fallas de Tenochtitlán, y el sistema de fallas de Chapala-Oaxaca (García-Palomo et al., 2002; Garduño-Monroy et al., 1998; Rosas-Elguera et al., 1997). Por lo que se reconoce que la diversificación de este género, parece estar relacionada con la compleja historia geológica de los sistemas hidrológicos en el área, en el que patrones de conexión y desconexión se han establecido de acuerdo a las relaciones filogenéticas de otras especies de peces (Domínguez-Domínguez et al., 2010; Ornelas-García et al., 2008; Pérez-Rodríguez et al., 2009; 2015; Piller et al., 2015; Webb, 2002).

Para el caso de *P. infans* el evento cladogenético que separó a los dos clados, coincide con la interrupción de la conexión antigua de la parte alta del río Ameca con las cuencas del Centro de México, influido por la actividad volcánica y tectónica ocurrida hace aproximadamente 3-1 millones de años (Ferrari et al., 1999), esta barrera pudo haber bloqueado la ruta de dispersión entre los dos clados (Beltrán-López et al., 2018). Este evento vicariante ocurrió cuando el sistema hidrológico formado por el complejo Paleosistema de Chapala-Lerma (Ríos Ameca y Lerma, Lagos Magdalena, Chapala y los lagos distribuidos a lo largo del graben de Colima), comenzaron su aislamiento debido a la actividad volcánica y tectónica en el área de la triple unión (mismo evento que pudo haber provocado la diferenciación genética de *Ilyodon* spp.), y las fallas de Ameca y San Marcos, hace aproximadamente 3.5-1.5 millones de años (Rosas-Elguera et al., 1998).

Por otra parte, para *X. variata*, el evento cladogenético que separó a los dos linajes coincide con la actividad geológica de la región que tuvo lugar durante el Plioceno y el Cuaternario debido a las fallas geológicas alrededor del

lago de Cuitzeo, así como a la zona de fallas de la ciudad de Morelia y al graben de Penjamillo (Ferrari et al., 1994), además de la formación de campos volcánicos monogenéticos, el más prominente, el campo volcánico de Michoacán-Guanajuato con cerca de 1000 conos volcánicos distribuidos a lo largo de 40,000 Km^2 en el sector central de la FVTM (Hasenaka, 1994; Hasenaka y Carmichael, 1985), el vulcanismo en este sector comenzó en el Plioceno tardío (hace aproximadamente 2.87 millones de años (Hasenaka y Carmichael, 1985).

Para todas las especies, los análisis de varianza molecular implementados mostraron el mismo patrón, maximizando el coeficiente Φ_{CT} cuando las muestras se agruparon de acuerdo a los clados y sub-clados obtenidos en los análisis filogenéticos y no de acuerdo a las regiones biogeográficas propuestas por Domínguez-Domínguez et al., (2006).

Aunque para *X. variata, Ilyodon* spp. y *P. infans* los eventos vicariantes por actividad volcánica y tectónica son la primer causa de aislamiento de los linajes genéticos encontrados, los factores biológicos y ecológicos parecen haber influido en la historia evolutiva de manera diferente en las especies co-distribuidas estudiadas.

En el caso de *Ilyodon* spp., se han reconocido grupos con alta divergencia morfológica asociada con la estructura trófica, con variación en la forma y arreglo del cráneo, la boca y los dientes, además el tamaño de estos peces está influenciado por el hábitat y por la alimentación (Grudzien y Turner, 1984a,b; Turner y Grosse, 1980), esta divergencia morfológica puede estar asociada a la diferenciación genética encontrada inclusive en una misma cuenca, ya que se ha mostrado recientemente que los goodeinos radiaron en un escenario de oportunidades ecológicas a lo largo de los diferentes hábitats disponibles, por lo que se mostró que ha existido una rápida especiación y diversificación morfológica dentro de la subfamilia Goodeinae (Foster y Piller, 2018).

Para el caso particular de *P. infans*, considerando que es una especie de ancestría neotropical (mientras que el resto son neárticas), el patrón filogeográfico encontrado y la alta estructura genética, mostró que los dos linajes divergentes, tienen correspondencia con la altitud a la que se distribuyen, es decir, uno de los clados se distribuye en altitudes menores a los

1300 msnm, mientras que el segundo se encuentra por arriba de los 1300 msnm, por lo que en esta especie, su origen biogeográfico y los ciclos glaciales e interglaciales durante el cuaternario (i. e. temperaturas más extremas en las partes más altas) pueden ser factores determinantes en la distribución de la variación genética, esto considerando que a altitudes menores la temperatura del agua puede ser más alta y estable que a altitudes mayores (Beltrán-López et al., 2018).

En el caso de *X. variata*, aunque un evento vicariante debido a la actividad volcánica y téctonica, separó a los dos clados divergentes de la misma manera que en *P. infans*, la explicación más plausible para la subsecuente diferenciación genética de los dos clados de *X. variata* es debida a la fuerte selección sexual (Beltrán-López et al., *en preparación*), dado que *X. variata* es altamente dimórfica, con cortejos complejos y diferenciados entre linajes (Ritchie et al., 2007; Villa-Villaseñor, 2013), aunado a esto se ha reconocido que las regiones biogeográficas Zacapu-Lerma (clado I) y Cuitzeo (clado II) han estado conectadas en tiempos recientes, como lo indican trabajos previos en los que estas regiones comparten haplotipos o están estrechamente relacionados (Beltrán-López et al., 2018, en preparación; Domínguez-Domínguez et al., 2008; 2010; Ornelas-García et al., 2012; Pérez-Rodríguez et al., 2009; 2015), sin embargo para *X. variata* no hay haplotipos compartidos ni estrechamente relacionados, obteniendo una alta diferenciación genética entre ambos (Beltrán-López et al., en preparación).

El alto dimorfismo sexual que presenta *Xenotoca variata* puede explicar el patrón filogeográfico encontrado, ya que en trabajos previos se ha mostrado que las especies dimórficas presentan una diferenciación genética más grande entre sus poblaciones, sugiriendo que aunque los machos sean capaces de moverse entre poblaciones de estas especies, la reproducción para los que migran es menos exitosa (Ritchie et al., 2007), por lo que la selección sexual puede influir en la especiación. Si las hembras son más selectivas y los machos presentan divergencias morfológicas en diferentes poblaciones, el flujo genético será menor (Ritchie et al., 2007).

