



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
POSGRADO EN CIENCIAS BIOLÓGICAS
INSTITUTO DE BIOLOGÍA
BIOLOGÍA EVOLUTIVA

Filogenómica de la especie-anillo *Euphorbia tithymaloides*
usando datos de RADseq

TESIS

QUE PARA OPTAR POR EL GRADO DE:
MAESTRO EN CIENCIAS

PRESENTA:

MAX DEMIAN MEDINA RODRÍGUEZ

TUTORA PRINCIPAL DE TESIS: DRA. NATALIA IVALÚ CACHO GONZÁLEZ
INSTITUTO DE BIOLOGÍA, UNAM

COMITÉ TUTOR: DRA. SOFÍA SOLÓRZANO LUJANO
FES IZTACALA, UNAM

DR. JUAN JOSÉ MORRONE
FACULTAD DE CIENCIAS, UNAM

TUTOR INVITADO: DR. CARLOS ALONSO MAYA LASTRA
COLUMBIA UNIVERSITY

CIUDAD UNIVERSITARIA, CD. MX., 2022.



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
POSGRADO EN CIENCIAS BIOLÓGICAS
INSTITUTO DE BIOLOGÍA
BIOLOGÍA EVOLUTIVA

Filogenómica de la especie-anillo *Euphorbia tithymaloides*
usando datos de RADseq

TESIS

QUE PARA OPTAR POR EL GRADO DE:
MAESTRO EN CIENCIAS

PRESENTA:

MAX DEMIAN MEDINA RODRÍGUEZ

TUTORA PRINCIPAL DE TESIS: DRA. NATALIA IVALÚ CACHO GONZÁLEZ
INSTITUTO DE BIOLOGÍA, UNAM

COMITÉ TUTOR: DRA. SOFÍA SOLÓRZANO LUJANO
FES IZTACALA, UNAM

DR. JUAN JOSÉ MORRONE
FACULTAD DE CIENCIAS, UNAM

TUTOR INVITADO: DR. CARLOS ALONSO MAYA LASTRA
COLUMBIA UNIVERSITY

CIUDAD UNIVERSITARIA, CD. MX., 2022.

COORDINACIÓN DEL POSGRADO EN CIENCIAS BIOLÓGICAS
INSTITUTO DE BIOLOGÍA
OFICIO CPCB/891/2022
ASUNTO: Oficio de Jurado

M. en C. Ivonne Ramírez Wence
Directora General de Administración Escolar, UNAM
P r e s e n t e

Me permito informar a usted que el Comité Académico del Posgrado en Ciencias Biológicas, celebrada el día **1° de agosto de 2022** se aprobó el siguiente jurado para el examen de grado de **MAESTRO EN CIENCIAS BIOLÓGICAS** en el campo de conocimiento de **BIOLOGÍA EVOLUTIVA** del estudiante **MEDINA RODRÍGUEZ MAX DEMIAN** con número de cuenta **416032330** con la tesis titulada **“Filogenómica de la especie-anillo *Euphorbia tithymaloides* usando datos de RADseq.”**, realizada bajo la dirección de la **DRA. NATALIA IVALÚ CACHO GONZÁLEZ**, quedando integrado de la siguiente manera:

Presidente: DR. DANIEL IGNACIO PIÑERO DALMAU
Vocal: DRA. ALICIA MASTRETTA YANES
Vocal: DRA. CAROLINA GRANADOS MENDOZA
Vocal: DR. JUAN PABLO JARAMILLO CORREA
Secretario: DRA. SOFÍA SOLORZANO LUJANO

Sin otro particular, me es grato enviarle un cordial saludo.

ATENTAMENTE
“POR MI RAZA HABLARÁ EL ESPÍRITU”
Ciudad Universitaria, Cd. Mx., a 26 de septiembre de 2022

COORDINADOR DEL PROGRAMA



DR. ADOLFO GERARDO NAVARRO SIGÜENZA



AGRADECIMIENTOS

Agradezco al Posgrado en Ciencias Biológicas de la Universidad Nacional Autónoma de México (UNAM) por brindarme la oportunidad para realizar mis estudios de Maestría.

Agradezco al Consejo Nacional de Ciencia y Tecnología (CONACyT) por el apoyo económico otorgado a través de la beca de Maestría (número de becario: 788747), con la que me fue posible dedicarme a mis estudios y continuar con mi crecimiento académico.

Al CONACYT por el financiamiento otorgado para el proyecto “Genómica y morfometría de *Euphorbia tithymaloides*, la única especie-anillo en plantas” (CB2015-00255829) a N. Ivalú Cacho.

Al Laboratorio de Genómica Funcional y Sistemas (Conacyt-INFR2016-268109) por proporcionar infraestructura para el trabajo de laboratorio.

A la Dra. Natalia Ivalú Cacho González, mi tutora principal, por todo el apoyo en mi formación académica.

A los miembros de mi Comité Tutor, la Dra. Sofía Solórzano Lujano, el Dr. Juan José Morrone Lupi y el miembro invitado el Dr. Carlos Alonso Maya Lastra, por la asesoría, la revisión del escrito y las sugerencias que enriquecieron el proyecto.

AGRADECIMIENTOS A TÍTULO PERSONAL

Agradezco a mi tutora principal, la Dra. Natalia Ivalú Cacho González, por darme la idea de titularme por estudios de posgrado, por siempre alentarme a continuar con el trabajo, por preocuparse por mí y tenerme suficiente paciencia durante todo el tiempo que duró la maestría hasta su conclusión.

Agradezco a los miembros de mi Comité Tutor, el Dr. Juan J. Morrone y la Dra. Sofía Solórzano por proporcionarme observaciones útiles, comentarios y sugerencias que me ayudaron a complementar y enriquecer mi trabajo.

Agradezco al miembro invitado de mi Comité Tutor, el Dr. Carlos Alonso Maya Lastra, por su completa disposición a ayudarme durante todo el proceso, por orientarme en la tarea de realizar los análisis y por darme consejos útiles e importantes.

Agradezco a los miembros de mi Jurado para el Examen de Grado: El Dr. Daniel Ignacio Piñero Dalmau, la Dra. Alicia Mastretta, la Dra. Sofía Solórzano Lujano, la Dra. Carolina Granados Mendoza y el Dr. Juan Pablo Jaramillo Correa.

Agradezco a los miembros del Cacho Lab, los otros estudiantes de mi tutora, Brenda Hernández, Itzel Piña de la Rosa, Clara López Marmolejo y Miguel Rosas, por brindarme su compañía, orientarme y ayudarme a realizar diversas actividades en la oficina y el laboratorio.

Agradezco a todos los profesores que me dieron clases durante los primeros tres semestres de la maestría.

Agradezco a mis papás, Francisca Rodríguez y César Gilberto Medina, por siempre ayudarme económicamente cuando lo necesité y permitirme vivir con ellos durante buena parte de la maestría, por preocuparse por mí, por tenerme paciencia durante los días de estrés, enojo, tristeza, y aguantar mis extraños ciclos de sueño.

Agradezco a las personas que me brindaron su compañía y que me permitieron dialogar y convivir con ellos durante muchas etapas de este proceso: Eduardo Sánchez, Nadia Ramón, Mich Montalvo y Julieta Vargas (Julieta Bosque).

Por último, agradezco a mi familia del alma, las personas que me han brindado su amistad y cariño incondicional durante muchos años, que siempre me ayudaron a encontrar la motivación y me dieron el apoyo emocional necesario que me permitió completar todo este proceso: Juan Pablo Mancera, Luis Fernando Valles, Alan Alfredo Nevárez (Fred Calavera), Bernardo Blanco (Bernie), Leonardo Rodríguez y Arath Sosa Gabaldón. Definitivamente de no ser por ellos no hubiera podido llegar tan lejos ni lograr nada de esto. Siempre les estaré eternamente agradecido por su valiosa amistad. Los quiero mucho, profundamente.

Índice

Resumen	1
Abstract.....	2
Introducción	3
Objetivos.....	6
<i>Objetivo general</i>	6
<i>Objetivos particulares</i>	6
Escenarios, hipótesis y resultados esperados	7
Métodos.....	9
<i>Muestras, extracción de ADN y secuenciación</i>	9
<i>Procesamiento de datos y construcción de matriz de SNPs</i>	10
<i>Visualización de datos de RADseq (PCA)</i>	12
<i>Enfoque filogenético</i>	13
<i>Análisis de flujo génico</i>	13
<i>Análisis de estructura genética</i>	14
Resultados.....	15
<i>Procesamiento de datos</i>	15
<i>Visualización de datos (PCA)</i>	15
<i>Enfoque filogenético</i>	15
<i>Análisis de estructura genética</i>	22
Discusión	27
Conclusiones	37
Referencias.....	38
Anexo I. Algunos conceptos, métodos y parámetros utilizados en este trabajo.....	47
<i>Especies-anillo y el concepto de especie</i>	47
<i>Filogenética, filogeografía y genética de poblaciones: enfoques y limitaciones</i>	48
<i>Filogenética</i>	48
<i>Genética de poblaciones</i>	49
<i>Filogeografía</i>	50
<i>Datos de RADseq</i>	51
<i>Parámetros importantes</i>	53
Anexo II. <i>Euphorbia tithymaloides</i>	55
<i>El Caribe y Euphorbia tithymaloides</i>	57
<i>Euphorbia tithymaloides en México</i>	58
Anexo III. Tabla detallada de los ensamblados realizados	59
Referencias.....	60

Índice de figuras

- Figura 1.** Distribución geográfica de los ejemplos mejor estudiados de especies-anillo. a) *Ensatina eschscholtzii* (tomado de Kuchta y Wake, 2006). b) *Phylloscopus trochiloides* (tomado de Wikimedia commons). 4
- Figura 2.** Representación gráfica de distintas hipótesis geográficas de colonización del Caribe por parte de *Euphorbia tithymaloides* desde su centro de origen en México/Guatemala y los resultados esperados correspondientes para análisis (filogenético y de estructura genética). A) Escenario de colonización a través de dos frentes. B) Escenario de colonización por un frente. C) Distintos eventos de colonización del continente a las islas. 8
- Figura 3.** Mapa que muestra los puntos de colecta de las muestras incluidas en este estudio. Clave de colores verde, accesiones continentales y la isla de Curazao; cian, muestras de las Antillas Menores (San Eustaquio, Guadalupe y Barbados); rojo las muestras de las Antillas Mayores (Bahamas, Jamaica, Puerto Rico, República Dominicana, St. Croix y St. John). La línea negra marca el Pasaje de Anegada. 10
- Figura 4.** Gráficas obtenidas del análisis de sensibilidad (u optimización) del parámetro `clust_threshold` (umbral de similitud) a partir de datos genómicos de *Euphorbia tithymaloides*. a) Número de RADloci obtenidos por muestra en función del valor del umbral de similitud. b) Número de SNPs totales en la matriz concatenada en función del valor del umbral de similitud. 15
- Figura 5.** Visualización de los resultados de análisis componentes principales (PCAs) realizados con datos genómicos de *Euphorbia tithymaloides*. Cada gráfica fue realizada con un ensamble diferente: a) `min50`, b) `min85`, c) `minsampleslocus02`, d) `min50_maxN10_maxH01`, e) `min50_maxN01_maxH01`, f) `min85_maxN01_maxH01`. Verde: Chiapas, Quintana Roo, Guatemala, Tamaulipas, San Luis Potosí y Oaxaca; verde claro: Costa Rica, Colombia, Venezuela y Curazao; cian: Antillas menores; rojo: Antillas Mayores. 17
- Figura 6a.** Árbol filogenético de *Euphorbia tithymaloides* abarcando su distribución por México, Centroamérica, el norte de Sudamérica y las islas del Caribe estimado a partir del ensamble `min50` (`min_samples_locus = 50`) analizado con RAxML (soporte bootstrap promedio de todos los clados = 72.22). Verde: Chiapas, Quintana Roo, Guatemala, Tamaulipas, San Luis Potosí y Oaxaca; verde claro: Costa Rica, Colombia, Venezuela y Curazao; cian: Antillas menores; rojo: Antillas Mayores. Este árbol representa un ejemplo de topología de 1 origen de colonización. 18
- Figura 6b.** Árbol filogenético de *Euphorbia tithymaloides* abarcando su distribución por México, Centroamérica, el norte de Sudamérica y las islas del Caribe estimado a partir del ensamble `min85` (`min_samples_locus = 85`) analizado con RAxML (soporte bootstrap promedio de todos los clados = 66.43). Verde: Chiapas, Quintana Roo, Guatemala, Tamaulipas, San Luis Potosí y Oaxaca; verde claro: Costa Rica, Colombia, Venezuela y Curazao; cian: Antillas menores; rojo: Antillas Mayores. Este árbol representa un ejemplo de topología de 2 orígenes de colonización. 19
- Figura 7.** Árbol filogenético de datos genómicos de *Euphorbia tithymaloides* obtenido con el ensamble `alldefault + outgroup` y el programa RAxML. En negro: el grupo externo conformado por los datos genómicos de especies emparentadas a *E. tithymaloides* obtenidos a partir de digestiones *in silico*. En verde: Chiapas, Quintana Roo, Guatemala, Tamaulipas, San Luis Potosí y Oaxaca; verde claro: Costa Rica, Colombia, Venezuela y Curazao; cian: Antillas Menores; rojo: Antillas Mayores. 21
- Figura 8.** Resultados de los análisis de Treemix de datos genómicos de *Euphorbia tithymaloides* al calcular dos eventos de flujo génico ($m = 2$): a) gráfica resultado de los ensambles `min50` y `min85`, b) `min50_maxN01_maxH01` y `min50_maxN10_maxH10`, c) `min85_maxN01_maxH01`, y d) `min_sample_locus02`. Las flechas indican eventos de flujo génico, su color solo es para diferenciar los dos principales eventos calculados. Ramas coloreadas de rojo y rectángulos rojos indican las muestras de las Antillas Mayores; ramas coloreadas en cian y rectángulos rojos denotan muestras de las Antillas Menores. 23

- Figura 9.** Resultados de los análisis de STRUCTURE realizados con datos genómicos de *Euphorbia tithymaloides*, asumiendo 2-5 grupos genéticos y usando el ensamble con el parámetro min_samples_locus02. 24
- Figura 10.** La prueba de Evanno sugiere que el valor más adecuado de grupos genéticos en *Euphorbia tithymaloides* es $K = 6$. En azul, valores de delta K 25
- Figura 11.** Resultados de los análisis de STRUCTURE realizados con datos genómicos de *Euphorbia tithymaloides*, asumiendo 6 grupos genéticos y usando el ensamble con el parámetro min_samples_locus02. La línea negra indica el Pasaje de Anegada. 25
- Figura 12.** Representaciones gráficas de las dos hipótesis soportadas por los dos tipos de topología resultantes de los distintos ensambles y sus implicaciones biogeográficas sobre la colonización del Caribe por *Euphorbia tithymaloides*. Las líneas discontinuas representan eventos de colonización hipotéticos que se asumen a partir de las topologías. a) escenario que implica dos frentes de colonización, pues las Antillas Mayores y Menores no aparecen como grupos hermanos y éstas últimas aparecen anidadas en un clado que tiene su origen en Centro y Sudamérica, b) escenario en el que las Antillas Mayores y Menores aparecen como grupos hermanos formando un clado, implicando un único evento de colonización del Caribe. 30
- Figura 13.** a) Árbol con disposición anidada de cuatro taxones (E (T1 (T2, T3))) en el que se basa el análisis ABBA-BABA. b) Escenario similar a la topología 1-origen y sus configuraciones ABBA y BABA. c) Escenario similar a la topología 2-orígenes y sus configuraciones ABBA y BABA. 34
- Figura 14.** Proceso de identificación y tipificación de marcadores RAD (ADN asociado a enzimas de restricción). Las muestras de ADN genómico S1 y S2 tienen la secuencia de reconocimiento de varias enzimas de restricción en distintos lugares a lo largo del genoma. Los triángulos azules representan los sitios de restricción de una enzima particular. Algunos de estos sitios de restricción no están presentes en una muestra debido a polimorfismos que causan diferencias en las secuencias (asteriscos rojos). Las dos muestras son digeridas con una enzima de restricción particular y luego ligadas a los enlazadores biotinilados (elipses azules). El ADN es cortado al azar dejando sólo los fragmentos flanqueantes al sitio de restricción ligados a los enlazadores de biotina, los cuales luego son purificados (Modificado de Miller et al., 2007). 52
- Figura 15.** Fotografía de un individuo de la especie *Euphorbia tithymaloides* con diversos ciatios, cortesía de la Dra. N. Ivalú Cacho. 55
- Figura 16.** Distribución geográfica de *Euphorbia tithymaloides* (los puntos representan localidades; tomado de Dressler, 1957). 56

Índice de tablas

Tabla 1. Número de individuos de <i>Euphorbia tithymaloides</i> muestreados por país o por Estado de la República Mexicana.	9
Tabla 2. Ensamblados realizados a partir de los datos genómicos de <i>Euphorbia tithymaloides</i> y sus respectivos valores de los cuatro parámetros modificados en este estudio. Una tabla más detallada se presenta en el Anexo III.	11
Tabla 3. Reconstrucciones filogenéticas obtenidas a partir de los ensamblados considerados como más adecuados (<i>alldefault</i> , <i>min_samples_locus02</i> y <i>alldefault + outgroup</i>), descripción de cada una de las reconstrucciones e hipótesis mejor soportada.	20
Tabla 4. Resumen de los resultados de los distintos análisis en este trabajo y sus implicaciones principales, que se abordan en la discusión.	26
Tabla 5. Resumen de la evidencia a favor y en contra de las hipótesis alternativas	36
Tabla 6. Ensamblados con los valores usados de los cuatro parámetros modificados, el número de loci resultante, el número de SNPs resultante, el porcentaje de sitios faltantes (<i>missing sites</i>) en la matriz de SNPs y la topología resultante a partir de análisis RAxML.	59

Resumen

La Cuenca del Caribe es una región geográfica altamente compleja que alberga una gran variedad de ecosistemas, y se le considera uno de los *hotspots* de diversidad biológica más importantes del mundo. *Euphorbia tithymaloides* L. es una planta que posee una distribución circular alrededor del Caribe, que comprende a México, Florida, Centroamérica, el norte de Sudamérica y la mayor parte de las islas del Caribe. En estudios anteriores se ha sugerido que *E. tithymaloides* pudo haber tenido un origen en México-Centroamérica para luego aumentar su distribución geográfica a partir de dos frentes: uno hacia el noreste culminando en las Antillas Mayores, y otro hacia el sureste pasando por Sudamérica y culminando en las Antillas Menores. Además, se ha sugerido que las poblaciones terminales de cada uno de los dos frentes presentan poco flujo génico directo entre ellas, a pesar de estar conectadas entre sí por una cadena de poblaciones a lo largo de toda la distribución. Estas características coinciden con el modelo de especie-anillo, que describe a una especie que aumenta su distribución geográfica alrededor de un área no habitable formando un anillo, de manera que las poblaciones terminales de cada extremo de la distribución presentan poco o nulo flujo génico a pesar de coexistir, comportándose como especies diferentes. Los estudios anteriores se basaron en análisis de dos genes nucleares, en este estudio se amplía el muestreo genético a un nivel genómico utilizando datos generados con tecnologías de secuenciación de alto rendimiento (RADseq). Por medio de este tipo de datos se realizaron análisis de flujo génico, filogenéticos, de agrupamiento (PCA) y de estructura genética con el propósito de profundizar nuestro entendimiento de la biogeografía histórica y patrones de divergencia de *Euphorbia tithymaloides*. Se utilizaron 95 muestras provenientes de 35 localidades distribuidas a lo largo de la mayor parte del área de distribución de *E. tithymaloides* en los países que conforman la cuenca del Caribe. A partir de los resultados se deduce un soporte parcial de la hipótesis filogeográfica de dos frentes de expansión, aunque no se descarta la hipótesis alternativa de una ruta de colonización más compleja.