Aunque consideramos que para *P. infans* el factor más importante es su origen biogeográfico, también es una especie dimórfica, sin embargo, a diferencia de los goodeinos, presenta un órgano intromitente, con el cual

pueden fertilizar a la hembra sin necesidad de la total cooperación de la hembra, por lo cual pueden se un factor adicional para explicar el patrón filogeográfico encontrado.

Las características biológicas pueden ser determinantes de la historia evolutiva y los patrones filogeográficos encontrados en cada grupo estudiado; mientras que *X. variata* es una especie dimórfica, *Goodea* spp., representa un grupo de especies monomórficas. Esta característica de *Goodea* puede explicar la falta de estructuración genética encontrada. Se ha establecido que las poblaciones de especies monomórficas presentan baja diferenciación genética debido a que presentan mayor flujo genético por que las hembras son menos selectivas para elegir pareja, por lo que los machos migrantes tienen mayores posibilidades de reproducción (Ritchie et al., 2007), lo cual es concordante con el patrón encontrado en *Goodea* (Beltrán-López et al., *en preparación*), aunado a que es una especie con alto potencial reproductivo que es capaz de albergar una gran cantidad de embriones (Uribe et al., 2005).

Con lo anteriormente descrito acerca de las características de las especies dimórficas y las monomórficas, es evidente que aunque *X. variata* y *Goodea* spp., comparten características ecológicas similares, como adaptación a diferentes medios acuáticos, amplia distribución en el Centro de México, adecuación al tipo de alimento disponible y tolerancia a cambios ambientales (De la Vega-Salazar, 2006; Miller et al., 2005), las características biológicas de ambas representan un factor más importante en el patrón filogeográfico encontrado para cada una.

Los patrones filogeográficos encontrados en este trabajo, muestran que a pesar de que todas las especies investigadas son peces dulceacuícolas, víviparos y endémicos al Centro de México, la actividad geológica y tectónica ha afectado de manera diferente la historia evolutiva de estas (*Ilyodon* spp. *Poeciliopsis infans*, *Goodea atripinnis* y *Xenotoca variata*). Además, una contribución importante de este trabajo, es que a pesar de que se ha establecido generalmente que la historia geológica del centro de México ha moldeado la distribución y variación genética a lo largo de la distribución de las poblaciones de las especies (Beltrán-López et al., 2017; 2018; Domínguez-Domínguez et al., 2008; 2010; Pedraza-Lara et al., 2012; Pérez-Rodríguez et al., 2009; 2015; 2016), aquí mostramos que es evidente que las características

biológicas y ecológicas también tienen un papel muy importante en los patrones filogeográficos de las especies (Beltrán-López et al., *en preparación*).

V. II Historia demográfica

La historia demográfica de las poblaciones para las especies *P. infans* mostraron un declive demográfico en el último interglacial (150,000-100,000 años) seguido de una expansión poblacional en el último máximo glacial (25,000-18,000 años), la expansión poblacional fue más evidente para el clado distribuido en las regiones biogeográficas de menor altitud. Para *Goodea*, se muestran tamaños efectivos poblacionales estables seguidos de expansión poblacional hace 30,000 años para un haplogrupo, mientras que para el segundo haplogrupo la expansión poblacional fue fechada hace aproximadamente 125,000 años, finalmente, para *X. variata*, el linaje genético de Cuitzeo mostró una expansión poblacional en los últimos 200,000 años, mientras que para el resto de los grupos la expansión poblacional se dio en los últimos 40,000 años.

Las diferencias encontradas en los resultados de expansión poblacional, pueden ser atribuidos a varios factores, como el origen biogeográfico, es evidente que las continuas fluctuaciones climáticas y los niveles de agua en sistemas hidrológicos debido a los ciclos glaciales e interglaciares (Caballero et al., 2010), durante los últimos 200,000 años han afectado la historia demográfica de manera diferente a las especies estudiadas y en la especie neotropical *P. infans* estas fluctuaciones han provocado declive con subsecuente expansión poblacional. En el caso contrario, para *Goodea* y *X. variata*, la expansión poblacional para ambas especies se da en diferentes periodos de tiempo, por lo que las caracteristicas biológicas y ecológicas de estas dos especies podrían ser un factor determinante para que las fluctuaciones climáticas en este periodo no tuvieran el mismo efecto que tuvieron en *P. infans* (Beltrán-López et al., 2018; *en preparación*).

V. III Implicaciones taxonómicas

Los resultados obtenidos para todos los grupos de especies analizadas (*Ilyodon* spp., *Poeciliopsis infans*, *Goodea* spp., y *Xenotoca variata*), tienen diferentes y nuevas implicaciones en su taxonomía.

En el caso de *Ilyodon* spp., y *Goodea* spp., en el que se han descrito cinco y tres diferentes especies respectivamente, se mostró que especies taxonomicamente reconocidas comparten haplotipos o están cercanamente relacionadas, con distancias genéticas menores al 1% con *cytb* (Beltrán-López et al., 2017; en preparación). En el caso contrario, *Poeciliopsis infans* y *Xenotoca variata* mostraron alta diferenciación genética entre los dos linajes recuperados, encontrando distancias genéticas entre estos dos linajes mayores al 3% con *cytb* (Beltrán-López et al., 2018; *en preparación*).

Para el caso de *Ilyodon* spp., se recuperaron seis grupos genéticos, que aunque tienen correspondencia geográfica, no tienen correspondencia taxonómica. Pruebas de delimitación de especies fueron implementadas, sin embargo, no fue posible resolver la taxonomía del género, debido a diferentes factores como la politomia basal obtenida y, a que se ha reportado que en especies incipientes o con sorteo de linajes incompleto, estas pruebas pueden no ser eficientes debido a la reciente diferenciación, las cuales son difíciles de distinguir únicamente con caracteres moleculares o morfológicos (Eberle et al., 2016; Heled et al., 2013).

No todas las especies de *Ilyodon* fueron identificadas como monofiléticas en los resultados filogenéticos, mientras que si consideramos que todas han sido descritas en base a su morfología, el presente trabajo, muestra que la actual taxonomía del género no tiene correspondencia con los resultados moleculares, encontrando estructura geográfica significativa pero que no es concordante con las cinco especies reconocidas previamente: *I. furcidens*, *I. xantusi*, *I. whitei*, *I. lennoni* e *I. cortesae* mostrando relaciones polifiléticas. Aunque algunas sugerencias taxonómicas fueron realizadas, consideramos que un estudio con mayor número de muestras, así como la inclusión de mas caracteres genéticos y morfológicos, permitirán tener un mejor panorama de la evolución de los grupos divergentes de *Ilyodon* (Beltrán-López et al., 2017).