Abstract

The Caribbean Basin is a highly complex geographic region that harbors a great diversity of ecosystems, and is considered one of the world's most important hotspots of biological diversity. *Euphorbia tithymaloides* L. is a plant that exhibits a circular geographic range around the Caribbean, which covers part of Mexico, Florida, Central America, the north of South America and most of the Caribbean islands. In previous studies it has been suggested that *E. tithymaloides* could've originated in Mexico-Central America and would later expand its geographic range along two fronts: one northeastward into the Greater Antilles, and another one southeastward passing through South America and into the Lesser Antilles. Additionally, it has been suggested that the terminal populations of each one of the fronts display very limited direct gene flow between them, despite being connected by a chain of populations throughout the whole geographic range. These characteristics coincide with the ring-species model, that describes a species that expands its geographic range around a non-habitable area forming a ring-like distribution. The terminal populations of each of the two ends of this distribution display little to no gene flow despite coexisting, thus acting like different species. While previous studies were based on analyses performed using two nuclear genes, the present study expands the genetic sampling to a genomic level using high-throughput sequencing data (RADseq). Multiple analyses were performed using this type of data: phylogenetic analyses, principal component analyses (PCA), genetic structure analyses and gene flow tests. These analyses were executed in order to deepen our understanding of the historical biogeography and divergence patterns of *Euphorbia tithymaloides*. A total of 95 samples from 35 localities distributed throughout most of the geographic range of *E. tithymaloides* were used. The results of the present study partially support a scenario of geographic expansion along two fronts, although an alternative hypothesis of a more complex pattern of colonization cannot be completely discarded.

Introducción

“Las formas que poseen en grado considerable el carácter de especies, pero que se parecen mucho a otras formas o están tan estrechamente vinculadas con ellas por graduaciones intermedias que los naturalistas no gustan de clasificarlas como especies distintas, son en varios sentidos las más importantes para nosotros.” -Darwin (1859)

La mayoría de las veces el proceso de divergencia que ocurre durante los eventos de especiación es de naturaleza continua (Dobzhansky 1940, De Queiroz 2005, Moritz y Bi 2011, Nosil 2012). Dicha continuidad puede ser estudiada a partir de diversos enfoques, como patrones en la variación clinal (Endler 1977), el grado de separación o segregación de linajes [*lineage sorting*] (Funk y Omland 2003) o el grado de aislamiento reproductivo (Craddock 1987, Castillo 2017, Stankowski y Ravinet 2021). Los enfoques utilizados para estudiar los procesos de especiación pueden variar según el grado de avance de dicho proceso (Tregenza 2002, Svensson et al. 2009). Incluso se han documentado casos en los que la divergencia gradual ocurre aún con presencia de flujo génico entre las poblaciones (Smith et al. 1997, Niemiller et al. 2008, Papadopulos et al. 2011, Martin et al. 2013).

Para de Queiroz (1998; 2005) el proceso de especiación o divergencia de linajes es continuo, y dentro de éste se pueden distinguir varios criterios de especiación relacionados a distintos conceptos de especie. Más recientemente, esta visión del proceso de especiación se ha relacionado con el término “continuo de especiación” [*speciation continuum*], el cual fue definido por primera vez por Shaw y colaboradores (2014) como “la secuencia continua de cambios genéticos que ocurren mientras dos linajes divergen el uno del otro en dirección al aislamiento reproductivo”. Para estos autores, la evolución divergente y el aislamiento reproductivo son los elementos principales del proceso de especiación. Stankowski y colaboradores (2021) definen al continuo de especiación como un continuo de aislamiento reproductivo. Bajo esta definición, la distancia entre los taxa dentro del continuo de especiación es proporcional a la fuerza del aislamiento reproductivo que hay entre ellos.

Las especies-anillo son sistemas que adquieren una distribución geográfica circular (similar a un anillo) resultado de una expansión histórica durante la cual el sistema rodea áreas no habitables. En las especies-anillo, las poblaciones terminales de cada uno de ambos frentes de expansión geográfica son las más divergentes del sistema, incluso llegando a comportarse como especies diferentes; estas poblaciones eventualmente se reencuentran en su trayectoria alrededor del hábitat no disponible, y pueden coexistir sin que su diferenciación colapse a causa de flujo génico (Irwin et al. 2001a, Fuchs et al. 2015, de Brito Martins y de Aguiar 2017). Algunos autores han planteado que las especies-anillo deben cumplir con cuatro características (Irwin et al. 2001b, Coyne y Orr 2004, Cacho y Baum 2012, Kuchta y Wake 2016): 1) una distribución circular que surgió por la expansión geográfica histórica a través de dos frentes; 2) coexistencia de las dos formas terminales; 3) una cadena de poblaciones que permite el flujo génico entre las dos poblaciones terminales, y; 4) suficiente aislamiento reproductivo para impedir que las dos formas terminales colapsen por hibridación.

Mayr (1940) se refirió a las especies-anillo como la “demostración perfecta de especiación”, porque permiten estudiar aspectos clave sobre el proceso de divergencia evolutiva. Entre otros

aspectos, estos sistemas permiten rastrear en el espacio la historia de la formación de las especies a través de la variación geográfica que conecta las formas aisladas reproductivamente. Para Mayr (1942), la primera fase de la especiación consiste en la subdivisión poblacional, y la segunda en el aislamiento reproductivo. Por otro lado, para Ayala y colaboradores (1974), la especiación alopátrica consiste en una primera fase de diferenciación genética (causada por una separación geográfica) y una segunda fase en la que las poblaciones entran en contacto. En los sistemas ideales de especie-anillo todas estas etapas o fases se pueden observar simultáneamente, por lo que estos sistemas se consideran importantes en biología evolutiva. Además, los sistemas de especie-anillo permiten evaluar si la diferenciación geográfica en la interfaz población-especie puede ocurrir en presencia de flujo génico (Irwin et al. 2001b, 2005).

Algunos ejemplos clásicos de especies-anillo son la salamandra *Ensatina eschscholtzii* que habita las zonas montañosas que rodean al Valle Central de California (Figura 1a, Kuchta y Wake 2016), y el ave *Phylloscopus trochiloides*, que habita las montañas del Altiplano Tibetano (Figura 1b, Irwin et al. 2001b). Estos sistemas que muestran una distribución circular adquirida por una divergencia histórica a través de dos frentes geográficos y su posterior reencuentro ha sido respaldado por estudios a nivel genómico en los últimos años (Alcaide et al. 2014, Irwin et al. 2016).

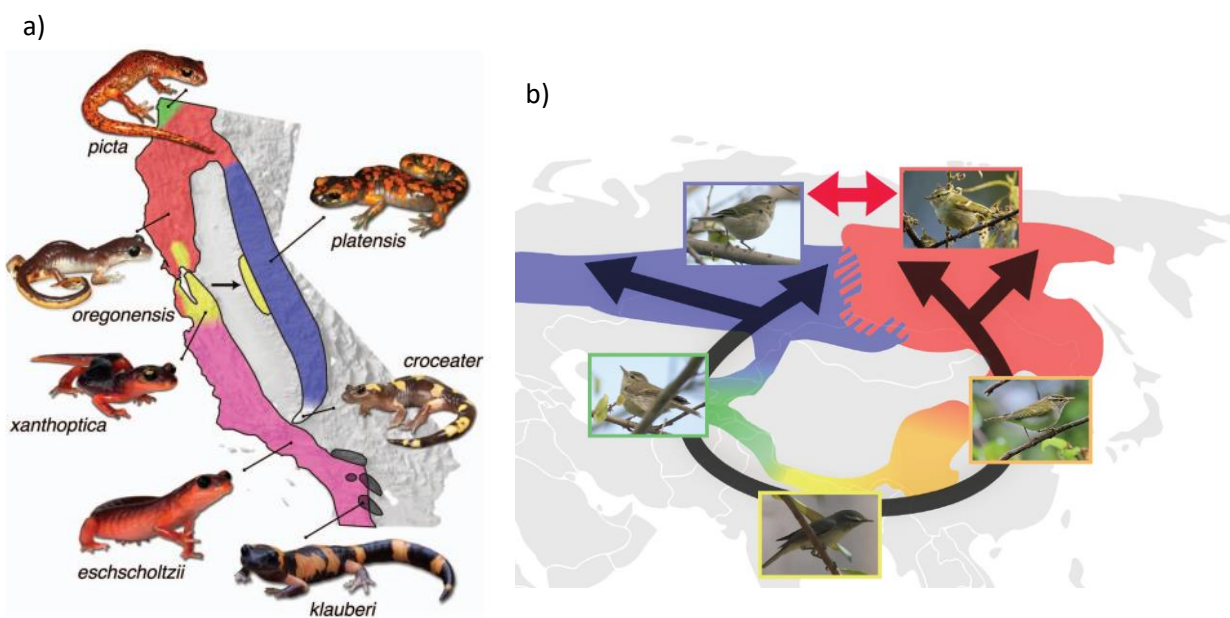


Figura 1. Distribución geográfica de los ejemplos mejor estudiados de especies-anillo. a) *Ensatina eschscholtzii* (tomado de Kuchta y Wake, 2006). b) *Phylloscopus trochiloides* (tomado de Wikimedia commons).

El proceso de divergencia tan particular que muestran los sistemas de especies-anillo puede ser difícil de representar desde un punto de vista taxonómico: muchas veces se han considerado como una especie con varias subespecies, o incluso como sistemas que consisten de dos o más especies diferentes (Irwin et al. 2001b).

El objeto de estudio de este proyecto es la planta *Euphorbia tithymaloides* L., la especie más variable y de distribución geográfica más amplia del clado *Pedilanthus* del género *Euphorbia*. Su área

de distribución abarca México, Florida, Centroamérica, el norte de Sudamérica y la mayoría de las islas del Caribe. Para sistematizar la variación morfológica en esta especie, Dressler (1957) sugirió seis subespecies en ambientes insulares en el Caribe (*E. tithymaloides* ssp. *angustifolia*, *E. t.* ssp. *bahamensis*, *E. t.* ssp. *jamaicensis*, *E. t.* ssp. *padifolia*, *E. t.* ssp. *parasítica* y *E. t.* ssp. *smalli*) y una forma continental (*E. tithymaloides* ssp. *tithymaloides*). Las diferencias morfológicas entre las subespecies son sutiles y hay intergradación entre ellas, al grado en que la distribución geográfica es un carácter diagnóstico para algunas subespecies (p.ej., *E. tithymaloides* ssp. *jamaicensis*; Cacho et al., 2012; Dressler, 1957). Información más detallada sobre este sistema se presenta en el Anexo II.

Un estudio filogenético de 42 poblaciones de *E. tithymaloides* del Caribe basado en secuencias de dos regiones nucleares de copia única (el Unigene SGN-U342009 del Sol Genomics Network (SGN) y el locus *GAPC-2* de la subunidad C de la gliceraldehído-3-fosfato deshidrogenasa) mostró que esta especie expandió su distribución geográfica mediante dos frentes, cada uno representado por un clado de poblaciones insulares hermano a distintas poblaciones del continente, de forma consistente con un patrón de especie-anillo (Cacho y Baum 2012). Dicho estudio ubica el área de origen de *E. tithymaloides* en México/Guatemala, y los dos frentes de expansión hacia el Caribe a través de las Antillas Menores y las Antillas Mayores, los cuales se reencuentran en el área de las Islas Vírgenes. Adicionalmente, los autores sugieren que el Pasaje de Anegada, caracterizado por una gran distancia entre islas (150 km) y corrientes de fuertes a moderadas, podría estar funcionando como una barrera al flujo génico entre las formas terminales de este sistema. Es decir, el escenario biogeográfico y de intergradación fenotípica observados en *E. tithymaloides* corresponden con el tipo de progresión que se ha observado en varias especies-anillo, y que se ha utilizado para fundamentar dicho modelo (Irwin 2000, Soltani et al. 2017). Una limitación del estudio anterior es su muestreo genético: sólo se secuenciaron dos regiones nucleares de copia única, lo cual podría limitar la capacidad de inferencia de los análisis realizados. En este trabajo se evaluará la biogeografía histórica y patrones de divergencia de la especie *Euphorbia tithymaloides* ampliando la perspectiva a una escala genómica.

Objetivos

Objetivo general

Evaluar, a partir de una perspectiva genómica, a la especie *Euphorbia tithymaloides* con el propósito de identificar los procesos y factores biogeográficos y evolutivos que han moldeado la divergencia poblacional de la especie en su distribución por la cuenca del Caribe.

Objetivos particulares

1. Estimar la variación genética en las secuencias de RADseq provenientes de un muestreo geográfico representativo del área de distribución de la especie *Euphorbia tithymaloides* en el Caribe, para revelar si existen patrones geográficos en dicha variación.
2. Examinar la hipótesis de que *Euphorbia tithymaloides* colonizó el Caribe a través de dos frentes geográficos, que es consistente con el fenómeno de especie-anillo, usando un enfoque filogenético y datos de RADseq.
3. Caracterizar la estructura genética de *Euphorbia tithymaloides* con base en datos de RADseq utilizando un enfoque de filogeografía y genética de poblaciones, para complementar el enfoque filogenético con elementos que puedan ofrecer información acerca de los procesos que pueden estar subyaciendo los patrones observados.

Escenarios, hipótesis y resultados esperados

A. ESCENARIO DE DOS FRENTES DE COLONIZACIÓN: *Euphorbia tithymaloides* expandió su área de distribución hacia el Caribe a través de dos frentes geográficos que se reencuentran, consistente con el modelo de especie-anillo (Cacho et al., 2012; Figura 2A).

ENFOQUE FILOGENÉTICO: Si la expansión de *E. tithymaloides* por el Caribe se dio a través de dos frentes geográficos, un enfoque filogenético mostrará un árbol con dos clados bien diferenciados que corresponden a los dos frentes geográficos de la especie. Cada clado incluirá poblaciones insulares, sucesivamente anidadas (en una secuencia que refleja el avance de dicho frente) y será hermano a poblaciones del centro de origen (México y Guatemala, Cacho et al., 2012).

ENFOQUE GENÉTICA DE POBLACIONES: En análisis de agrupamiento [*clustering*] las poblaciones terminales de ambos frentes forman grupos bien diferenciados y no deben mostrar señales de hibridación o de pertenencia a un mismo grupo.

B. ESCENARIO DE UN SOLO FRENTE DE COLONIZACIÓN: *Euphorbia tithymaloides* colonizó el Caribe por medio de una 'progresión geográfica', partiendo de un solo punto del continente. Es decir, la colonización siguió un patrón 'lineal' desde el continente pasando el norte de Sudamérica, las Antillas Menores y las Antillas Mayores (o el caso contrario, una colonización pasando por México y Florida, las Antillas Mayores y luego las Antillas Menores; Figura 2B).

ENFOQUE FILOGENÉTICO: La filogenia mostrará dos clados principales, uno correspondiente a las accesiones continentales y otro a las insulares. El clado insular estará anidado en el clado continental, y mostrará un patrón sucesivamente anidado en el que en las ramas más externas estarán representadas las poblaciones continentales, y en las más internas las poblaciones insulares, mostrando así la colonización lineal sucesiva de las islas.

ENFOQUE GENÉTICA DE POBLACIONES: Los métodos de agrupamiento darán como resultado una serie lineal sucesiva de individuos entre las poblaciones más diferenciadas.

C. ESCENARIO DE EVENTOS DE COLONIZACIÓN INDEPENDIENTES: *Euphorbia tithymaloides* colonizó el Caribe, no por medio de una 'progresión geográfica', sino a través de eventos independientes (Figura 2C).

ENFOQUE FILOGENÉTICO: El árbol filogenético mostrará una estructura en la que distintas poblaciones insulares son hermanas a distintas poblaciones continentales, sin un patrón geográfico claro. No será aparente en la filogenia una resolución más allá de varios clados conformados por poblaciones insulares y poblaciones continentales que representan sus parientes más cercanos.

ENFOQUE GENÉTICA DE POBLACIONES: Los análisis de agrupamiento mostrarán grupos independientes formados por las poblaciones insulares y ciertas poblaciones continentales. Es decir, casi no habrá agrupamientos determinados por la cercanía geográfica.

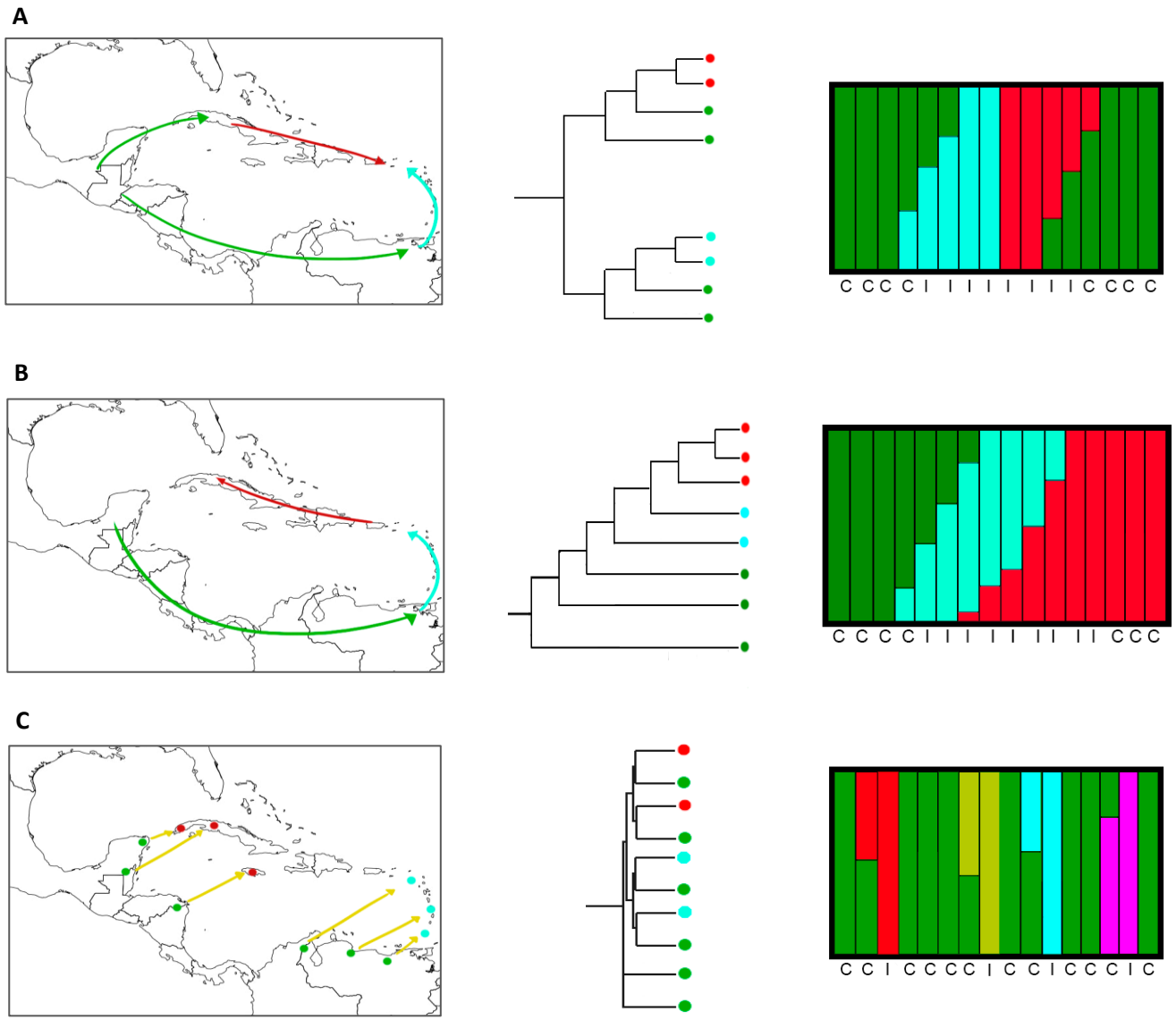


Figura 2. Representación gráfica de distintas hipótesis geográficas de colonización del Caribe por parte de *Euphorbia tithymaloides* desde su centro de origen en México/Guatemala y los resultados esperados correspondientes para análisis (filogenético y de estructura genética). A) Escenario de colonización a través de dos frentes. B) Escenario de colonización por un frente. C) Distintos eventos de colonización del continente a las islas.

Métodos

Muestras, extracción de ADN y secuenciación

Se utilizó material genético de 95 individuos provenientes de 35 localidades (con 1-6 individuos por localidad), que abarcan México, Centroamérica, el norte de Sudamérica, las Antillas Mayores y las Antillas Menores, y que colectivamente cubren la mayor parte del Caribe (Tabla 1). Se realizó un mapa con los puntos de colecta (Figura 3) utilizando R v4.0.3 (R Core Team 2020) con los paquetes MAPS, RGDAL y RASTER (Hijmans y van Etten 2012, Bivand et al. 2016).

Tabla 1. Número de individuos de *Euphorbia tithymaloides* muestreados por país o por Estado de la República Mexicana.

País (región)	Muestras
Bahamas	1
Barbados	3
Colombia	5
Costa Rica	4
Curazao	4
Guadeloupe	7
Guatemala	9
Jamaica	3
México (Chiapas)	3
México (Oaxaca)	11
México (Quintana Roo)	17
México (San Luis Potosí)	2
México (Tamaulipas)	1
Puerto Rico	6
República Dominicana	5
San Eustaquio	6
Islas Vírgenes St. Croix	3
Islas Vírgenes St. John	4
Venezuela	1
Suma total	95

Para la extracción de ADN se utilizó una colección de tejido seco en sílica gel. Las extracciones de ADN se realizaron con el *DNeasy Plant Mini Kit* (Qiagen, USA) y las cuantificaciones se realizaron con métodos fluorimétricos (*Qbit*). Las bibliotecas y la generación de datos de RADseq se realizaron en *Florgenex* (www.florgenex.com), de acuerdo con el protocolo descrito en Baird et al. (2008) y usando la enzima de restricción *SbfI* (5'-CCTGCAGG-3'). Se generaron lecturas *single-end* de 100 bp

(1x100 pb) y la secuenciación fue realizada en el HiSeq 4000 en las instalaciones GC3F de la Universidad de Oregon, en Estados Unidos.

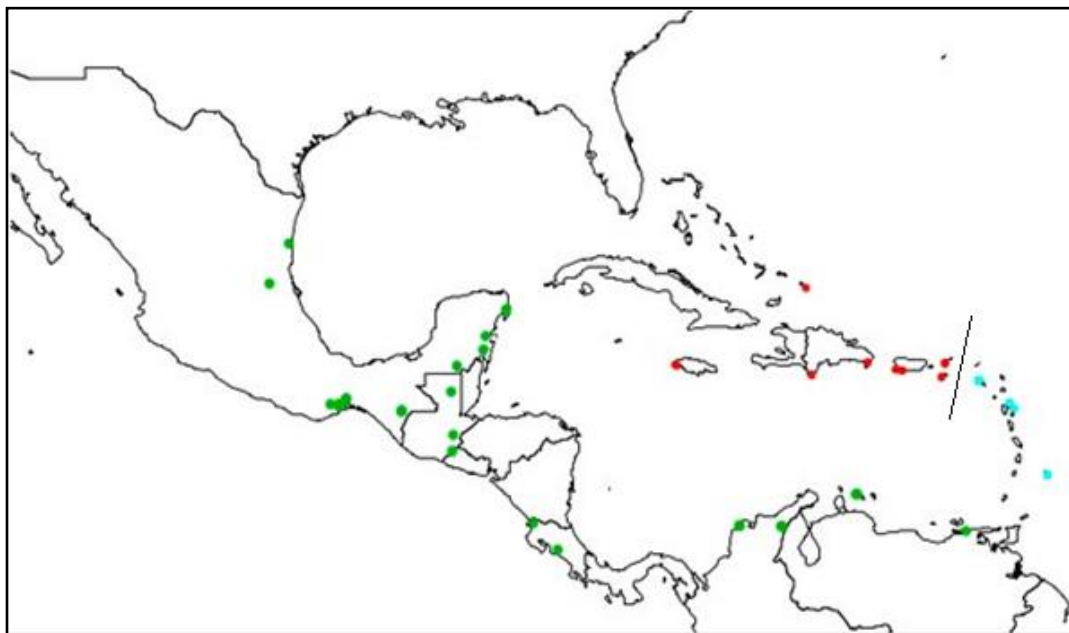


Figura 3. Mapa de los puntos de colecta de las muestras de *Euphorbia tithymaloides* utilizadas en este estudio. Clave de colores: verde, accesiones continentales y la isla de Curazao; cian, muestras de las Antillas Menores (San Eustaquio, Guadalupe y Barbados); rojo las muestras de las Antillas Mayores (Bahamas, Jamaica, Puerto Rico, República Dominicana, St. Croix y St. John). La línea negra marca el Pasaje de Anegada.

Procesamiento de datos y construcción de matriz de SNPs

El procesamiento de las secuencias y varios de los análisis posteriores se realizaron con el programa *ipyrad* 0.9 (Eaton y Overcast 2020) en una interfaz de programación de aplicaciones (API) por medio de JupyterLab 3.4.2 (Corlay et al. 2016). Primero, se realizó un ensamble *de novo* siguiendo la *pipeline* del programa, que consiste en: demultiplexar [*demultiplex*] los datos a partir de los códigos de barras, filtrarlos, editarlos y agruparlos en lecturas; se estimó la heterocigosidad y la tasa de error de cada una de las muestras, se filtraron y se puso formato para análisis subsecuentes a los archivos de salida. Se obtuvieron una matriz de SNPs y una matriz de loci, las cuales fueron los archivos de entrada para los análisis posteriores.

Para construir el ensamble más adecuado para estos datos, se evaluaron distintos valores (0.83 a 0.96) del umbral de similitud (*clust_threshold*; figuras 5a y 5b) y se eligió el valor que maximizara el número de SNPs y de loci recuperados. Este criterio es consistente con la propuesta de Mastretta-Yanes et al. (2015), la cual también fue probada por McCartney-Melstad et al. (2019). También se llevó a cabo una exploración con el *min_samples_locus* realizando distintos ensambles con diferentes valores asignados a dicho parámetro (2, 4, 12, 25, 50, 75, 85, 95; *default* = 4) y realizando los árboles filogenéticos para evaluar aquellos que tenían mayor sentido biogeográfico y mayores valores de *bootstrap*.

Además, se realizaron exploraciones variando los valores de los parámetros *max_Ns_consens* y *max_Hs_consens*, evaluando valores de 0.01 a 0.2 (teniendo como valor *default* 0.05). Todos los ensamblajes producidos y los respectivos valores de los parámetros se muestran en la Tabla 2.

Tabla 2. Ensamblajes realizados a partir de los datos genómicos de *Euphorbia tithymaloides* y sus respectivos valores de los cuatro parámetros modificados en este estudio. Una tabla más detallada se presenta en el Anexo III.

Nombre del ensamble	<i>clust_threshold</i>	<i>min_samples_locus</i>	<i>max_Ns_consens</i>	<i>max_Ns_consens</i>
<i>alldefault</i>	0.85	4	0.05	0.05
<i>min_sample_locus 02</i>	0.85	2	0.05	0.05
<i>euphorbia</i>	0.88	4	0.05	0.05
<i>min12</i>	0.88	12	0.05	0.05
<i>min25</i>	0.88	25	0.05	0.05
<i>min50</i>	0.88	50	0.05	0.05
<i>min75</i>	0.88	75	0.05	0.05
<i>min85</i>	0.88	85	0.05	0.05
<i>min95</i>	0.88	95	0.05	0.05
<i>min50_maxN01_maxH01</i>	0.88	50	0.01	0.01
<i>min50_maxN01_maxH10</i>	0.88	50	0.01	0.10
<i>min50_maxN10_maxH01</i>	0.88	50	0.10	0.01
<i>min50_maxN10_maxH10</i>	0.88	50	0.10	0.10
<i>min85_maxN01_maxH01</i>	0.88	85	0.01	0.01
<i>min85_maxN01_maxH10</i>	0.88	85	0.01	0.10
<i>min85_maxN10_maxH01</i>	0.88	85	0.10	0.01
<i>min85_maxN10_maxH10</i>	0.88	85	0.10	0.10
<i>min85_maxN20_maxH20</i>	0.88	85	0.20	0.20
<i>maxH01</i>	0.85	4	0.05	0.01
<i>maxH10</i>	0.85	4	0.05	0.10
<i>maxH25</i>	0.85	4	0.05	0.25
<i>maxH50</i>	0.85	4	0.05	0.50
<i>maxH75</i>	0.85	4	0.05	0.75

Los ensamblajes *min_samples_locus02* y *alldefault* fueron considerados como los más adecuados para llevar a cabo todos los análisis filogenéticos y de estructura genética. Esto fue por varias razones, en primer lugar, el ensamblaje *min_samples_locus02* es el menos “restrictivo” al momento de definir las relaciones entre las accesiones, ya que acepta que para cada locus deben existir por lo menos dos muestras que lo compartan. De esta manera se toma en consideración la mayor cantidad de loci posibles, lo cual es útil al momento de reconstruir las relaciones entre individuos y poblaciones a nivel infraespecífico porque se conserva la mayor cantidad de variación.

Por otro lado, el ensamble *alldefault* es el que posee todos los valores de los parámetros predeterminados, y es similar al ensamble *min_samples_locus02* excepto por el hecho de que el parámetro *min_samples_locus* posee su valor predeterminado de 4, que significa que para cada locus debe haber por lo menos 4 muestras que lo compartan. Ambos ensambles son muy similares y retienen mucha variación en comparación con el resto de los ensambles realizados. Aunque estos ensambles se consideraron los más adecuados, aun así se siguieron utilizando los otros ensambles para realizar los diversos análisis (excepto el análisis de STRUCTURE) con el propósito de comparar resultados (o de ejemplificar los tipos de topología encontrados; figuras 6a y 6b).

Con el propósito de investigar si los datos podrían reflejar herencia maternal (mitocondria y plástidos) o biparental (núcleo) se realizó una serie de digestiones *in silico* a partir de genomas y plastomas (cloroplasto y mitocondria) publicados de distintas especies de la familia Euphorbiaceae (*Euphorbia esula*, *Manihot esculenta*, *Jatropha curcas*, *Ricinus communis* y *Hevea brasiliensis*) utilizando la enzima de restricción *SbfI* (5'-CCTGCAGG-3'). De esta exploración se concluyó que la gran mayoría parte de los sitios de corte para la enzima utilizada (*SbfI*) se encuentran en el núcleo. El número más alto de sitios de corte en un plastoma fue de 7, en la mitocondria de *Manihot esculenta*, mientras que el número de sitios de corte en el genoma de las distintas especies varió de 1,647 a 16,126. Además de explorar el lugar de los sitios de corte, las digestiones *in silico* de *Ricinus communis*, *Hevea brasiliensis*, *Manihot esculenta* y *Euphorbia esula* se utilizaron para complementar el ensamble *alldefault* (que resultó en el ensamble *alldefault + outgroup*) con el propósito de incorporar un grupo externo a las reconstrucciones filogenéticas.

Visualización de datos de RADseq (PCA)

El análisis de componentes principales (PCA) es un método estadístico en el cual una gran cantidad de mediciones pueden ser reducidas a unos cuantos componentes principales (PCs) que explican los principales patrones de variación. En un análisis exitoso el primer componente principal del análisis es la combinación de mediciones que explica la mayor cantidad de variación en los datos (Pearson 1901, Reich et al. 2008). Este análisis puede ser llevado a cabo con datos genéticos para inferir estructura poblacional y variación genética (Novembre y Stephens 2008, Reich et al. 2008). Se realizó un análisis de componentes principales usando el toolkit de *ipyrad* (Eaton y Overcast 2020) en la interfaz de programación de aplicaciones (API) por medio de JupyterLab (Corlay et al. 2016) y se visualizó utilizando el programa Toyplot (<https://toyplot.readthedocs.io/en/stable/>). Los análisis de PCA se llevaron a cabo utilizando los datos de SNPs de la variedad de ensambles realizados con distinto valor de los parámetros *min_samples_locus*, *max_Ns_consens* y *max_Hs_consens*. Se agruparon las muestras a partir de localidades asignadas según al país que pertenecen. Al parámetro *mincov* (que filtra los SNPs que son compartidos en menos de una cierta proporción de muestras) se le asignó un valor de 0.5, lo cual quiere decir que se filtraron los SNPs que no eran compartidos en al menos la mitad de las muestras. Se decidió no utilizar algoritmos de imputación, es decir, no se reemplazaron los datos faltantes con algún valor consenso calculado a partir de los datos disponibles.

Enfoque filogenético

Para inferir las relaciones de ancestro-descendencia entre los individuos, se estimaron árboles filogenéticos con el método de máxima verosimilitud usando los diferentes ensambles y los programas RAxML (Stamatakis 2014) y IQTree (Nguyen et al. 2015). Además, para los ensambles *min_samples_locus02* y *alldefault* se utilizó el programa ASTRAL-III v 5.7.1 (Zhang et al. 2020), con el propósito de crear reconstrucciones filogenéticas a partir de colecciones de árboles de genes individuales. Resumir árboles de genes puede ayudar a resolver mejor las discordancias causadas por la separación incompleta de linajes o *incomplete lineage sorting* (Zhang et al. 2020). Para estos dos ensambles, además, se obtuvieron secuencias consenso por localidad para crear reconstrucciones filogenéticas con un solo representante por cada localidad.

Para realizar los análisis de RAxML se utilizó el modelo GTRGAMMA (GTR + optimización de tasa de sustitución + modelo GAMMA de heterogeneidad) y 100 réplicas de *bootstrap* (algoritmo de análisis rápido). Para los análisis con IQTree se escogieron 1000 réplicas de *bootstrap*, con valores predeterminados en el resto de los parámetros. Para las reconstrucciones filogenéticas en las que se utilizó un solo representante por cada localidad, se empleó la herramienta de análisis *window_extractor* dentro de ipyrad para obtener las secuencias consenso. El formato de los archivos de salida de estos análisis fue procesado con Toytree (Eaton y Overcast 2020) y Toyplot (Shead 2014) usando JupyterLab (Corlay et al. 2016) y las filogenias fueron visualizadas con FigTree v1.4.4 (<http://tree.bio.ed.ac.uk/software/figtree/>).

Para realizar los análisis de ASTRAL primero se utilizó la herramienta de análisis treeslider dentro de ipyrad para producir un árbol filogenético para cada uno de los RADloci. Como método de inferencia de los árboles se escogió “raxml” y se usaron 100 repeticiones de *bootstrap*. Posteriormente se llevó a cabo el análisis ASTRAL realizando un “resumen” de los árboles de los RADloci individuales. En el parámetro de anotaciones se escogió una probabilidad posterior local, y el resto de los parámetros se dejó por defecto.