En el caso de *Goodea* se han descrito tres especies: *G. atripinnis*, *G. luitpoldii* y *G. gracilis*. Sin embargo, organismos analizados pertenecientes a las tres especies descritas no fueron recuperadas como monofiléticas y comparten haplotipos entre ellas, sugiriendo que reconocer tres especies para el género no es taxonómicamente adecuado, aunque sugerimos que una revisión morfológica detallada es necesaria para robustecer los resultados y por lo tanto llevar a cabo las adecuaciones nomenclaturales pertinentes, sin embargo en el presente trabajo con datos moleculares y bajo el principio de prioridad, sugerimos que el género debería reconocerse como monotípico, siendo *G. atripinnis* la única especie válida.

Ambos géneros *Ilyodon* y *Goodea* exhiben alta variación morfológica, lo que ha llevado a la descripción de especies nuevas, sin embargo, se ha reconocido, que las poblaciones muestran respuestas fenotípicas similares de acuerdo al ambiente donde se distribuyen las poblaciones y al flujo de las corrientes, mostrando respuestas adaptativas a hábitats divergentes, considerando que la plasticidad fenotípica puede promover respuestas morfológicas como resultado de los cambios ambientales (Crispo, 2008; Foster et al., 2015).

Para el caso de *P. infans* y *X. variata*, los dos linajes divergentes encontrados (distancias genéticas de 3.3% y 5% respectivamente con *cytb*), sugieren eventos de especiación, debido a que incluso presentan valores más altos que el promedio mostrado para las especies de goodeinos (Doadrio y Domínguez, 2004; Domínguez-Domínguez et al., 2010) y son superiores al valor del 2% que se ha sugerido como el valor mínimo propuesto para el límite entre dos especies hermanas con el *cytb* para vertebrados (Avise, 1998, 2000; Bradley y Baker, 2001).

El evento cladogenético que separó los dos linajes de estas especies tiene correspondencia con la actividad volcánica y tectónica durante el Plioceno y Pleistoceno en el Centro de México, sin embargo, la diferenciación genética después del aislamiento por barreras geográficas, puede deberse a factores ecológicos y biológicos propios de cada especie. En el caso de *P. infans*, su origen biogeográfico puede haber influenciado la diferenciación genética de los dos linajes, considerando que uno se distribuye en altitudes por debajo de 1300

msnm y el otro por arriba de esta altitud, además las condiciones climáticas pudieron haber permitido esta diferenciación (Beltrán-López et al., 2018).

De manera contraria, para *X. variata*, la fuerte selección sexual que existe al ser una especie dimórfica, ha ocasionado que los ejemplares de los dos linajes presenten comportamientos diferentes en el cortejo, como se ha mostrado en trabajos previos (Fitzsimons, 1976; Villa-Villaseñor, 2013), aunado a esto, se ha mostrado también que las especies dimórficas presentan menos flujo genético entre las poblaciones, y por lo tanto mayor diferenciación genética de las poblaciones con subsecuentes eventos de especiación (Ritchie et al., 2007), como es el caso de *X. variata*.

Sugerimos para *Ilyodon* spp. *P. infans* y *X. variata* una revisión taxonómica que incluya análisis morfológicos, en el caso de *Ilyodon* para tener un panorama más completo del número y las especies que lo componen, mientras que para *P. infans* y *X. variata*, la revisión taxonómica es necesaria, para que en su caso, las especies nuevas sean descritas.

V. IV Implicaciones para la conservación de los peces del Centro de México

El Centro de México, ha sido fuertemente impactado por las actividades antropogénicas provocando pérdida y degradación del hábitat debido a sobreexplotación, contaminación, degradación del hábitat y la introducción de especies no nativas (De la Vega-Salazar, 2006; Domínguez-Domínguez et al., 2005, 2007; Mercado-Silva et al., 2002). Estas actividades podrían provocar la pérdida de diversidad genética de las poblaciones, reduciendo así su potencial evolutivo. Bajo este escenario, la diversidad genética calculada para *P. infans*, *G. atripinnis* y *X. variata*, mostró que en *P. infans* la diversidad genética más alta se encontró para las regiones biogeográficas de menores altitudes.

A diferencia de *P. infans*, para *G. atripinnis* la diversidad genética en general presentó valores moderados a bajos, con algunas excepciones (Beltrán-López et al., *en preparación*) y para *X. variata*, los valores de diversidad genética tienden a ser bajos, debido a esto, sugerimos que para *Goodea atripinnis* que es una especie ampliamente distribuida con baja diferenciación genética, no se observan Unidades Operacionales de Conservación (OCUs por sus siglás en inglés).

Para *X. variata* y *P. infans* varias regiones biogeográficas pueden considerarse como unidades operacionales de conservación (OCUs por sus siglás en inglés) debido a los valores de diversidad genética y a la diferenciación genética encontrada entre las regiones biogeográficas.

En el caso de *Ilyodon*, aunque la diversidad genética no fue calculada, debe considerarse la aplicación de medidas de conservación para los grupos genéticos encontrados, debido a que se han establecido las especies de este género como especies vulnerables (NOM-059) y considerando la diferenciación genética encontrada, el establecimiento de medidas de conservación permitirá la conservación del pool genético encontrado (Beltrán-López et al., 2017).

VI. CONCLUSIONES GENERALES

1. La actividad volcánica y téctonica en el Centro de México durante el Plioceno y el Pleistoceno, ha sido el principal factor en la historia evolutiva y subsecuente diferenciación genética de tres de los cuatro grupos de especies que se analizaron (*Ilyodon* spp., *Poeciliopsis infans* y *Xenotoca variata*), encontrando patrones filogeográficos complejos y fuerte estructura genética.

2. Las características ecológicas y biológicas de cada grupo y especie analizados, representan un factor importante en la diferenciación genética encontrada en todos los grupos, mostrando diferencias importantes y por lo tanto patrones filogeográficos particulares para cada especie analizada.

3. *Goodea* spp., representa una excepción a los patrones filogeográficos encontrados generalmente para los peces dulceacuícolas del centro de México, encontrando baja diferenciación y variabilidad genética entre todas las regiones biogeográficas en donde se distribuye, este patrón es atribuible a las características ecológicas y biológicas presentes en este grupo, mientras que la actividad volcánica y tectónica no son factores que han contribuido a la estructura genética de las poblaciones.