Con base en estudios que sugieren que el origen de expansión de la distribución geográfica de *E. tithymaloides* fue la región de México y Centroamérica (Cacho y Baum 2012), el enraizamiento de los árboles filogenéticos se realizó con muestras del sur de México (p. ej. Quintana Roo, Chiapas) o de Guatemala. Esta interpretación del origen de la colonización coincide con la hipótesis de que las especies del subgénero al que pertenece *Euphorbia tithymaloides* evolucionaron en México (Dressler 1957, Dressler y Sacamano 1992, Lomelí-Sención 2006). Adicionalmente se realizaron análisis filogenéticos con RAxML y IQTree utilizando el ensamble que aporta grupos externos representados por especies de la familia Euphorbiaceae (*alldefault* + *outgroup*).

Análisis de flujo génico

Para inferir patrones de divergencia y de flujo génico entre individuos de distintas localidades a partir de las matrices de SNPs se utilizó el programa Treemix (Pickrell y Pritchard 2012). Este programa se basa en el hecho de que no siempre se puede representar la historia poblacional por medio de un árbol bifurcante, a causa de que puede existir flujo génico entre las distintas poblaciones que conforman una especie. Por lo tanto, el programa primero infiere un árbol bifurcante de poblaciones

a partir del método de máxima verosimilitud. Posteriormente, a partir de las poblaciones que se adecuan pobremente al árbol modelo generado, se infiere la existencia de flujo génico (o eventos de migración) entre poblaciones. Estos análisis se llevaron a cabo con varios de los ensambles realizados (p. ej. *min50*, *min85*, *min50_maxN01_maxH01*, *min_samples_locus02*, etc.) y dejando el valor predeterminado de 2 al parámetro *m* que indica el número de eventos de migración (o en este caso, flujo génico). A pesar de que no existe un número “correcto” de eventos de migración en un set de datos, se puede calcular el valor más adecuado de este parámetro por medio de pruebas estadísticas (Fitak 2021). Sin embargo, para este trabajo se decidió usar el valor predeterminado de 2 por considerarse suficiente para encontrar evidencia de flujo génico entre los dos grupos de interés: el grupo de las Antillas Mayores y el de las Antillas Menores.

Análisis de estructura genética

Para analizar la estructura genética del conjunto de muestras, se realizó un análisis de “agrupamiento de poblaciones con mezcla” usando el programa STRUCTURE v2.3.4 (Pickrell y Pritchard 2012). Se probaron los números de *K* de 2-10 para considerar distintos escenarios biogeográficos [poblaciones insulares vs. continentales (*K*=2); poblaciones continentales, de las Antillas Mayores y Antillas Menores (*K*=3); etc.]. Se utilizó una longitud de periodo de *burnin* de 500,000 y un número de repeticiones MCMC post-*burnin* de 750,000. El valor de *K* más adecuado para nuestro set de datos se escogió con base en una prueba de Evanno (Evanno et al. 2005), y los resultados se graficaron con Toyplot (toyplot.readthedocs.io). El ensamble utilizado para el análisis de STRUCTURE fue el *min_samples_locus02*, debido a que este es el menos “restrictivo” ya que para cada locus deben existir por lo menos dos muestras que lo compartan.

Resultados

Procesamiento de datos

El análisis de sensibilidad para evaluar cómo varía el número de RADloci y de SNPs en función del parámetro *clust_threshold* muestra que los valores óptimos de este parámetro son entre 0.89 y 0.91 para RADloci (figura 4a), y entre 0.86 y 0.88 para SNPs (figura 4b).

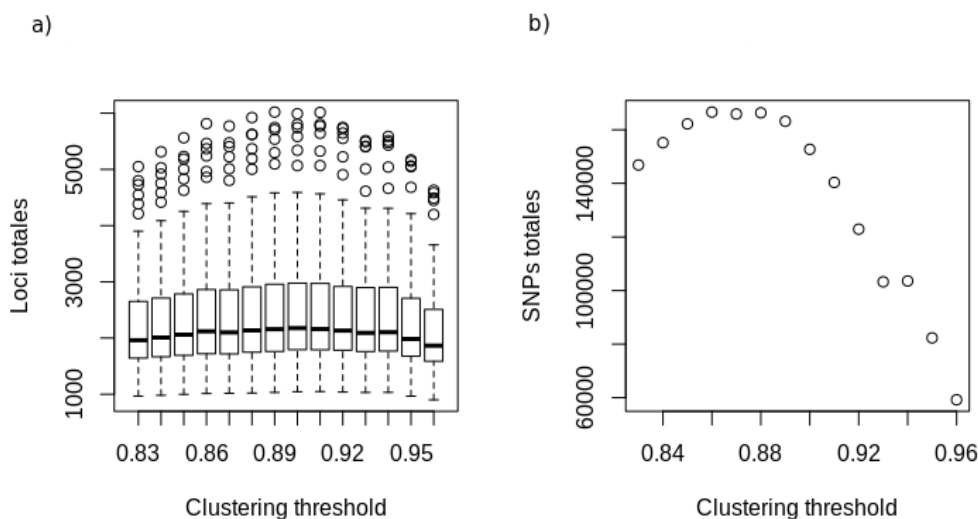


Figura 4. Gráficas obtenidas del análisis de sensibilidad (u optimización) del parámetro *clust_threshold* (umbral de similitud) a partir de datos genómicos de *Euphorbia tithymaloides*. a) Número de RADloci obtenidos por muestra en función del valor del umbral de similitud. b) Número de SNPs totales en la matriz concatenada en función del valor del umbral de similitud.

Visualización de datos (PCA)

Se realizó un análisis de componentes principales por cada ensamble, y por cada análisis se retuvieron los primeros dos componentes principales. A lo largo de todos los ensambles, la varianza explicada del primer componente principal (PC0) tuvo un rango de 13.6-29.8%, mientras que la varianza explicada del segundo componente principal (PC1) tuvo un rango de 10.9-23.7%. El número de SNPs totales utilizados variaron en un rango de 251 (por ejem. en el ensamble *min85_maxN01_maxH01*) a 4684 (por ejem. en el ensamble *min50*). A lo largo de los PCAs obtenidos con todos los ensambles, los resultados fueron bastante similares y consistentes. Para la mayor parte de los ensambles las muestras se agrupan de manera definida según su área geográfica (continente, Antillas Mayores y Antillas Menores). Las gráficas revelan algunas muestras que se ubican en el espacio intermedio entre la agrupación del continente y alguno de los dos grupos de las islas (como las correspondientes a Costa Rica). En todos los ensambles las muestras de las Antillas Mayores parecen agruparse de manera más definida y separada del resto de las muestras (figura 5).

Enfoque filogenético

A partir de los análisis filogenéticos llevados a cabo con todos los ensambles (Tabla 2) se obtuvieron siempre dos topologías contrastantes con respecto a sus clados principales. La primera de éstas fue denominada *1-origen* (se muestra un ejemplo de ésta en la Figura 6a, obtenida a partir del ensamble *min50* analizado con RAxML) que muestra 6 clados principales diferenciados por la distribución

geográfica de sus accesiones: i) Guatemala y Quintana Roo; ii) Guatemala y Chiapas; iii) San Luis Potosí, Oaxaca y Tamaulipas; iv) Costa Rica, Colombia, Venezuela y Curazao; v) las Antillas Mayores; y vi) las Antillas Menores. Esta topología se distingue por el hecho de que el clado de las muestras provenientes de Antillas Mayores y el de las provenientes de las Antillas Menores aparecen como grupos hermanos. Por otra parte, se obtuvo la topología *2-origenes* (un ejemplo de ésta se muestra en la Figura 6b, obtenida a partir del ensamble *min85* analizado con RAxML), en la que se recuperan 5 clados principales: i) Guatemala y Quintana Roo; ii) Guatemala y Chiapas; iii) San Luis Potosí, Oaxaca y Tamaulipas; iv) las Antillas Mayores; y v) las Antillas Menores, Costa Rica, Colombia, Venezuela y Curazao. Esta topología se distingue por el clado que agrupa a las accesiones de Costa Rica, el norte de Sudamérica y las Antillas Menores en una disposición anidada que parece indicar una direccionalidad en la colonización desde el continente a las islas, aunque los valores de soporte de este clado siempre fueron relativamente bajos (entre 22-38). Ambas topologías implican dos patrones distintos de colonización de las islas (2 posibles eventos de colonización para *2-origenes* y 1 único evento para *1-origen*).

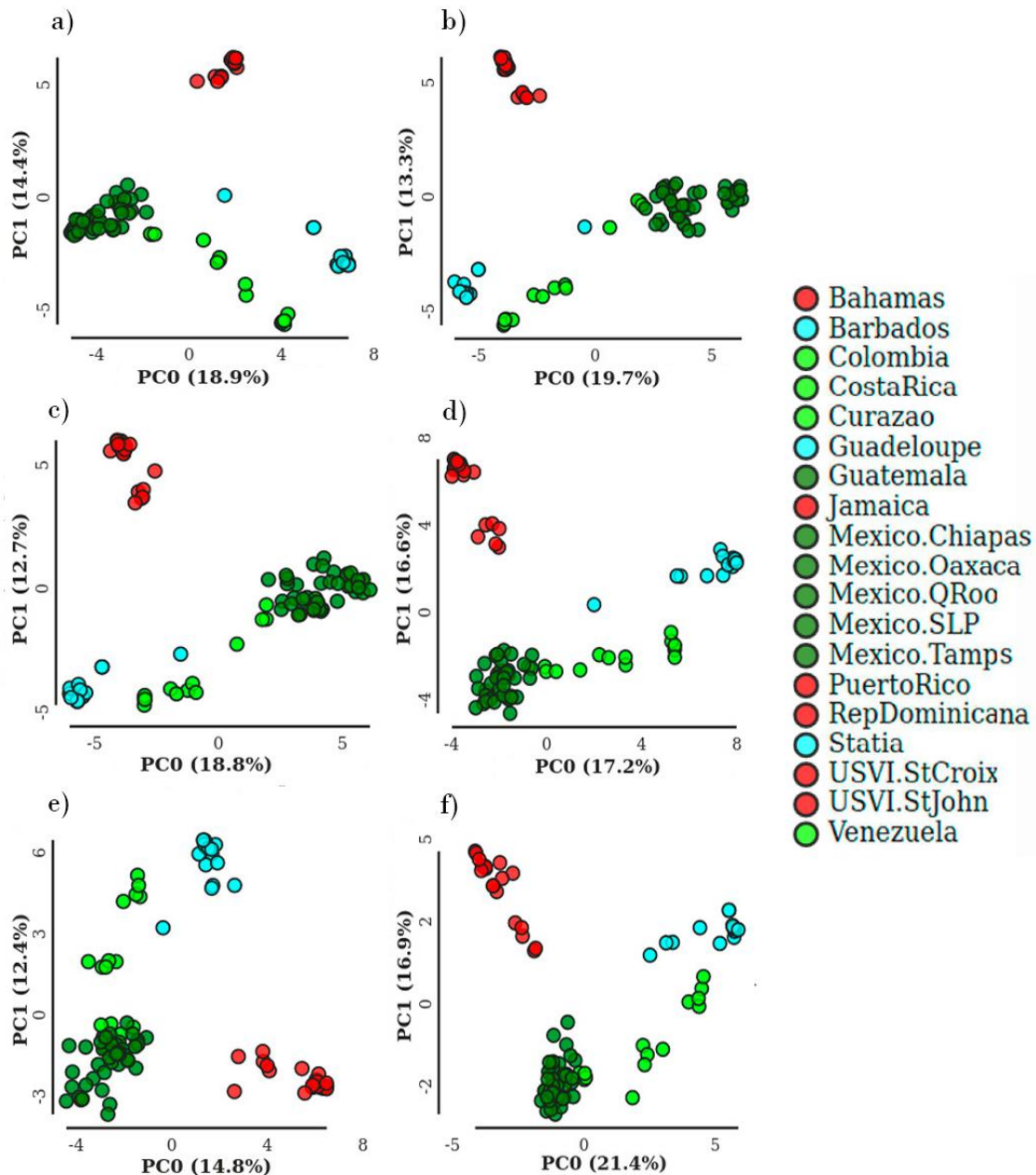


Figura 5. Visualización de los resultados de análisis componentes principales (PCAs) realizados con datos genómicos de *Euphorbia tithymaloides*. Cada gráfica fue realizada con un ensamble diferente: a) min50, b) min85, c) minsampleslocus02, d) min50_maxN10_maxH01, e) min50_maxN01_maxH01, f) min85_maxN01_maxH01. Verde: Chiapas, Quintana Roo, Guatemala, Tamaulipas, San Luis Potosí y Oaxaca; verde claro: Costa Rica, Colombia, Venezuela y Curazao; cian: Antillas menores; rojo: Antillas Mayores.

Los resultados de las reconstrucciones filogenéticas realizadas a partir de los programas RAXML, IQTree y ASTRAL para los ensambles que se consideraron más adecuados (*alldefault* y *min_samples_locus02*) se muestran en la tabla 3 junto con la hipótesis mejor soportada por cada una de las reconstrucciones. Asimismo, en la tabla 3 y figura 7 se muestran los resultados de las reconstrucciones obtenidas con el ensamble *alldefault + outgroup*.

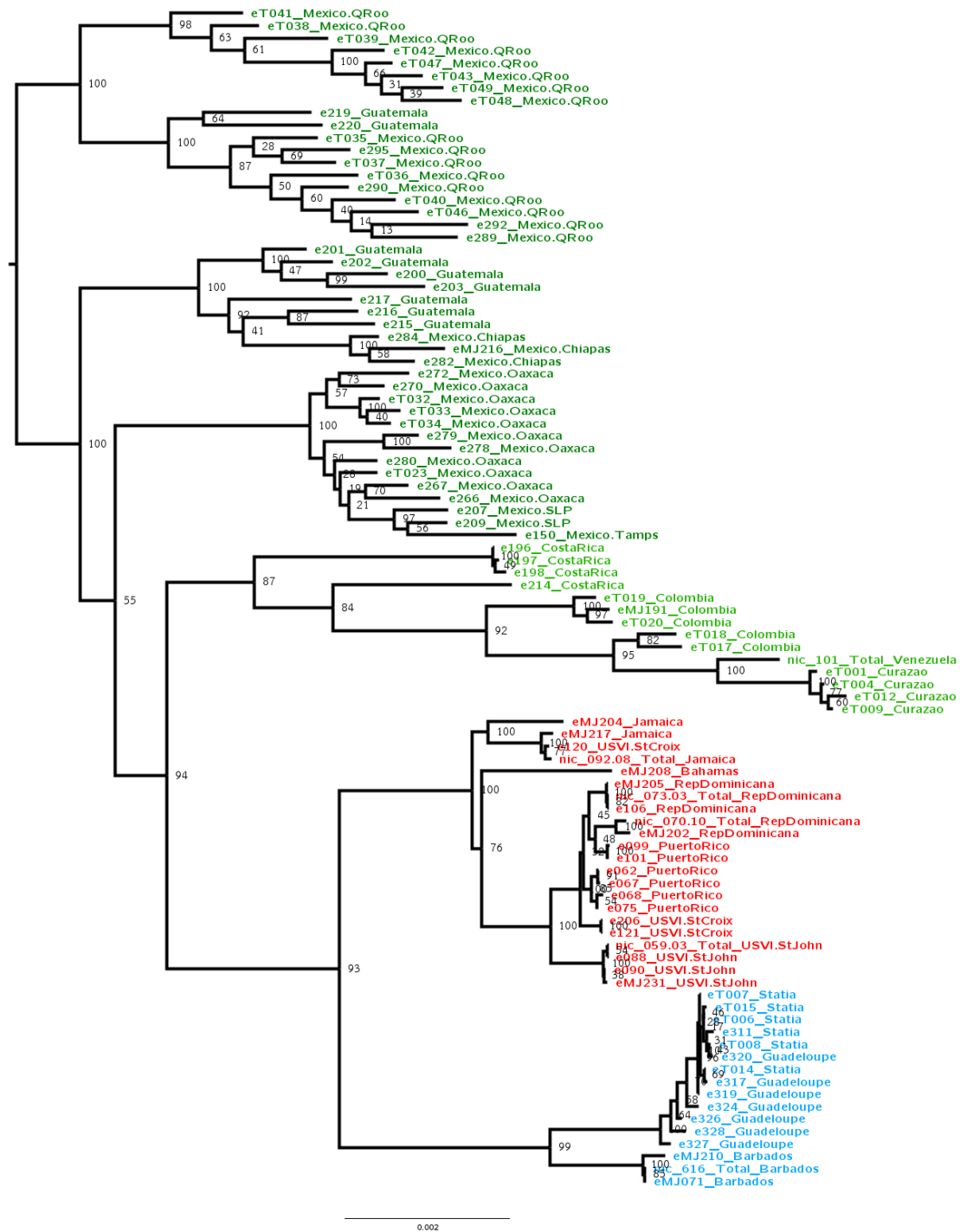


Figura 6a. Árbol filogenético de *Euphorbia tithymaloides* abarcando su distribución por México, Centroamérica, el norte de Sudamérica y las islas del Caribe estimado a partir del ensamble min50 (min_samples_locus = 50) analizado con RAxML (soporte bootstrap promedio de todos los clados = 72.22). Verde: Chiapas, Quintana Roo, Guatemala, Tamaulipas, San Luis Potosí y Oaxaca; verde claro: Costa Rica, Colombia, Venezuela y Curazao; cian: Antillas menores; rojo: Antillas Mayores. Este árbol representa un ejemplo de topología de 1 origen de colonización.

Tabla 3. Reconstrucciones filogenéticas obtenidas a partir de los ensamblajes considerados como más adecuados (*alldefault*, *min_samples_locus02* y *alldefault + outgroup*), descripción de cada una de las reconstrucciones e hipótesis mejor soportada.

Ensamblaje	Reconstrucción filogenética	Descripción	Hipótesis soportada
<i>min_sample_locus02</i>	min_sample_locus02	Ensamble con min_sample_locus = 2 analizado con IQTree	1-origen
	raxml_min_sample_locus02	Ensamble con min_sample_locus = 2 analizado con RAXML	1-origen
	IQtree_consensus_poblacion_iv_min2	Ensamble con min_sample_locus = 2, 1 representante de cada localidad (consenso de todas las muestras de dicha localidad), analizado con IQTree	2-origenes
	raxml_consensus_poblacion_iv_min2	Ensamble con min_sample_locus = 2, 1 representante de cada localidad (consenso de todas las muestras de dicha localidad), analizado con RAXML	2-origenes
	astral_consensus_min2	Ensamble con min_sample_locus = 2, 1 representante de cada localidad (consenso de todas las muestras de dicha localidad), analizado con ASTRAL	2-origenes
<i>alldefault</i>	secondIQ	Ensamble con todos los parámetros <i>default</i> analizado con IQTree	2-origenes
	raxml	Ensamble con todos los parámetros <i>default</i> analizado con RAXML	1-origen
	second_astral_consensus	Ensamble con todos los parámetros <i>default</i> , 1 representante de cada localidad (consenso de todas las muestras de dicha localidad), analizado con ASTRAL	2-origenes
	consensoIQ	Ensamble con todos los parámetros <i>default</i> , 1 representante de cada localidad (consenso de todas las muestras de dicha localidad), analizado con IQTree	2-origenes
	consensus_poblacion_iv	Ensamble con todos los parámetros <i>default</i> , 1 representante de cada localidad (consenso de todas las muestras de dicha localidad), analizado con RAXML	2-origenes
<i>alldefault + outgroup</i>	big_outgroup	Digestión <i>in silico</i> de todos los genomas disponibles de Euphorbiaceae a nivel de cromosoma + <i>alldefault</i> y reconst. en IQtree	2-origenes
	big_outgroup_raxml	Digestión <i>in silico</i> de todos los genomas disponibles de Euphorbiaceae a nivel de cromosoma + <i>alldefault</i> y reconst. en RAXML	2-origenes

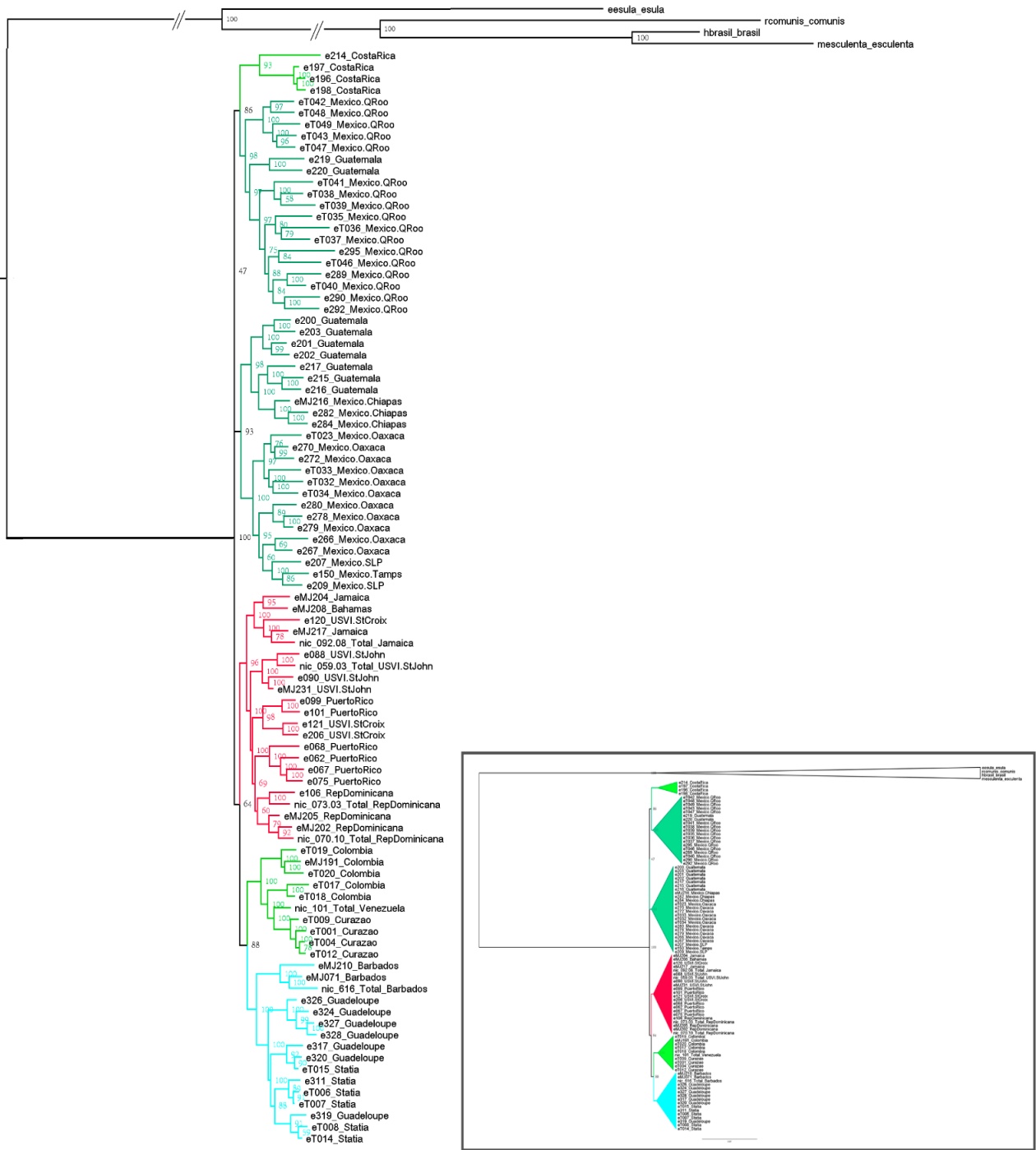


Figura 7. Árbol filogenético de datos genómicos de *Euphorbia tithymaloides* obtenido con el ensamble *alldefault* + *outgroup* y el programa RAxML. En negro: el grupo externo conformado por los datos genómicos de especies emparentadas a *E. tithymaloides* obtenidos a partir de digestiones *in silico*. En verde: Chiapas, Quintana Roo, Guatemala, Tamaulipas, San Luis Potosí y Oaxaca; verde claro: Costa Rica, Colombia, Venezuela y Curazao; cian: Antillas Menores; rojo: Antillas Mayores.

Para los análisis de Treemix se obtuvo un cladograma por cada ensamble, sin embargo, para algunos ensambles las topologías resultantes fueron las mismas (por ejemplo para min50 y min85). En la figura 8 se presentan cuatro ejemplos de topologías obtenidas a partir de una variedad de ensambles distintos (min50, min85, min50_maxN01_maxH01, min50_maxN10_maxH10, min85_maxN01_maxH01 y min_sample_locus02), las flechas de color indican eventos de “migración” o flujo génico. La mayor parte de las topologías resultantes coinciden con la hipótesis de dos eventos de colonización, algunas presentando inconsistencias como las muestras de Costa Rica apareciendo muy poco emparentadas una de otra (por ejem. figura 8a). Para los ensambles min50 y min85 (figura 8a) se obtuvo un resultado que sugiere flujo génico entre los individuos de las dos localidades muestreadas de Costa Rica y entre las muestras de la localidad costarricense de La Hacha y un clado compuesto por individuos de México y Guatemala. Para los ensambles min50_maxN01_maxH01 y min50_maxN10_maxH10, los resultados sugieren flujo génico entre el ancestro del subclado Costa Rica-Sudamérica-Antillas Menores y poblaciones de Guatemala (figura 8c, flecha naranja) y entre poblaciones de México (Quintana Roo) y un clado que contiene al grupo anteriormente mencionado (Costa Rica-Sudamérica-Antillas Menores; figura 9c, flecha verde). Por otra parte, los resultados obtenidos a partir de diversos ensambles (por ejem. min85_maxN01_maxH01 y min_sample_locus02) sugieren la existencia de flujo génico entre los individuos de las Antillas Menores y las Antillas Mayores (figuras 8b y 8d).

Análisis de estructura genética

De entre los análisis de STRUCTURE realizados con $K = 2-8$, los más significativos fueron los resultantes de $K = 2$ (Figura 9a) y $K = 3$ (Figura 9b) por la clara correspondencia biogeográfica de los grupos genéticos o *clusters* estimados, y el resultante de $K = 6$ (Figura 11) por ser 6 el valor de grupos genéticos considerado más “adecuado” a partir de la prueba de Evanno (Figura 10).

El resultado obtenido de $K = 2$ (Figura 9a) muestra un grupo genético correspondiente a los individuos de México, Guatemala y Costa Rica (en verde), y otro grupo correspondiente a los individuos del norte de Sudamérica y las Antillas Menores (en cian). En estos resultados, los individuos de las Antillas Mayores comparten casi en proporciones iguales alelos de ambos grupos. Esto sugiere que uno de los grupos genéticos podría estar relacionado a los individuos del continente, y el otro grupo relacionado a los individuos isleños o con distribución cercana a las poblaciones de las islas (como las del norte de Sudamérica).

En el resultado obtenido de $K = 3$ (Figura 9b) se distinguen grupos genéticos que corresponden a los individuos de las Antillas Mayores, las Antillas Menores y el continente. A partir de este punto, al aumentar el número de K únicamente se multiplican los grupos genéticos de las muestras del continente, pero la división en tres grupos principales (A. Mayores, A. Menores y el continente) se mantiene.

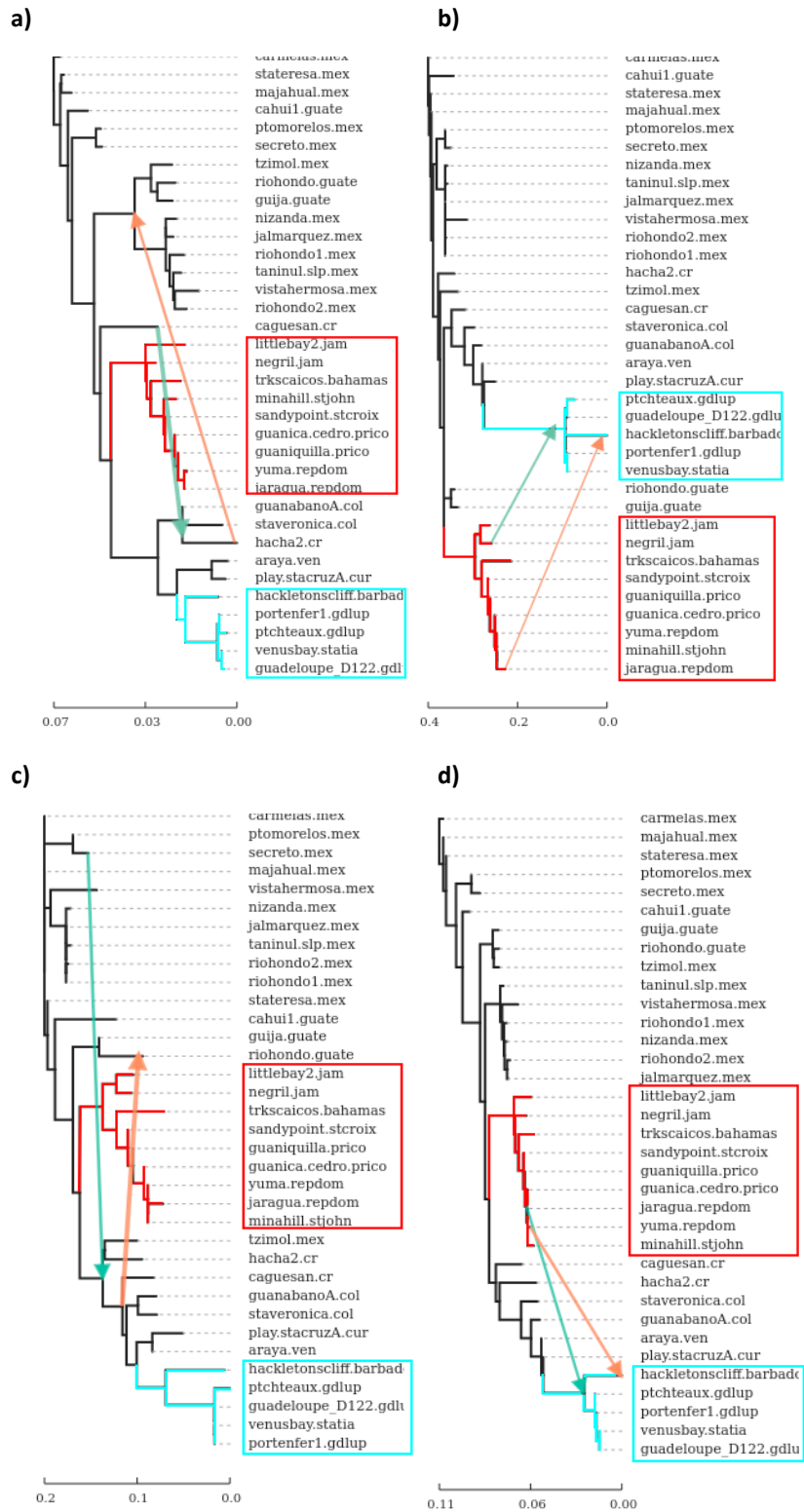


Figura 8. Resultados de los análisis de Treemix de datos genómicos de *Euphorbia tithymaloides* al calcular dos eventos de flujo génico ($m = 2$): a) gráfica resultado de los ensambles *min50* y *min85*, b) *min50_maxN01_maxH01* y *min50_maxN10_maxH10*, c) *min85_maxN01_maxH01*, y d) *min_sample_locus02*. Las flechas indican eventos de flujo génico, su color solo es para diferenciar los dos principales eventos calculados. Ramas coloreadas de rojo y rectángulos rojos indican las muestras de las Antillas Mayores; ramas coloreadas en cian y rectángulos rojos denotan muestras de las Antillas Menores.

En el caso de la división en 6 grupos genéticos (Figura 11), las muestras que más presentaron pertenencia parcial a varios grupos genéticos [*clusters*], fueron las que provienen de Chiapas, Guatemala, Costa Rica, Colombia, Venezuela y las Antillas Menores, lo cual podría deberse a la gradación continua de frecuencias alélicas en esta región, implicando 'entremezclamiento' genético, o *admixture*. Por otro lado, las muestras que menos presentan pertenencia parcial a distintos grupos genéticos, sino que presentan pertenencia total o casi total a únicamente un grupo genético, fueron las que pertenecen a las poblaciones de Quintana Roo, Oaxaca, Tamaulipas, San Luis Potosí y las Antillas Mayores.

La compartición de alelos entre las Antillas Mayores y las Antillas Menores parece ocurrir, pero de forma unidireccional, pues todas las muestras de las Antillas Menores presentan una pertenencia parcial (aunque muy baja) al único grupo genético al que pertenecen las muestras de las Antillas Mayores (Figura 9b, 9c, 9d y figura 11). Sin embargo, las muestras de las Antillas Mayores parecen compartir muy pocos alelos con el resto de las muestras.

La tabla 4 presenta un resumen de los escenarios derivados de los análisis de este trabajo. En esta tabla se puede observar que los resultados y dichos escenarios presentan algunas discordancias, así como elementos comunes importantes.

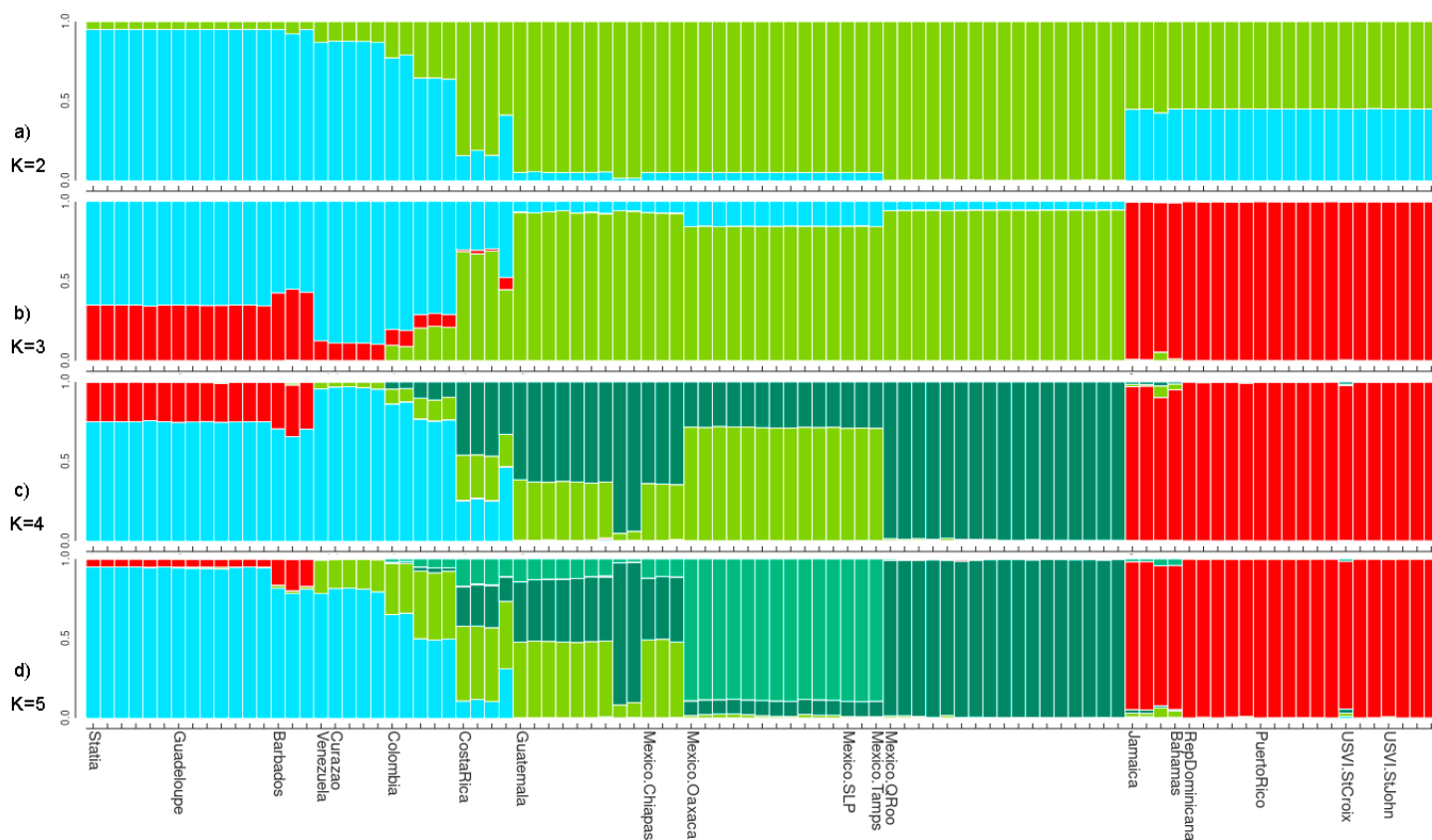


Figura 9. Resultados de los análisis de STRUCTURE realizados con datos genómicos de *Euphorbia tithymaloides*, asumiendo 2-5 grupos genéticos y usando el ensamble *min_samples_locus02*.