4. Para *P. infans*, su origen biogeográfico, a diferencia de las especies de origen neártico (goodeinos) ha influido en su estructura genética, su diversidad

genética y su historia demográfica, mostrando que aunque es capaz de colonizar regiones de altas altitudes, las poblaciones que se distribuyen en áreas con menor altitud, presentan la mayor diversidad genética, por lo tanto mayor potencial evolutivo, a diferencia de las poblaciones distribuidas en las partes altas.

5. Para los goodeinos (*Ilyodon* spp., *Goodea* spp., y *X. variata*), el origen biogeográfico no parece tener una influencia en su patrón filogeográfico, pero las caraterísticas ecológicas como tipo de alimentación, capacidad de colonización y dispersión, rango de distribución y ambientes en donde se distribuye, así como las características biológicas como marcado o bajo dimorfismo sexual, selección sexual, cortejo prenupcial, y potencial reproductivo, parecen ser factores de mayor importancia en la baja o fuerte estructura genética encontrada, así como en la diferenciación y variabilidad genética de estas especies.

6. Las fluctuaciones climáticas durante el cuaternario, en particular los ciclos glaciares e interglaciares han moldeado la diversidad genética y la demografía histórica de *P. infans*, mientras que para los goodeinos, los cambios climáticos que han ocurrido en los últimos 200,000 años en el Centro de México, han influenciado de manera positiva a estas especies con expansiones poblacionales.

7. Una revisión taxonómica es necesaria para todos los grupos estudiados, con el fin de evaluar la variación genética encontrada desde una perspectiva taxonómica.

REFERENCIAS

Arellano, E., González-Cozátl, F. X. & Rogers, D. S. (2005). Molecular systematics of Middle American harvest mice *Reithrodontomys* (Muridae), estimated from mitochondrial cytochrome b gene sequences. *Molecular Phylogenetics and Evolution*. 37(2): 529–540.

Allendorf, F. W. (1983). Isolation, gene flow, and genetic differentiation among populations. *Genetics and Conservation*. 18(3): 51-65.

Andrés, J. A. & Bogdanowicz. (2011). Isolating microsatellite loci: looking back, looking ahead. Methods in Molecular Biology. 722: 211-232.

Antonelli, A. & Sanmartín, I. (2011). Mass extinction, gradual cooling, or rapid radiation? Reconstructing the spatiotemporal evolution of the ancient angiosperm genus *Hedyosmun* (Chloranthaceae) using empirical and simulated approaches. *Systematic Biology.* 60(5): 596-615.

Avise, J. C. (1998). The history and purview of phylogeography: a personal reflection. *Molecular Ecology*. *7:* 371-379.

Avise, J. C. (2000). Phylogeography: The history and formation of species. Harvard Univ. Press, Cambridge. 464 pp.

Avise, J. C. (2008). Phylogeography: retrospect and prospect. *Journal of Biogeography*. 36: 3–15.

Avise, J. C. & Ball, Jr, R. M. (1990). Principles of genealogical concordance in species concepts and biological taxonomy. *Oxford Surveys in Evolutionary Biology*. 7: 45–67.

Avise, J. C., Arnold, J., Ball, R. M., Bermingham, E., Lamb, T., Neigel, J. E., Reeb, C. A. & Saunders, N. C. (1987). Intraspecific phylogeography: The mitochondrial DNA bridge between population genetics and systematics. *Annual Review in Ecology and Systematics*, 18, 489-522.

Beltrán-López, R. G., Domínguez-Domínguez, O., Guerrero, J. A., Corona-Santiago, D. K., Mejía-Mojica, H., & Doadrio, I. (2017). Phylogeny and taxonomy of the genus *Ilyodon* Eigenmann, 1907 (Teleostei: Goodeidae), based on mitochondrial and nuclear DNA sequences. *Journal of Zoological Systematics and Evolutionary Research*. 55(4): 340-355. DOI: 10.1111/jzs.12175

Beltrán-López, R. G., Domínguez-Domínguez, O., Pérez-Rodríguez, R., Piller, K. R., & Doadrio, I. (2018). Evolving in the highlands: The case of the Neotropical Lerma Live-bearing *Poeciliopsis infans* (Woolman, 1894) (Cyprinodontiformes: Poeciliidae) in central México. *BMC Evolutionary Biology*. 18: https://doi.org/10.1186/s12862-018-1172-7.

Beltrán-López, R. G., Domínguez-Domínguez, O., Piller, K. R., Mar-Silva, A. F. & Doadrio, I. Genetic differentiation among populations of the freshwater

fish *Goodea atripinnis* Jordan, 1880 (Cyprinodontiformes: Goodeidae), implications for its evolutionary history and taxonomy. *En preparación.*

Beltrán-López, R. G., Domínguez-Domínguez, O., Piller, K. R., Perea, S., Pérez-Ponce de León, G. & Doadrio, I. Processes determining the cladogenetic events among *Xenotoca variata* (Bean, 1887) (Cyprinodontiformes: Goodeidae). *En preparación*.

Berra, T. M. (2001). Feshwater fish distribution. Academic Press, Nueva York-Londres. pp 615.

Bermingham, E. & Martin, A. P. (1998). Comparative mtDNA phylogeography of neotropical freshwater fishes: testing shared history to infer the evolutionary landscape of lower Central America. *Molecular Ecology*. 7(4): 499-517.

Betancourt-Resendes, I., Pérez-Rodríguez, R. & Domínguez-Domínguez, O. (2018). Speciation of silverside *Chirostoma attenuatum* (Pisces: Atheriniformes) in Central Mexico. *Journal of Zoological Systematics and Evolutionary Research*, 56(3), 323-334. DOI: 10.1111/jzs.12216

Bradley, R. D. & Baker, R. J. (2001). A test of the genetic species concept: Cytochrome-b sequences and mammals. *Journal of Mammalogy*. 82: 960-973.

Brawand, D., Wagner, C. E., Li, Y. I., Malinsky, M., Keller, I., Fan, S. *et al*. (2014). The genomic substrate for adaptive radiation in African cichlid fish. *Nature.* 513: 375-381.

Brusatte, S. L., Benton, M. J., Ruta, M. & Lloyd, G. T. (2011). The first 50Myr of dinousar evolution: macroevolutionary pattern and morphological disparity. *Biology Letters*. 4(6): 733-736.