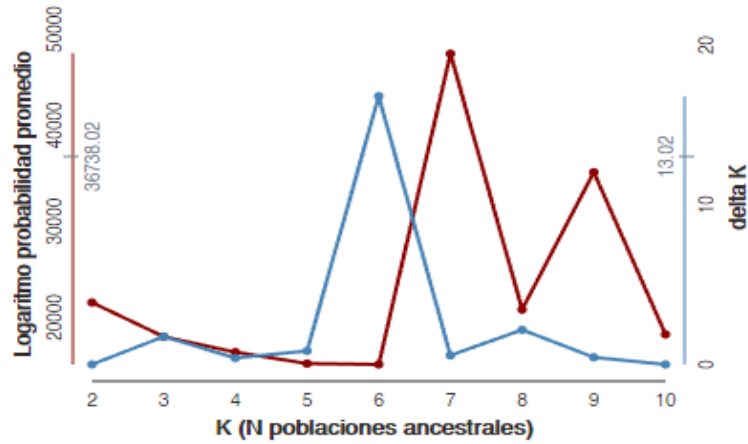


Figura 10. La prueba de Evanno sugiere que el valor más adecuado de grupos genéticos en *Euphorbia tithymaloides* es K = 6. En azul, valores de delta K; en rojo, el logaritmo probabilidad promedio.

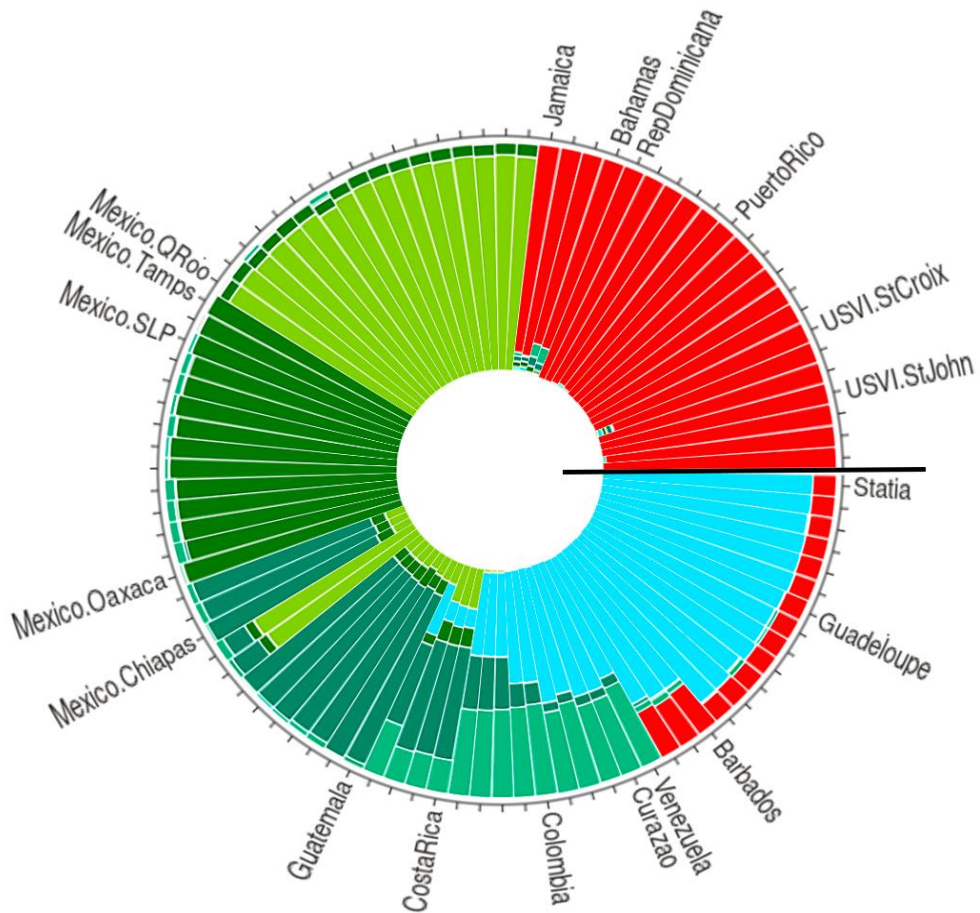


Figura 11. Resultados de los análisis de STRUCTURE realizados con datos genómicos de *Euphorbia tithymaloides*, asumiendo 6 grupos genéticos y usando el ensamble *min_samples_locus02*. La línea negra indica el Pasaje de Anegada.

Tabla 4. Resumen de los resultados de los distintos análisis en este trabajo y sus implicaciones principales, que se abordan en la discusión.

Análisis	Resultados e implicaciones	
RAxML	<p>En varias filogenias las Antillas Menores y Mayores aparecen como grupos hermanos, en un clado que aparece como grupo hermano de un grupo monofilético compuesto por Costa Rica, Colombia, Venezuela y Curazao, lo que implica que las Antillas fueron colonizadas a partir de un único evento que resulta difícil de conciliar con la geografía del Caribe (topología <i>1-origen</i>).</p>	<p>Varias filogenias indican que hubo dos eventos de colonización independientes de las Antillas Mayores y Menores. Costa Rica, Colombia, Venezuela y Curazao se agrupan con las Antillas Menores en un solo clado con disposición anidada que tiene sentido biogeográfico (topología <i>2-origenes</i>).</p>
	<p>Las poblaciones de Costa Rica a veces aparecen como un grupo monofilético y otras veces no.</p>	
PCAs	<p>Antillas Menores y Antillas Mayores aparecen muy separadas. Muestras de poblaciones continentales aparecen cercanas al grupo de las Antillas Menores. El grupo de las Antillas Mayores aparece alejado del resto de las muestras.</p>	
TreeMix	<p>En la mayor parte de los árboles se observa que las Antillas Menores forman un clado con poblaciones del continente, lo que es consistente con la hipótesis de dos eventos de colonización independientes (topología <i>2-origenes</i>). Varios resultados sugieren flujo génico entre muestras continentales y también desde las Antillas Mayores hacia las Antillas Menores. Las poblaciones de Costa Rica a veces aparecen como un grupo monofilético y otras veces no.</p>	
STRUCTURE	<p>Cuando se asumen dos grupos genéticos, un grupo tiene correspondencia con las muestras de México, Guatemala y Costa Rica, y el otro con correspondencia con las muestras del norte de Sudamérica y las Antillas Menores. Cuando se asumen de 3 a 6 grupos genéticos, las Antillas Mayores presentan pertenencia casi únicamente a un solo grupo, las Antillas Menores presentan alelos pertenecientes al de las Antillas Mayores.</p>	

Discusión

Este es el estudio que aborda la historia filogeográfica y patrones de divergencia de *Euphorbia tithymaloides* en su distribución circuncaribeña de manera más amplia, tanto en cuestión de muestreo molecular como de diversidad de análisis.

Los análisis de este estudio fueron abordados desde una perspectiva genómica, utilizando datos de RADseq, para ampliar nuestro conocimiento de un sistema que únicamente había sido abordado con la información de dos genes nucleares (*SGN-U34200* y *G3PDH*; Cacho et al., 2019, 2012; Monteverde-Suárez, 2021). El uso de tecnologías de secuenciación de alto rendimiento, y en particular con datos de RADseq, ha demostrado proporcionar nuevas perspectivas y mayor resolución a la hora de reconstruir la historia evolutiva de los taxones, y han resultado ser particularmente útiles en estudios acerca de especies-anillo (p.ej., Alcaide et al., 2014; Bouzid et al., 2021; Hernández et al., 2022; Wen et al., 2021).

En comparación con los estudios previos, en este trabajo se cuenta con más áreas geográficas representadas, pues se incluyen muestras provenientes de Colombia y la isla de Curazao. Este trabajo, además, amplía la variedad de análisis que han sido realizados para este sistema, abordándolo no solamente desde un enfoque filogenético, sino también realizando análisis de estructura genética, análisis de componentes principales y análisis de flujo génico.

Varios de los resultados y las conclusiones de este trabajo son consistentes con las de estudios previamente realizados. Por ejemplo, entre los resultados de este estudio se corrobora la evidencia de cohesión genética entre individuos de determinadas zonas geográficas (p.ej., monofilia en individuos de las Antillas Mayores y Antillas Menores), lo cual es consistente con estudios filogenéticos previos (Cacho y Baum 2012).

Desde un punto de vista filogeográfico, este estudio proporciona mayor evidencia de que el centro de expansión geográfica de *Euphorbia tithymaloides* es la región de México/Guatemala. Esta conclusión se obtiene de los análisis filogenéticos, sobre todo de los árboles en los que se incluyen datos genómicos de otras euforbiáceas, pues en estos resultados aparecen las muestras de esta zona geográfica como grupo hermano del resto de las muestras. La evidencia de México/Guatemala como centro de origen hipotético es consistente con los resultados de análisis filogenéticos anteriormente realizados (Cacho y Baum 2012).

Este estudio también es congruente con estudios previos en cuanto a que proporciona evidencia clara de una expansión geográfica de *E. tithymaloides* desde México/Guatemala hacia el sur, pasando por Centroamérica y llegando al norte de Sudamérica, desde donde se colonizaron las Antillas Menores en una progresión sur a norte (de Barbados hacia Guadalupe y St. Eustaquio). Esta hipótesis aparece fuertemente soportada por los resultados de este trabajo, pues en las filogenias se aprecia una disposición anidada indicativa de direccionalidad en el rumbo de la colonización en cuestión, y en las gráficas de los PCAs y los análisis de estructura genética se aprecia un patrón de ancestría mezclada que sugiere una conexión en forma de una transición gradual entre los grupos referidos. Por otro lado, los individuos de las Antillas Mayores parecen formar un grupo genético y linaje cohesivo cuyo origen permanece aún poco claro: los PCAs no identifican poblaciones continentales que muestren una cercanía genética clara con las poblaciones de las Antillas Mayores, y

los análisis de estructura genética no revelan variación genética compartida con grupos continentales. Estos patrones son consistentes con resultados filogenéticos (Cacho y Baum 2012), así como de PCA y de estructura poblacional (Cacho et al. 2019, Monteverde-Suárez 2021).

Adicionalmente, este trabajo aporta evidencia que proporciona nuevas perspectivas sobre el sistema de *E. tithymaloides* alrededor de la cuenca del Caribe. Por ejemplo, es destacable la presencia de alta diversidad genética en las poblaciones de Costa Rica. Esto se ve reflejado en la gráfica del análisis de estructura genética, en la que los individuos pertenecientes a estas poblaciones presentan pertenencia parcial a un mayor número de grupos genéticos, en comparación con las otras muestras. Otro resultado que podría relacionarse con esta diversidad es el hecho de que para algunas reconstrucciones filogenéticas los individuos de Costa Rica formaron un grupo monofilético o aparecían cercanamente relacionados (min75), pero para otras reconstrucciones no (astral_consensus_min2).

Otro aporte de este estudio es la evidencia clara de diversas pozas génicas en las poblaciones de México/Guatemala, lo cual es consistente con los resultados de los análisis de estructura genética por la aparición de distintos grupos genéticos en estas zonas. Por ejemplo, en algunas poblaciones existe alta compartición de alelos (Chiapas y Guatemala) y en otras poca o nula (el resto de las poblaciones de México, que forman dos grupos genéticos bien definidos). Un resultado similar se puede observar en los árboles filogenéticos, en los que las distintas poblaciones de Quintana Roo y Guatemala no forman grupos monofiléticos, sino que se encuentran dispersas en distintos clados. Esta evidencia es congruente con análisis filogenéticos previos basados en genes nucleares, en los cuales las poblaciones de México y Centroamérica presentaron el nivel más alto de heterocigosidad (Cacho y Baum 2012). La diversidad genética que poseen las poblaciones de esta zona es consistente con la hipótesis de que la región de México/Guatemala es el centro de origen de *E. tithymaloides*, y desde donde ocurrió la expansión geográfica de esta especie. Es importante tomar en cuenta que las 15 especies que pertenecen al clado o subgénero *Pedilanthus* se pueden encontrar en México, y la mayoría son endémicas a este país (Dressler 1957, Webster 1994, Steinmann 2003, Cacho et al. 2010). Desde la perspectiva del género *Euphorbia*, se ha hipotetizado que este es el segundo género botánico con mayor número de especies en México, sólo por detrás de *Salvia* (Steinmann 2002). Además, desde el nivel de familia, la familia Euphorbiaceae ocupa el sexto lugar en diversidad de familias de plantas en México, y de las 782 especies existentes en el país, el 57% son endémicas a este país (Steinmann 2002). La región de México/Centro América también es centro de diversificación de otros grupos pertenecientes a la familia Euphorbiaceae. Por ejemplo, aunque se ha hipotetizado que el género *Croton* tiene su origen en lo que hoy es Sudamérica, el grupo *Croton* sección *Corylocroton* parece tener su origen de diversificación en México/Centro América (Van Ee et al. 2008). Asimismo, análisis filogenéticos han sugerido que el género *Manihot* se pudo haber originado y diversificado en la región de Mesoamérica, para posteriormente aumentar su distribución geográfica por Sudamérica (Duputié et al. 2011). Algo similar ocurre con el clado Acalyphoideae *sensu* Wurdack, el mayor grupo dentro de Euphorbiaceae (Wurdack et al. 2005), pues análisis de áreas ancestrales sugieren que diversas especies de este grupo que habitan en las islas del Caribe se originaron a partir de poblaciones ancestrales provenientes de la región de México-Mesoamérica (Cervantes et al. 2016).

Desde una perspectiva filogenética, una contribución novedosa de este estudio es que se recuperan dos topologías contrastantes con implicaciones distintas sobre la historia de expansión

biogeográfica de *Euphorbia tithymaloides*. Por un lado, la topología en la que las secuencias provenientes de poblaciones insulares se agrupan en un clado (con dos subclados que corresponden a las Antillas Mayores y Antillas Menores), el cual es hermano a un clado formado por secuencias de Centro y Sudamérica, sugiere que la colonización de la mayoría de las islas del Caribe pudo haber ocurrido en un único evento a partir de las poblaciones continentales (topología *1-origen*). Por otro lado, la topología en la que las Antillas Menores forman un clado con distintas muestras continentales, es consistente con la hipótesis de que la colonización del Caribe por *E. tithymaloides* se dio a través de dos eventos o frentes: un origen incierto de colonización de las Antillas Mayores, y una colonización de las Antillas Menores desde el norte de Sudamérica (topología *2-origenes*).

La topología *1-origen* no implica del todo una progresión biogeográfica lineal debido a que no muestra una disposición anidada, es decir, no observamos un orden de ramificación que parezca indicar que la colonización de *E. tithymaloides* haya comenzado en las Antillas Menores y posteriormente haya progresado hacia las Antillas Mayores, o viceversa. Además, esta topología sugiere que en ambos clados de las Antillas, las poblaciones que parecen haber sido las primeras en divergir en comparación con las otras poblaciones son las que pertenecen a las islas más cercanas al continente (Barbados en el caso de las Antillas Menores, y Jamaica y Bahamas en el caso de las Antillas Mayores). Este patrón no es consistente con una progresión biogeográfica lineal, pero sí es consistente con una progresión a través de dos frentes. Hay que destacar, además, que la hipótesis de un solo evento de colonización que sugiere la topología *1-origen* parece ser poco probable en términos biogeográficos, pues es inverosímil (aunque no imposible) que dos grupos de islas tan extensos geográficamente hayan sido colonizados simultáneamente y desde sus puntos más lejanos (Jamaica y Barbados) en un solo evento de movilización desde el continente (o dos eventos muy cercanos en el tiempo desde la misma población ancestral; Figura 12). Por otro lado, la hipótesis de dos frentes de colonización (la cual parece sugerir la topología *2-origenes*) ha sido fundamentada por resultados de estudios biogeográficos previos (Cacho y Baum 2012, Cacho et al. 2019, Monteverde-Suárez 2021). Dressler (1957) fue el primero en hipotetizar una colonización del Caribe por *E. tithymaloides* a través de dos frentes, pero las formas terminales que sugirió para dichos frentes fueron las subespecies *parasitica* y *angustifolia*, planteando que *angustifolia* (que se distribuye por las A. Mayores) y *padifolia* (que se distribuye por las A. Menores) eran formas cercanamente relacionadas utilizando como referencia su morfología floral. Posteriormente se sugirió que las subespecies *padifolia* y *angustifolia*, lejos de ser hermanas, son las formas terminales de la distribución. La similitud morfológica entre estas dos subespecies, que causó que Dressler (1957) originalmente las considerara hermanas, parece sugerir un proceso de convergencia evolutiva en la variación morfológica del ciatio (Cacho et al. 2019, Monteverde-Suárez 2021). Esta clase de semejanzas o paralelismo que poseen los individuos de las poblaciones terminales de dos frentes de expansión geográfica también se ha reportado para otras especies-anillo, como en el caso de la coloración del lirio mariposa *Calochortus venustus* en California (Hernández et al. 2022) y la longitud y complejidad del canto, la morfología y el plumaje del mosquitero verdoso *Phylloscopus trochiloides* en el Tíbet (Irwin et al. 2001b).

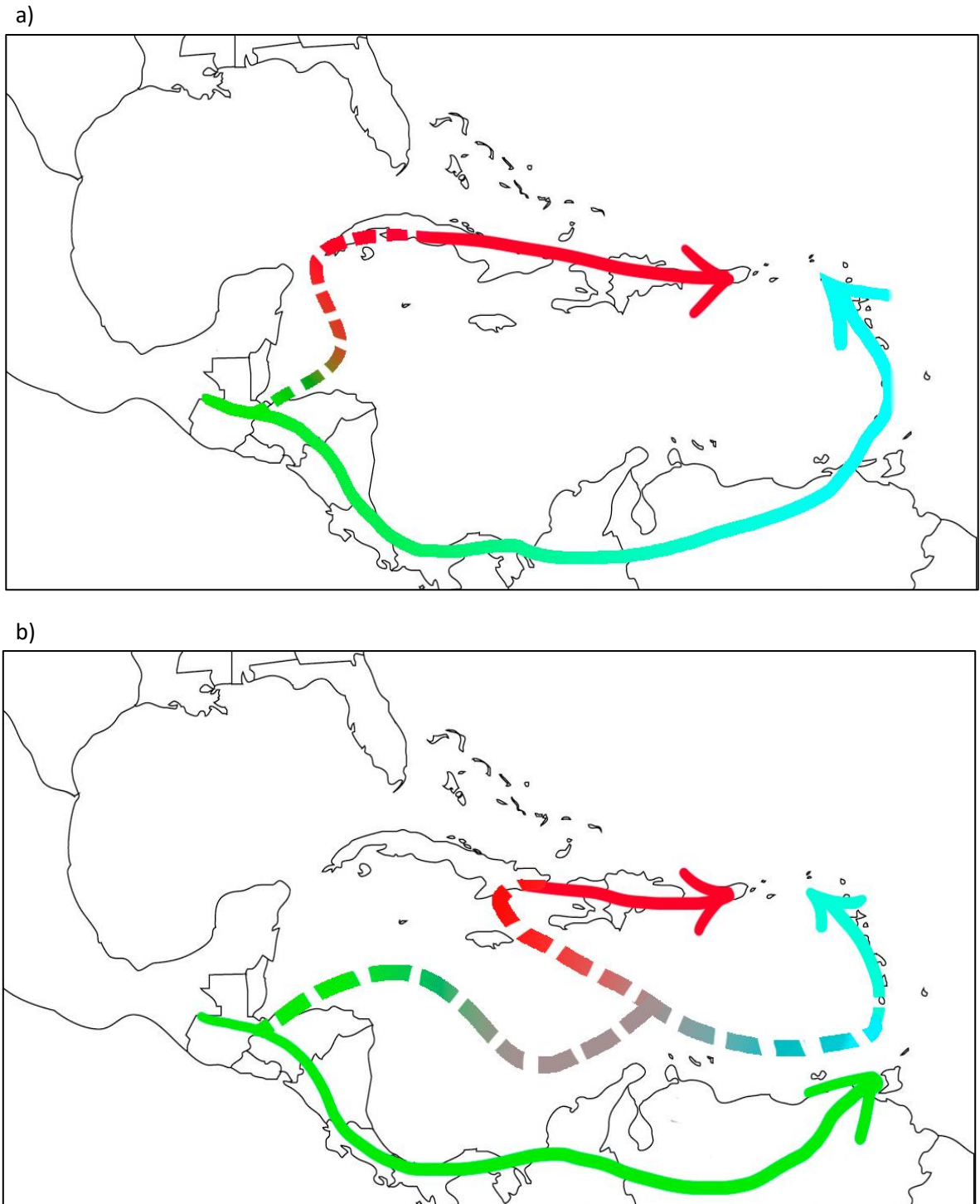


Figura 12. Representaciones gráficas de las dos hipótesis soportadas por los dos tipos de topología resultantes de los distintos ensambles y sus implicaciones biogeográficas sobre la colonización del Caribe por *Euphorbia tithymaloides*. Las líneas discontinuas representan eventos de colonización hipotéticos que se asumen a partir de las topologías. a) escenario que implica dos frentes de colonización, pues las Antillas Mayores y Menores no aparecen como grupos hermanos y éstas últimas aparecen anidadas en un clado que tiene su origen en Centro y Sudamérica, b) escenario en el que las Antillas Mayores y Menores aparecen como grupos hermanos formando un clado, implicando un único evento de colonización del Caribe.

Otro aporte importante de este estudio son los análisis para caracterizar explícitamente el flujo génico entre poblaciones (estadístico m , TreeMix). Estos análisis sugieren la existencia de flujo génico entre zonas continentales, y de forma importante, entre los grupos de las Antillas Mayores y las Antillas Menores, algo que anteriormente únicamente había sido inferido por dos resultados de estudios previos: a) la evidencia de una población aparentemente híbrida en Vieques o polimorfismos compartidos potencialmente ancestrales (Cacho y Baum 2012), y b) un análisis de estructura poblacional que muestra evidencia de compartición de alelos (Monteverde-Suárez 2021). Indicios de esta transferencia de material genético potencial entre las Antillas Mayores y las Antillas Menores también aparecen en el análisis de STRUCTURE realizado en este estudio, que se ve reflejado por el hecho de que las muestras de las Antillas Menores contienen alelos pertenecientes al grupo genético de las Antillas Mayores, pero no viceversa. Esto podría ser indicio de flujo génico unidireccional o de que el grupo génico de las Antillas Mayores se pudo haber originado a partir de las poblaciones de las Antillas Menores (aunque esta última hipótesis no es soportada por ninguno de los otros análisis, como los análisis filogenéticos o los PCAs).

Estos indicios de flujo génico entre las muestras de las Antillas Mayores y Menores podrían explicar por qué se obtienen dos topologías contrastantes. Más específicamente, podría ser la razón por la que la topología *1-origen* presenta a los dos clados de las Antillas como grupos hermanos. Además, la existencia de flujo génico entre los grupos de las A. Mayores y Menores podría explicar (por lo menos de forma parcial) la tendencia de las poblaciones de las islas a tener una morfología de la inflorescencia más parecida al acercarse al Pasaje de Anegada, la zona que divide a las dos Antillas y que se ha hipotetizado que podría fungir como barrera al flujo génico (Cacho y Baum 2012, Cacho et al. 2019). Esta explicación podría ser la razón por la que Dressler (1957) consideraba que estas dos subespecies (*padifolia* y *angustifolia*) eran hermanas. La existencia de flujo génico entre las poblaciones terminales de dos frentes geográficos se ha observado anteriormente en diversos ejemplos de especies-anillo putativas, como ocurre en los peces trompeta del género *Aulostomus* en el suroeste del Océano Atlántico (Bowen et al. 2001), los lagartos *Sceloporus occidentalis* en el oeste de Norte América (Bouzid et al. 2021), y en las salamandras *Ensatina eschscholtzii* en California (Stebbins 1957, Brown y Stebbins 1964, Wake et al. 1986, Devitt et al. 2011). Este último caso es relevante debido a que el flujo genético entre los dos taxones terminales es unidireccional, tal como parece ocurrir en *E. tithymaloides* según los resultados del análisis de estructura genética.

Este estudio cuenta con varias limitaciones. Una de ellas es la posibilidad de que el flujo génico que aparentemente existe entre los dos grupos de las Antillas pueda causar cierta interferencia al momento de inferir la historia evolutiva de estas poblaciones, lo que podría contribuir a explicar la existencia de las dos topologías contrastantes obtenidas a partir de los análisis filogenéticos de máxima verosimilitud. También, existen ciertos aspectos de este estudio que podrían limitar la capacidad de inferencia de la historia filogeográfica de la especie. Por ejemplo, una diferencia de este estudio con estudios anteriores es que aquellos utilizaban información de un mayor número de muestras (42 poblaciones, cada una de las cuales se obtuvieron 1–16 individuos). Asimismo, es posible que una mayor representación de ciertas zonas geográficas (como Florida, Centroamérica y otras islas del Caribe) pueda aumentar la resolución en los resultados de este estudio. Quizás las de mayor relevancia son las poblaciones de Cuba, y posiblemente también las de Florida, cuya inclusión podría resolver aspectos que no pudieron ser resueltos en este trabajo, por

ejemplo, el hecho de que genéticamente no se encontró ningún vínculo directo o indicio de continuidad de las poblaciones de las Antillas Mayores y el Continente tanto en los PCAs como en los análisis filogenéticos ni de estructura genética. Es decir, no existe evidencia clara sobre cómo pudo haber ocurrido la colonización de las Antillas Mayores, a pesar de que estudios previos sugieren un frente de colonización a partir del continente independiente al frente de colonización de las Antillas Menores (Cacho y Baum 2012, Cacho et al. 2019, Monteverde-Suárez 2021). La discontinuidad entre los grupos del continente y las Antillas Mayores no necesariamente invalida la hipótesis de los dos frentes de colonización de las Antillas. Existen casos de especies anillo putativas en las que también se han encontrado discontinuidades genéticas que se cree son un artefacto de limitaciones en el muestreo, en particular de muestreo escaso o ausente en ciertas zonas de la distribución. Un ejemplo es el mosquitero verdoso *Phylloscopus trochiloides* que habita alrededor de la meseta del Tíbet, cuya distribución presenta discontinuidades genéticas entre las regiones de Kirguistán y Pakistán, Nepal y China, y China y el este de Siberia, las cuales son atribuidas a falta de muestreo. Sin embargo, existe una discontinuidad genética en el suroeste de la distribución de este mosquitero, que no parece ser causada por limitaciones en el muestreo, sino por divergencia alopátrica y posterior contacto secundario (Alcaide et al. 2014). Otro ejemplo es la abeja *Megachile concinna* que habita alrededor del Mediterráneo, y que presenta una ausencia significativa de muestreo en la región del norte de África, entre Egipto y Túnez (Soltani et al. 2017).

Por otro lado, con respecto al muestreo molecular, se debe tomar en cuenta que a pesar de que los datos producidos a partir de las técnicas de secuenciación de alto rendimiento en diversos casos han demostrado ser mejores para informar y responder preguntas en comparación con los datos de unos cuantos genes, éstos también presentan limitaciones importantes. Una limitación al momento de utilizar datos de RADseq son los errores de genotipación, los cuales pueden tener causas técnicas o humanas (para un resumen de posibles errores, ver Mastretta-Yanes et al., 2015). Para resolver el problema de una genotipación inadecuada, en este trabajo se llevaron a cabo distintos ensambles a partir de los datos de RADseq pero evaluando diferentes valores para determinados parámetros. Para el parámetro de umbral de similitud se realizaron pruebas con un intervalo de valores para escoger el valor que proporcionara la mayor cantidad de información (SNPs y RADloci o secuencias) en el ensamble final (método sugerido en: Mastretta-Yanes et al., 2015; McCartney-Melstad et al., 2019). Para el parámetro de número mínimo de muestras por locus, se utilizó el valor de 2 pues este es el menos “restrictivo” al momento de definir las relaciones entre las muestras, es decir, acepta que para cada locus deben existir por lo menos dos muestras que lo compartan. Este enfoque de maximizar la cantidad de información y permitir la mayor cantidad de relaciones entre las muestras proporcionó buenos resultados al momento de realizarse los diversos análisis, pero no garantiza la ausencia de errores o artefactos.

Otra posible limitación de los datos de RADseq es que los ensambles *de novo* que se realizan utilizando esa clase de datos no permiten inferir haplotipos, sino sólo genotipos. Esta problemática se puede resolver a partir de un método llamado *phasing*, que busca inferir la relación correcta entre los alelos (*cis* o *trans*) en múltiples loci y por lo tanto permite la inferencia de haplotipos a partir de información genómica (Browning y Browning 2011, Martin et al. 2016). Sin embargo, para realizar *phasing* es necesario contar con lecturas (*reads*) de mayor longitud en comparación con las producidas por los métodos de secuenciación reducida del genoma (Wenger et al. 2019). Esta

limitante puede influir en las inferencias de tipo filogenético y filogeográfico debido a que la consideración de los dos alelos de cada muestra (y su debido orden en cada cromosoma) proporciona más información para resolver preguntas a escala evolutiva, como la interfaz entre población y especie. Por ejemplo, la información proporcionada por la inferencia de haplotipos (*phasing*) ha sido utilizada para detectar selección positiva (p. ej. Sabeti et al., 2007) y para reconstruir historia poblacional (p. ej. Harding et al., 1997; Yunusbaev et al., 2019).

Con respecto al tipo de análisis, se debe tener en cuenta que los análisis filogenéticos de máxima verosimilitud que fueron realizados asumen una hipótesis de parentesco de forma bifurcante. Sin embargo, los resultados de este trabajo sugieren la existencia de flujo génico entre las Antillas Mayores y Menores, lo cual daría como resultado las topologías contrastantes obtenidas. El intercambio genético entre linajes no puede ser representado de manera bifurcante, y puede considerarse un caso de evolución reticulada (Linder *et al.*, 2004). La evolución reticulada (o en red) engloba fenómenos y eventos como la recombinación meiótica y sexual, la transferencia horizontal de genes, la especiación por hibridación, la introgresión y la separación incompleta de linajes (Linder *et al.*, 2004; Baum, 2007; Eaton & Ree, 2013). Para analizar sistemas biológicos que presentan esta clase de evolución, se han desarrollado métodos que toman en cuenta la diferencia entre los árboles de genes y los árboles de especies o linajes, como los análisis de concordancia de árboles (Baum, 2007; Larget *et al.*, 2010), los análisis de coalescencia (Liu, 2008), análisis de redes (Allman *et al.*, 2019), y diversos métodos estadísticos. Uno de estos métodos es el estadístico D de Patterson, el cual sirve para detectar y medir la presencia de introgresión causada por mezcla o *admixture* ancestral, definiendo *admixture* como la mezcla que ocurre por contacto secundario de dos poblaciones que anteriormente estuvieron aisladas (Durand *et al.*, 2011). La prueba del estadístico D, o prueba ABBA-BABA, detecta y mide la introgresión por medio de la comparación en la ocurrencia de dos escenarios diferentes (“ABBA” y “BABA”) definidos por la configuración de los alelos “A” y “B” en un árbol anidado de cuatro terminales (E (T₁ (T₂, T₃))) (Fig. 13a, Durand *et al.*, 2011; Eaton & Ree, 2013). De esta forma, en el escenario “ABBA” la terminal T₃ compartiría el alelo “A” con el grupo externo, mientras que las terminales T₁ y T₂ compartirían el alelo derivado “B”. Por otra parte, en el escenario “BABA” la terminal T₂ compartiría el alelo “A” con el grupo externo, mientras que el alelo derivado “B” sería compartido por las terminales T₁ y T₃. La ocurrencia de los escenarios ABBA y BABA se cuenta en diferentes sitios a través de las *n* pares de bases con las que se cuenta, y posteriormente dicho conteo se compara usando la fórmula:

$$D(T_1, T_2, T_3, E) = \Sigma (C_{ABBA} - C_{BABA}) / \Sigma (C_{ABBA} + C_{BABA}).$$

Para poder aplicar la prueba estadística en este trabajo, se tendrían que considerar dos escenarios biogeográficos distintos tomando en cuenta cuatro terminales diferentes, las cuales podrían representar distintas áreas geográficas de la distribución de *E. tithymaloides*. En el primer escenario se tomaría en cuenta una topología similar a *1-origen* (Fig. 13b), y en el segundo una topología similar a *2-orígenes* (Fig. 13c). En la topología *1-origen*, el grupo externo podría estar representado por un individuo de México/Guatemala o un individuo de la especie hermana *E. personata*, T₁ por un individuo del norte de Sudamérica, T₂ por un individuo de las Antillas Menores, y T₃ por un individuo de las Antillas Mayores. En la topología *2-orígenes*, T₁ representaría a las Antillas Mayores, T₂ a Sudamérica y T₃ a Antillas Mayores. En el caso de que no hubiera nada de introgresión, lo cual sería un resultado muy poco verosímil dadas las dos topologías contrastantes que frecuentemente eran inferidas a partir de los datos, la D de Patterson tendría un valor de 0 para alguna de las dos topologías (la que representara mejor las relaciones de parentesco de las muestras). En el caso más verosímil, obtendríamos un valor distinto a 0 para las dos topologías, causado por una sobrerrepresentación de la configuración ABBA o BABA. Con la topología *1-origen* habría una sobrerrepresentación de la configuración ABBA causada por la compartición de alelos derivados (“B”)

entre la muestra proveniente de Sudamérica y la de las Antillas Menores (dicha compartición de alelos sería la causante de que surja la topología *2-origenes* en los análisis filogenéticos). Con la topología *2-origenes* se encontraría una sobrerrepresentación de la configuración BABA causada por la compartición de alelos entre las Antillas Mayores y Menores (posible explicación de que surja la topología *1-origen*).

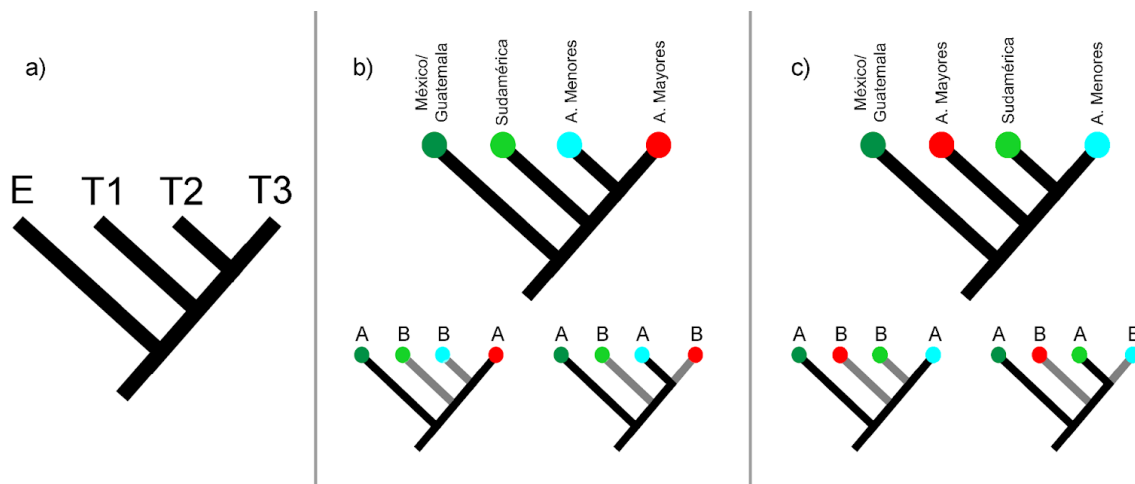


Figura 13. a) Árbol con disposición anidada de cuatro taxones (E (T1 (T2, T3))) en el que se basa el análisis ABBA-BABA. b) Escenario similar a la topología *1-origen* y sus configuraciones ABBA y BABA. c) Escenario similar a la topología *2-origenes* y sus configuraciones ABBA y BABA.