Bryson, R. W. Jr., García-Vázquez, U. O. & Riddle, B. R. (2011a). Phylogeography of Middle American gopher-snakes: mixed responses to biogeographical barriers across the Mexican Transition Zone. *Journal of Biogeography*. 38: 1570–1584.

Bryson, R. W. Jr., García-Vázquez, U. O. & Riddle, B. R. (2011b). Relative roles of Neogene vicariance and Quaternary climate change on the historical diversification of bunchgrass lizards (*Sceloporus scalaris* group) in Mexico. *Molecular Phylogenetics and Evolution*. 62: 447–457.

Bryson, R. W. Jr., Murphy, R. W., Lathrop, A. & Lazcano-Villareal, D. (2011c).

Evolutionary drivers of phylogeographical diversity in the highlands of Mexico: a case study of the *Crotalus triseriatus* species group of montane rattle-snakes. *Journal of Biogeography*. 38: 697–710.

Bryson, R. W. Jr., Murphy, R. W., Graham, M. R., Lathrop, A. & Lazcano-Villareal, D. (2011d). Ephemeral Pleistocene woodlands connect the dots for highland rattlesnakes of the *Crotalus intermedius* group. *Journal of Biogeography*. 38: 2299–2310.

Burbrink, F. T. & Pyron, R. A. (2010). How does ecological opportunity influences rates of speciation, extinction, and morphological diversification in new world ratsnakes (Tribe Lampropeltini)?. *Evolution*. 64(4): 934-943.

Caballero, M., Lozano-García, M. S., Vázquez-Selem, L. & Ortega, B. (2010). Evidencias de cambio climático y ambiental en las cuencas de altura del centro de México durante el último máximo glacial. *Boletin de la Sociedad Geológica Mexicana*. 62(3): 359–377.

Castoe, T. A., Daza, J. M., Smith, E. N., Sasa, M. M., Kuch, U., Campbell, J. A. et al. (2009). Comparative phylogeography of pitvipers suggests a consensus of ancient Middle American highland biogeography. *Journal of Biogeography*. 36(1): 88–103.

Corona-Santiago, D. K., Doadrio, I. & Domínguez-Domínguez, O. (2015) Evolutionary History of the Live-bearing endemic *Allotoca diazi* species complex (Actinopterygii, Goodeinae): evidence of founder effect events in the Mexican Prehispanic period. *Plos One*. 10: e0124138.

Correa-Ramírez, M. M. (2010). Análisis de la diversidad genética de Pardosa sierra Banks, 1898 (Araneae: Lycosidae) en la península de Baja California, México. Tesis Doctoral. Centro de Investigaciones Biológicas del Noroeste, S. C.

Cracraft, J. (1983). Species concepts and speciation analysis. *Current Ornithology*. 1: 159–187.

Crispo, E. (2008). Modifying effects of phenotypic plasticity on interactions among natural selection, adaptation and gene flow. *Journal of Evolutionary Biology*. 21: 1460–1469.

Darwin, C. (1859). On the origin of species by means of natural selection or the preservation of favoured races in the struggle for life. John Murray, London. 449 pp.

Davey, J. W., Hohenlohe, P. A., Etter, P. D., Boone, J. Q., Catchen, J. M. & Blaxter, M. L. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics*. 12: 499–510.

Davis, J. I. & Nixon, K. C. (1992). Population, genetic variation, and delimitation of phylogenetic species. *Systematic Biology*. 41: 421–435.

De la Vega-Salazar, M. Y. (2006). Estado de Conservación de los peces de la familia Goodeidae (Cyprinodontiformes) en la mesa central de México. *Revista de Biología Tropical*. 54(1): 163-177.

Derek, E., Briggs, G., Fortey, R. A. & Wills, M. (1992). Morphological disparity in the Cambrian. *Science*. 256: 5064.

Doadrio, I. & Domínguez, O (2004). Phylogenetic relationships whitin the fish family Goodeidae based on cytochrome b sequence data. *Molecular Phylogenetics and Evolution*. 31: 416–430. PMID: 15062784

Dobzhansky, T. (1937). Genetics and the origin of species. Columbia University Press. 364 pp.

Domínguez-Domínguez, O., Alda, F., Pérez-Ponce de León, G., García-Garitagoitia, J. L. & Doadrio, I. (2008a). Evolutionary history of the endangered fish *Zoogoneticus quitzeoensis* (Bean, 1898) (Cyprinodontiformes: Goodeidae) using a sequential approach to phylogeography based on mitochondrial and nuclear DNA data. *BMC Evololutionary Biology.* 8: 161.

Domínguez-Domínguez, O., Bernal-Zuñiga, D. M., & Piller, K. R. (2016). Two new species of the genus *Xenotoca* Hubbs and Turner, 1939 (Teleostei, Goodeidae) from central-western México. *Zootaxa*. 4189: 81-98.

Domínguez-Domínguez, O., Boto, L., Alda, F., Pérez-Ponce de León, G. & Doadrio, I. (2007). Human impacts on drainages of the Mesa Central, Mexico, and its genetic effects on an endangered fish, *Zoogoneticus quitzeoensis*. *Conservation Biology*. 21(1): 168-180.

Domínguez-Domínguez, O., Doadrio, I., Martínez-Meyer, E., Zambrano, L. & Pérez-Ponce de León, G. (2006). Using ecological niche modeling as a conservation tool for freshwater species: the live-bearing fishes (Goodeidae) in Central Mexico. *Conservation Biology*. 20: 1730–1739. PMID: 17181808

Domínguez-Domínguez, O., Mercado-Silva, N. & Lyons, J. (2005). Conservation status of Mexican goodeids: problems, perspectives, and solutions. En: *Viviparous Fishes*. Uribe MC, Grier HJ, (eds). New Life Publications, Florida, pp 495-504.

Domínguez-Domínguez, O., Pedraza-Lara, C., Gurrola-Sánchez, N., Perea, S., Pérez-Rodríguez, R., Israde-Alcántara, I. et al. (2010). Historical Biogeography of the Goodeinae (Cyprinodontiformes). En:, Viviparous Fishes II. Uribe-Aranzabal MC, Grier HJ (eds). New Life Publications, Florida, pp 13–30.

Domínguez-Domínguez, O. & Pérez Ponce de León, G. (2007). Los goodeidos, peces endémicos del centro de México. CONABIO. *Biodiversitas*. 75: 12–15.

Domínguez-Domínguez, O., Pérez-Rodríguez, R. & Doadrio I. (2008b) Morphological and genetic comparative analyses of populations of *Zoogoneticus quitzeoensis* (Cyprinodontiformes: Goodeidae) from Central Mexico, with description of a new species. *Revista Mexicana de Biodiversidad*. 79(2): 373-383.

Domínguez-Domínguez, O. & Vázquez-Domínguez, E. (2009). Filogeografía: aplicaciones en taxonomía y conservación. Animal Biodiversity and Conservation. 32(1): 59-70.

Eberle, J., Warnock, R. C. M. & Ahrens, D. (2016) Bayesian species delimitation in *Pleophylla* chafers (Coleoptera) – the importance of prior choice and morphology. *BMC Evolutionary Biology*. 16: 94. doi: 10.1186/s12862-016-0659-3

Egan, A. N., Schlueter, J. & Spooner, D. M. (2012). Applications of next-generation sequencing in plant biology. *American Journal of Botany*. 99: 175–185.

Ekblom, R. & Galindo, J. (2011). Applications of next generation sequencing in molecular ecology of non-model organisms. *Heredity*. 107: 1–15.

Emerson, K. J., Merz, C. R., Catchen, J. M., Hohenlohe, P. A., Cresko, W. A., Bradshaw, W. E. & Holzapfel, C. M. (2010). Resolving postglacial phylogeography using high-throughput sequencing. *Proceedings of the National Academy of Sciences of the United States of America*. 107: 16196–16200.

Eschmeyer WN, Fricke R Dan der Laan R (eds). **CATALOG OF FISHES: GENERA, SPECIES, REFERENCES.** (http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp). Accessed online 1 March 2016. [This version was edited by Bill Eschmeyer.]

Ferrari, L., Conticelli, S., Vaggelli, G., Potrone, C. M. & Manetti, P. (2000) Late Miocene volcanism and intra-arc tectonics during the early development on the Trans-Mexican volcanic belt. *Tectonophysics*. 318: 161–185.

Ferrari, L., Garduño, V., Pasqueré, G. & Tibaldi, A. (1994). Volcanic and tectonic evolution of central Mexico: Oligocene to present. *Geofísica Internacional*. *33*: 107–123.

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.

Ferrari, L. & Rosas-Elguera, J. (1999). Late Miocene to Quaternary extension and the northern boundary of the Jalisco block, western Mexico. The Tepic-Zacoalco rift revised. *Geological Society of American Bulletin Special Paper*. 334: 1-23.

Fitzsimons, J. M. (1976). Ethological isolating mechanisms in goodeid fishes of the genus *Xenotoca* (Cyprinodontiformes, Osteichthyes). *Bulletin southern California academy of sciences*. 75: 84-99.

Foster, K., Bower, L., & Piller, K. (2015). Getting in shape: habitat-based morphological divergence for two sympatric fishes. *Biological Journal of the Linnean Society*. 114,:152-162.

Foster, K. L. & Piller, K. R. (2018). Disentangling the drivers of diversification in an imperiled group of freshwater fishes (Cyprinodontiformes: Goodeidae). *BMC Evolutionary Biology*. 18: 116.

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.* Boulder, Colorado, Geological Society of America Map and Chart Series MCH089, 26 p.

Garduño-Monroy, V., Saucedo-Girón, R., Jiménez, Z., Gavilanes-Ruiz, J., Cortés, A. & Uribe-Cifuentes, R. (1998). La falla Tamazula-límite suroriental del bloque Jalisco y sus relaciones con el complejo volcánico

de Colima, México. *Revista Mexicana de Ciencias Geológicas*. 15: 132–144.

Godoy, J. A. (2009). La genética, los marcadores moleculares y la conservación de especies. *Ecosistemas*. 18(1): 23-33.

Goto, A. & Andoh, T. (1990). Genetic divergence between the sibling species of river-sculpin, *Cottus amblystomopsis* and *C. Nozawae*, with special reference to speciation. *Environmental Biology of Fishes*, 28, 257-266.

Grudzien, T. A. & Turner, B. J. (1984a). Genic identity and Geographic Differentiation of Trophically Dichotomous *Ilyodon* (Teleostei: Goodeidae). *Copeia*. 1984: 102–107.

Grudzien, T. A. & Turner, B. J. (1984b). Direct evidence that the *Ilyodon* morphs are a single biological species. *Evolution*. 38: 402–407.

Guichoux, E., Lagache, L., Wagner, S., Chaumeil, P., Léger, P., Lepais, O. Et al. (2011). Current trends in microsatellite genotyping. *Molecular Ecology Resources*. 11: 591-611.

Guzmán, A. F. (2015). El registro fósil de los peces mexicanos de agua dulce. *Revista Mexicana de Biodiversidad*. 86(3): 661-673.

Halffter, G. (1987). Biogeography of the montane entomofauna of Mexico and Central America. *Annual Review of Entomology*. 32: 95-114.

Halffter, G. (2003). Biogeografía de la entomofauna de montaña de México y América Central. En: Morrone JJ, Llorente-Bousquets J (eds). Una perspectiva latinoamericana de la biogeografía. Las Prensas de Ciencias, UNAM, Mexico City. pp. 217–220.

Hallgrimsson, B., Jamniczky, H. A., Young, N. M., Rolian, C., Schmidt-Ott, U. & Marcucio, R. S. (2012). The generation of variation and the developmental basis for evolutionary novelty. *Journal of Experimental Zoology*. 318(6): 501-571.

Harmon, L. J., Schulte, J. A., Larson, A. & Losos, J. B. (2003). Tempo and mode of evolutionary radiation in Iguanian lizards. *Science*. 301(5635): 961-964.

Hasenaka, T. (1994). Size, distribution, and magma output rate for shield volcanoes of the Michoacan-Guanajuato volcanic field, Central Mexico. *Journal of Volcanology and Geothermal Research*. 63: 13–31. doi:10.1016/0377-0273(94)90016-7.

Hasenaka, T. & Carmichael, I. (1985). The cinder cones at Michoacan–Guanajuato, Central Mexico: their age, volume, and distribution, and magma discharge rate. *Journal of Volcanology and Geothermal Research*. 25: 105–124. doi:10.1016/0377-0273(85) 90007-1.

Heled, J., Bryant, D. & Drummond, A. J. (2013). Simulating genes trees under the multispecies coalescent and time-dependent migration. *BMC Evolutionary Biology*. 13: 44. doi: 10.1186/1471-2148-13-44

Hey, J. & Machado, C. A. (2003). The study of structured populations–new hope for a difficult and divided science. *Nature Reviews Genetics.* 4: 53–543.

Hewitt, G. M. (2000). The genetic legacy of the Quaternary ice ages. *Nature*. 405(6789): 907–913.

Hewitt, G. M. (2004). Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society B*. 359(1442): 183-195.

Hrbek, T., Seckinger, J. & Meyer, A. (2007). A phylogenetic and biogeographic perspective on the evolution of poeciliid fishes. *Molecular Phylogenetics and Evolution*. 43: 986-998.

Huidobro, L., Morrone, J. J., Villalobos, J. L. & Álvarez, F. (2006). Distributional patterns of freshwater taxa (fishes, crustaceans and plants) from the Mexican transition zone. *Journal of Biogeography*, 33, 731-741.

Israde-Alcántara, I., Velázquez-Durán, R., Lozano-García, M. S., Bischoff, J., Domínguez-Vázquez, G., Garduño-Monroy, V. H. (2010). Evolución paleolimnológica del lago Cuitzeo, Michoacán durante el Pleistoceno-Holoceno. *Boletín de la Sociedad Geológica Mexicana*. 62(3): 345-357.

Iturralde-Vinent, M. A. & MacPhee, R. D. E. (1999). Paleogeography of the Caribbean region: Implications for Cenozoic biogeography. *Bulletin of the American Museum of Natural History*. 238.

Kelly, J. K. (2000). Epistasis, linkage, and balancing selection. En: Wolf JB, Brodie IED, Wade MJ (eds.) Epistasis and the Evolutionary Process. Oxford University Press, New York, USA, pp. 146-157.

Lanteri, A. & Confalonieri, V. (2003). Filogeografía: objetivos, métodos y ejemplos. En: Morrone JJ, Llorente J (eds.) Una prespectiva latinoamericana de la biogeografía. CONABIO, México, pp. 185–194.

Lexer, C., Mangili, S., Bossolini, E., Forest, F., Stölting, K. N., Pearman, P. B., et al. (2013). 'Next generation' biogeography: Towards understanding the drivers of species diversification and persistence. *Journal of Biogeography*. 40: 1013–1022.

Lombardi, J. & Wourms, J. P. (1985). The trophotaenial placenta of a viviparous goodeid fish. I. Ultrastructure of the internal ovarian epithelium, the maternal component. *Journal of Morphology.* 184: 277-292.

Lucinda PHF. Family Poeciliidae. In: Reis RE, Kullander SO, Ferraris CJJr, editors. Check list of the freshwater fishes of South and Central America. Porto Alegre: EDIPUCRS. 2003. p. 555-581.

Lucinda, P.H.F. & Reis, R.E. (2005). Systematics of the subfamily Poeciliinae Bonaparte (Cyprinodontiformes: Poeciliidae), with an emphasis on the tribe Cnesterodontini Hubbs. *Neotropical Ichthyology*. 3: 1–60.

Mateos, M., Sanjur, O. I. & Vrijenhoek, R. C. (2002). Historical biogeography of the livebearing fish genus *Poeciliopsis* (Poeciliidae: Cyprinodontiformes). *Evolution*. 56: 972–984.

Mejía-Madrid, H., Vásquez-Domínguez, E. & Pérez Ponce de León, G. (2007). Phylogeography and freshwater basins in central Mexico: recent history as revealed by the fish parasite *Rhabdochona lichtenfelsi* (Nematoda). *Journal of Biogeography*. 34: 787-801.

McCormack, J. E., Hird, S. M., Zellmer, A. J., Carstens, B. C., & Brumfield, R. T. (2013). Applications of next-generation sequencing to phylo- geography and phylogenetics. Molecular Phylogenetics and Evolution, 66, 526–538.

McKitrick, M. C. & Zink, R. M. (1988). Species concept in ornithology. *Condor*. 90: 1–14.

Mercado-Silva, N., Lyons, J. D., Maldonado, G. S. & Nava, M. M. (2002). Validation of a fish-based index of biotic integrity for streams and rivers of Central Mexico. *Reviews in Fish Biology and Fisheries*. 12(2): 179-191.

Miller, R. R., Minckley, W. L. & Norris, S. M. (2005). Freshwater Fishes of México. Chicago, IL: The University of Chicago Press.

Miller, R. R. & Smith, M. L. (1986). Origin and geography of the fishes of Central Mexico. En: Hocutt CH, Wiley EO, (eds). The Zoogeography of North American freshwater fishes. Wiley-Intersciences publication. pp. 487-519.

Mittelbach, G. G., Schemske, D. W., Cornell, H. V., Allen, A. P., Brown, J. M. & Bush, M. K. *et al*. (2007). Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. *Ecology Letters*. 10(4): 315-331.

Morris, A. B. & Shaw, J. (2018). Markers in time and space: A review of the last decade of plant phylogeographic approaches. *Molecular Ecology*. 27: 2317-2333.

Nelson, G. G. (1975). Anatomy of the male urogenital organs of *Goodea atripinnis* and *Characodon lateralis* (Atheriniformes: Cyprinodontoidei), and *G. atripinnis* courtship. *Copeia*. 3: 475-482.

Nixon, K. C. & Wheler, Q. D. (1990). An amplification of the phylogenetics species concept. *Cladistics*. 6: 211–223.

Ornelas-García, C. P., Alda, F., Díaz-Pardo, E., Gutiérrez-Hernández, A. & Doadrio, I. (2012). Genetic diversity shaped by historical and recent factors in the live-bearing twoline skiffia *Neotoca bilineata*. *Journal of Fish Biology*. 81(6): 1963-1984.

Ornelas-García, C. P., Domínguez-Domínguez, O. & Doadrio, I. (2008). Evolutionary history of the fish genus *Astyanax* Baird & Girard (1854) (Actinopterygii, Characidae) in Mesoamerica reveals multiple morphological homoplasies. *BMC Evololutionary Biology*. 8: 340.

Parenti, L. R. (1981). A phylogenetic and biogeographic analysis of cyprinodontiform fishes (Teleostei, Atherinomorpha). *Bulletin of the American Museum of Natural History*. 168(4): 335-557.

Pedraza-Lara, C., Doadrio, I., Breinholt, J. W. & Crandall, K. A. (2012). Phylogeny and Evolutionary Patterns in the Dwarf Crayfish Subfamily (Decapoda: Cambarellinae). *Plos One*. 7: e48233.

Pérez-Rodríguez, R., Domínguez-Domínguez, O., Doadrio, I., Cuevas- García, E., & Pérez-Ponce de León, G. (2015). Comparative historical biogeography of three groups of Neartic freshwater fishes across central México. *Journal of Fish Biology*. 86: 993–1015.

Pérez-Rodríguez, R., Domínguez-Domínguez, O., Pérez-Ponce de León, G. &Doadrio, I. (2009). Phylogenetic relationships and biogeography of the genus *Algansea* Girard 1856 (Cypriniformes:Cyprinidae) from central Mexico inferred from molecular data. *BMC Evolutionary Biology*. 9: 223.

Pérez-Rodríguez, R., Domínguez-Domínguez, O., Mar-Silva, F., Doadrio, I. & Pérez-Ponce de León, G. (2016). The historical biogeography of the southern group of the sucker genus *Moxostoma* (Teleostei: Catostomidae) and the colonization of Central México. *Zoological Journal of the Linnean Society*. 177: 633-647.

Piller, K. R., Kenway-Lynch, C. S., Camak, D. T. & Domínguez-Domínguez, O. (2015) Phylogeography and population structure of the imperilled redtail splitfin (Goodeidae: *Xenotoca eiseni*): Implications for Conservation. *Copeia*. 103: 440–454. doi: http://dx.doi.org/10.1643/CI-14-067

Prince, L. M. (2015). Plastid primers for angiosperm phylogenetics and phylogeography. *Applications in Plant Sciences*. 3: 1400085.

Rabosky, D. L. & McCune, A. R. (2010). Reinventing species selection with molecular phylogenies. *Trends in Ecology and Evolution*. 25(2): 68-74

Reznick, D. N. & Miles, D. B. (1989). Review of life history patterns in poeciliid fishes. En: Ecology and Evolutionof Livebearing Fishes (Poeciliidae). Meffe GK, Snelson JJ Jr. (eds). Prentice Hall, Englewood Cliffs, NJ. Pp 125-148.

Ritchie, M. G., Webb, S. A., Graves, J. A., Magurran, A. E. & Macías-García, C. (2005). Patterns of speciation in endemic Mexican Goodeid fish: sexual conflict or early radiation?. *Journal of Evolutionary Biology.* 18: 922-929.

Ritchie, M. G., Hamill, R. M., Graves, J. A., Magurran, A. E., Webb, S. A. & Macías-García, C. (2007). Sex and differentiation: population genetic divergence and sexual dimorphism in Mexican goodeid fish. *Journal of Evolutionary Biology*. 20: 2048-2055.

Rosas-Elguera, J., Ferrari, L., Lopez-Martinez, M. & Urrutia-Fucugauchi, J. (1997). Stratigraphy and tectonics of the Guadalajara region and the triple junction area, western Mexico. *International Geology Review*. 39: 125–140.

Rosas-Elguera, J. & Urrutia-Fucugauchi, J. (1998). Tectonic control on the volcano-sedimentary sequence of the Chapala graben, western Mexico. *International Geology Review*. 40(4): 350–362.

Rosen, D.E. (1975). A vicariance model of Caribbean biogeography. *Systematic Zoology*. 24: 431–464.

Sanger, F. & Coulson, A. R. (1975). A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *Journal of Molecular Biology*. 94(3): 441-448.

Seehausen, O. (2004). Hybridization and adaptive radiation. *Trends in Ecology and Evolution*. 19(4): 198-207.

Shaw, J., Shafer, H. L., Leonard, O. R., Kovach, M. J., Schorr, M. & Morris, A. B. (2014). Chloroplast DNA sequence utility for the lowest phylogenetic and phylogeographic inferences in angiosperms: The tortoise and the hare IV. *American Journal of Botany*. 101. 1987–2004.

Shubin, N., Tabin, C., & Carrol, S. (2009). Deep homology and the origins of evolutionary novelty. *Nature*. 457: 818-823.

Soto-Galera, E., Díaz-Pardo, E., López-López, E. & Lyons, J. (1999). Fish as indicators of environmental quality in the Río Lerma Basin, México. *Aquatic Ecosystems Health & Management*. 1: 267-276.

Stearns, S. C. & Hoekstra, R. F. (2000). Evolution: an introduction. Oxford University press, New York, USA. 381 pp.

Templeton, A. R. (2006). Population genetics and microevolutionary theory. John Wiley & Sons. New Jersey, USA. 720 pp.

Turner, B. J. & Grosse, D. J. (1980) Trophic differentiation in *Ilyodon,* a genus of stream dwelling Goodeid fishes: speciation versus ecological polymorphism. *Evolution*. 34: 259–270.

Uribe, M. C., De la Rosa-Cruz, G. & García-Alarcón, A. (2005). The ovary of viviparous teleost. Morphological differences between the ovaries of *Goodea atripinnis* and *Ilyodon whitei* (Goodeidae). En: *Viviparous Fishes*. Uribe, MC. Grier, HJ (eds), New Life Publications, Florida, pp 217-236.

Villa-Villaseñor, I. M. (2013). Variación intraespecífica en la elección de pareja en Xenotoca variata Bean 1887; (Cyprinodontiformes: Goodeidae) entre dos poblaciones de las cuencas de Cuitzeo y Zacapu. Tesis de Licenciatura. Universidad Michoacana de San Nicolás de Hidalgo, México. 62 pp.

Webb, S. A. (2002) Molecular systematic of the genus *Allodontichthys* (Cyprinodontiformes: Goodeidae). *Reviews in Fish Biology and Fisheries*. 12: 193–205.

Webb, S. A., Graves, J. A., Macías-García, C., Magurran, A. E., Foighil, D. O., & Ritchie, M.G. (2004). Molecular phylogeny of the livebearing Goodeidae (Cyprinodontiformes). *Molecular Phylogenetics and Evolution*. 30: 527–544.

Zimmer, E. A. & Wen, J. (2015). Using nuclear gene data for plant phylogenetics: Progress and prospects II. Next-gen approaches. *Journal of Systematics and Evolution*. 53: 371–379.

Zhang, A. B., Kubota, K., Takami, Y., Kim, J. L., Kim, J. K. & Sota, T. (2005). Species status and phylogeography of two closely related *Coptolabrus* species (Coleoptera: Carabidae) in South Korea inferred from mitochondrial and nuclear gene sequences. *Molecular Ecology*. 14: 3823–3841.