Desde una perspectiva más amplia, este trabajo aporta elementos importantes sobre nuestro entendimiento de la utilidad del modelo conceptual de especie-anillo. *Euphorbia tithymaloides* fue considerada formalmente como especie-anillo por primera vez a partir de estudios filogenéticos que mostraron evidencia de una expansión a través de dos frentes alrededor del Caribe (Cacho y Baum 2012). La cantidad y la especificidad de los “requisitos” para que una especie sea considerada especie-anillo pueden variar dependiendo de los autores. Sin embargo, los requisitos mínimos o más básicos reconocidos por la mayor parte de los autores son dos: i) una serie de formas progresivamente intermedias deben poseer una distribución geográfica en forma de anillo resultado de una expansión histórica, y ii) las formas terminales de la distribución deben mostrar aislamiento reproductivo (Irwin et al. 2001a, Patten y Pruett 2009, Cacho y Baum 2012, Smyth et al. 2015). Un tercer requisito, considerado por ciertos autores, es el de la existencia de flujo génico entre las poblaciones intermedias del anillo (Irwin et al. 2001a, Cacho y Baum 2012). Este tercer requisito está muy relacionado con el de la existencia de intergradación entre las poblaciones intermedias, y proporciona mayores implicaciones al modelo. Por ejemplo, a pesar de que no existe flujo génico directo entre las formas terminales, se puede considerar que existe el potencial de flujo génico indirecto entre ellas por medio de las poblaciones intermedias. A pesar de que hay una cierta cantidad de especies que han sido propuestas como especies-anillo, y el número de estas ha ido aumentando, es difícil que las especies cumplan con los requisitos ya mencionados al pie de la letra, lo cual ha sido discutido por otros autores (Irwin et al. 2001a). Eso puede ocurrir ya sea por causas intrínsecas al sistema (por ejemplo, ausencia de aislamiento reproductivo entre las poblaciones terminales) como por artefactos causados por cuestiones metodológicas (por ejemplo, falta de muestreo en ciertas regiones de la distribución). A partir de la evidencia de este estudio, se vuelven complicadas las cuestiones de si *E. tithymaloides* es un ejemplo de especie-anillo y si el modelo de especie-anillo es útil o informativo

para describir a este tipo de sistemas biológicos (en los que existe diversidad resultado de un proceso de divergencia en curso).

Mayr (1963) menciona que en casi todos los casos bien analizados de especies-anillo existen huecos significativos en las cadenas de poblaciones que conforman los frentes geográficos, o por lo menos hay evidencia de la existencia pasada de esos huecos. Esta idea continúa siendo vigente incluso con los casos mejor estudiados de especies-anillo, como ya se mencionó anteriormente sobre *Phylloscopus trochiloides* (Alcaide et al. 2014). Otro ejemplo de esto se observa en la falta de intergradación genética continua en distintas zonas de la distribución de la salamandra *Ensatina eschscholtzii* (Wake y Yanev 1986, Highton 1998). A raíz de esta evidencia, se ha llegado a sugerir que *E. eschscholtzii* no puede ser considerada como especie-anillo (Highton 1998), a pesar de que incluso en estudios más recientes se sigue considerando como tal (p.ej., Devitt et al., 2011; Kuchta et al., 2009). Este tipo de desacuerdos ocurre por proporcionarle más peso o énfasis a ciertas características del modelo que a otras (p.ej., dar más peso a la historia biogeográfica de la formación del anillo que a los requisitos sobre flujo génico). A pesar de que se ha demostrado a lo largo del tiempo que la adecuación al modelo de especie-anillo pocas veces se cumple al pie de la letra, Mayr (1942, 1963, 1970) mencionaba a las especies-anillo como “la demostración perfecta de especiación”, debido a que “el proceso de especiación geográfica puede ser seguido paso por paso”.

De manera general, este estudio aporta datos a escala genómica para profundizar nuestro conocimiento sobre la biogeografía histórica de *Euphorbia tithymaloides* en la cuenca del Caribe, analizados desde perspectivas filogenética, filogeográfica y de genética de poblaciones. Este trabajo aporta elementos nuevos que permiten evaluar escenarios alternativos sobre cómo *E. tithymaloides* colonizó el Caribe, un resumen de los cuales se presenta en la tabla 5.

Tabla 5. Resumen de la evidencia a favor y en contra de los dos escenarios de colonización de *Euphorbia tithymaloides* en el Caribe (Un origen vs. dos orígenes).

	HIPÓTESIS DE UN SÓLO FRENTE DE COLONIZACIÓN	HIPÓTESIS DE DOS FRENTE DE COLONIZACIÓN
EVIDENCIA A FAVOR	<ul style="list-style-type: none"> ● Análisis filogenéticos de máxima verosimilitud con diversos ensambls resultan en topologías en las que las Antillas Menores y las Antillas Mayores son grupos hermanos. ● En el análisis de STRUCTURE, el grupo de las Antillas Menores presenta alelos del grupo de las Antillas Mayores, pero no al revés. Esto podría ser consistente (aunque no necesariamente) con que el grupo de las Antillas Mayores pudo haberse originado de las Antillas Menores. 	<ul style="list-style-type: none"> ● Hipótesis soportada por estudio de Cacho y Baum (2012). ● Análisis filogenéticos de máxima verosimilitud con diversos ensambls resultan en topologías en las que las Antillas Menores y las Antillas Mayores no son grupos hermanos. ● Topología de las filogenias es la más verosímil biogeográficamente (Figura 11a). ● La mayor parte de los análisis de Treemix no muestran a las Antillas Mayores y Menores como grupos hermanos, y varios de éstos muestran evidencia de flujo génico entre las dos Antillas. ● La evidencia de flujo génico sugerida por Treemix podría explicar las topologías contrastantes y la aparición de alelos de las Antillas Mayores en las Antillas Menores que se aprecia en los resultados de STRUCTURE. ● En el análisis de componentes principales las muestras de las Antillas Menores y las de las Antillas Mayores siempre aparecen distantes y no se agrupan.
EVIDENCIA EN CONTRA	<ul style="list-style-type: none"> ● Un escenario de dos eventos de colonización desde una misma área geográfica y muy cercanos en el tiempo desde el continente hasta el Caribe resulta inverosímil (Figura 11b). 	<ul style="list-style-type: none"> ● En ningún resultado de los análisis se aprecia una conexión clara entre las accesiones del continente y las de las Antillas Mayores.

Conclusiones

Este estudio ha profundizado nuestro entendimiento de la historia evolutiva de *Euphorbia tithymaloides* en su distribución circuncaribeña. En primer lugar, ha proporcionado soporte al hecho de que México y Centroamérica representan el centro de inicio de la expansión geográfica de la especie, o que los individuos de las Antillas Mayores y de las Antillas Menores forman grupos linajes evolutivos independientes. Este estudio además ha proporcionado nueva información, como la existencia de flujo génico entre áreas continentales e incluso entre las Antillas Mayores y las Antillas Menores. También, se destaca la existencia de una alta diversidad genética dentro de las poblaciones pertenecientes a Costa Rica, y la evidencia de diversas pozas génicas entre las poblaciones de México y Guatemala.

A pesar de que diversos resultados de este estudio son consistentes con trabajos anteriores, también hay diferencias importantes con estudios previos. Quizás el más destacable es la evidencia de dos topologías contrastantes: una que sugiere un único evento de colonización del Caribe y otra que sugiere dos eventos de colonización de las Antillas. Globalmente, los análisis de este trabajo sugieren que la hipótesis de dos frentes parece tener mayor evidencia a favor. Otra aportación importante de este estudio es la documentación de la falta de una conexión clara entre las poblaciones de las Antillas Mayores y las poblaciones continentales; de forma que con los resultados de estos análisis no es posible concluir cómo pudo haberse colonizado este grupo de islas a partir del continente. Es importante resaltar que a partir de lo encontrado en este estudio no es posible descartar por completo la hipótesis de un solo evento de colonización de las Antillas.

Aunque una parte de la evidencia de este estudio sugiere que *E. tithymaloides* cumple con los requisitos del modelo especie-anillo sólo parcialmente, los resultados son consistentes con la hipótesis de una ampliación de la distribución geográfica a través de dos frentes y la evolución de cierto aislamiento reproductivo entre las formas terminales. A pesar de que el modelo de especie-anillo no se cumpla a la perfección, se puede llegar a la conclusión de que utilizar dicho marco teórico como hipótesis para estudiar procesos de divergencia y especiación con flujo génico, y divergencia y especiación por distancia.

De manera general, este estudio aporta datos de escala genómica para profundizar nuestro conocimiento sobre la biogeografía histórica de *Euphorbia tithymaloides* en el Caribe, analizados desde una visión que integra análisis desde una perspectiva filogenética, filogeográfica y de genética de poblaciones. Si bien no existe soporte inequívoco para un escenario biogeográfico único de cómo *Euphorbia tithymaloides* colonizó la cuenca del Caribe, este trabajo aporta elementos nuevos que permiten evaluar escenarios alternativos, y esto a su vez es una contribución a nuestro entendimiento de la biogeografía del Caribe, un 'hotspot' de diversidad en un área con una historia geológica altamente compleja (Santiago-Valentin y Olmstead 2004, Roncal et al. 2020).

Referencias

- Aguirre-Liguori JA, Luna-Sánchez JA, Gasca-Pineda J, Eguiarte LE (2020) Evaluation of the Minimum Sampling Design for Population Genomic and Microsatellite Studies: An Analysis Based on Wild Maize. *Frontiers in Genetics* 11: 1–16. <https://doi.org/10.3389/fgene.2020.00870>
- Alcaide M, Scordato ESC, Price TD, Irwin DE (2014) Genomic divergence in a ring species complex. *Nature* 511: 83–85. <https://doi.org/10.1038/nature13285>
- Aldrich J (1997) R. A. Fisher and the making of maximum likelihood 1912 - 1922. *Statistical Science* 12: 162–176. <https://doi.org/10.1214/ss/1030037906>
- Allman E, Baños H, Rhodes J (2019) NANUQ: a method for inferring species networks from gene trees under the coalescent model. *Algorithms for Molecular Biology* 14. 10.1186/s13015-019-0159-2.
- Amorim A (2013) 5 Brenner's Encyclopedia of Genetics: Second Edition Population Genetics. Elsevier Inc., 407–411 pp. <https://doi.org/10.1016/B978-0-12-374984-0.01195-5>
- Avice JC (2000) *Phylogeography: The History and Formation of Species*. Harvard University Press, Cambridge, MASS., USA, 464 pp.
- Avice JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE, Reeb CA, Saunders NC (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual review of ecology and systematics* 18: 489–522. <https://doi.org/10.1146/annurev.ecolsys.18.1.489>
- Ayala FJ, Tracey ML, Hedgecock D, Richmond RC (1974) Genetic Differentiation During the Speciation Process in *Drosophila*. *Evolution* 28: 576–592.
- Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, Selker EU, Cresko WA, Johnson EA (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE* 3: 1–7. <https://doi.org/10.1371/journal.pone.0003376>
- Baum DA (2007) Concordance trees, concordance factors, and the exploration of reticulate genealogy. *Taxon* 56(2), 417–426.
- Baxter SW, Davey JW, Johnston JS, Shelton AM, Heckel DG, Jiggins CD, Blaxter ML (2011) Linkage mapping and comparative genomics using next-generation rad sequencing of a non-model organism. *PLoS ONE* 6. <https://doi.org/10.1371/journal.pone.0019315>
- Bivand R, Keitt T, Rowlingson B, Pebesma E (2016) *rgdal*: Bindings for the geospatial data abstraction library. R package version. Available from: <https://cran.r-project.org/web/packages/rgdal/index.html>.
- Bouzid NM, Archie JW, Anderson RA, Grummer JA, Leaché AD (2021) Evidence for ephemeral ring species formation during the diversification history of western fence lizards (*Sceloporus occidentalis*). *Molecular Ecology*: 1–12. <https://doi.org/10.1111/mec.15836>
- Bowen BW, Bass AL, Rocha LA, Grant WS, Robertson DR (2001) Phylogeography of the trumpetfishes (*Aulostomus*): Ring species complex on a global scale. *Evolution* 55: 1029–1039. <https://doi.org/10.1111/j.0014-3820.2001.tb00619.x>
- de Brito Martins A, de Aguiar MAM (2017) Barriers to gene flow and ring species formation. *Evolution* 71: 442–448. <https://doi.org/10.1111/evo.13121>
- Brown CW, Stebbins RC (1964) Evidence for hybridization between the blotched and unblotched subspecies of the salamander *Ensatina eschscholtzii*. *Evolution* 18: 706–707.
- Browning SR, Browning BL (2011) Haplotype phasing: Existing methods and new developments. *Nature Reviews Genetics* 12: 703–714. <https://doi.org/10.1038/nrg3054>
- Cacho NI, Baum DA (2012) The Caribbean slipper spurge *Euphorbia tithymaloides*: The first example of a ring species in plants. *Proceedings of the Royal Society B: Biological Sciences* 279: 3377–3383. <https://doi.org/10.1098/rspb.2012.0498>
- Cacho NI, Monteverde-Suárez MJ, McIntyre PJ (2019) Convergent evolution in floral morphology in a plant ring species, the Caribbean *Euphorbia tithymaloides*. *American Journal of Botany* 106: 1032–1045. <https://doi.org/10.1002/ajb2.1318>
- Cacho NI, Berry PE, Olson ME, Steinmann VW, Baum DA (2010) Are spurred cyathia a key innovation? Molecular systematics and trait evolution in the slipper spurge (Pedilanthus clade: *Euphorbia*, Euphorbiaceae). *American Journal of Botany* 97: 493–510. <https://doi.org/10.3732/ajb.0900090>
- Case JE, Holcombe TL, Martin RG (1984) Map of geologic provinces in the Caribbean region. *Geological Society of America Memoir* 162: 1–30.
- Castillo DM (2017) Factors contributing to the accumulation of reproductive isolation: A mixed model approach.

- Ecology and Evolution 7: 5808–5820. <https://doi.org/10.1002/ece3.3093>
- Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013) Stacks: An analysis tool set for population genomics. *Molecular Ecology* 22: 3124–3140. <https://doi.org/10.1111/mec.12354>
- Cécile A, Larget B, Baum DA, Smith SD, Rokas A (2007) Bayesian Estimation of Concordance among Gene Trees. *Molecular Biology and Evolution* 24(2): 412–426. <https://doi.org/10.1093/molbev/msl170>
- Cervantes A, Fuentes S, Gutiérrez J, Magallón S, Borsch T (2016) Successive arrivals since the Miocene shaped the diversity of the Caribbean Acalyphoideae (Euphorbiaceae). *Journal of Biogeography* 43: 1773–1785. <https://doi.org/10.1111/jbi.12790>
- Chakraborty R (1993) Analysis of Genetic Structure of Populations: Meaning, Methods, and Implications. In: *Human Population Genetics*. Springer, Boston, MA, 189–206.
- Chan AP, Crabtree J, Zhao Q, Lorenzi H, Orvis J, Puiu D, Melake-berhan A, Jones KM, Redman J, Chen G, Cahoon EB, Gedil M, Stanke M, Haas BJ, Wortman JR, Fraser-liggett CM, Ravel J, Rabinowicz PD (2010) Draft genome sequence of the ricin-producing oilseed castor bean. *Nature biotechnology* 28: 951–956. <https://doi.org/10.1038/nbt.1674>.Draft
- Collins RA, Hrbek T (2018) An in silico comparison of protocols for dated phylogenomics. *Systematic Biology* 67: 633–650. <https://doi.org/10.1093/sysbio/syx089>
- Combosch DJ, Vollmer S V. (2015) Trans-Pacific RAD-Seq population genomics confirms introgressive hybridization in Eastern Pacific Pocillopora corals. *Molecular Phylogenetics and Evolution* 88: 154–162. <https://doi.org/10.1016/j.ympev.2015.03.022>
- Coney PJ (1982) Plate tectonic constraints on the biogeography of middle America and the Caribbean region (Central America). *Annals - Missouri Botanical Garden* 69: 432–443. <https://doi.org/10.2307/2399080>
- Corlay S, Ivanov P, Avila D, Abdalla S, Willing C, Jupyter Development Team (2016) Jupyter Notebooks – a publishing format for reproducible computational workflows. In: Loizides F, Schmidt B (Eds), *Positioning and Power in Academic Publishing: Players, Agents and Agendas*. IOS Press., Amsterdam, 87–90. <https://doi.org/doi:10.3233/978-1-61499-649-1-87>
- Coyne JA, Orr HA (2004) *Speciation*. Sinauer Associates, Inc., Sunderland, MA, 545 pp.
- Craddock EM (1987) Degrees of reproductive isolation between closely related species of Hawaiian *Drosophila*. In: *Genetic Mechanisms of Speciation in Insects*. Springer, Dordrecht, 111–139.
- Cutter AD (2013) Integrating phylogenetics, phylogeography and population genetics through genomes and evolutionary theory. *Molecular Phylogenetics and Evolution* 69: 1172–1185. <https://doi.org/10.1016/j.ympev.2013.06.006>
- Darwell CT, Rivers DM, Althoff DM (2016) RAD-seq phylogenomics recovers a well-resolved phylogeny of a rapid radiation of mutualistic and antagonistic yucca moths. *Systematic Entomology* 41: 672–682. <https://doi.org/10.1111/syen.12185>
- Darwin C (1859) *On the origin of species by means of natural selection, or, the preservation of favoured races in the struggle for life*. J. Murray, London.
- Davey JL, Blaxter MW (2010) RADseq: Next-generation population genetics. *Briefings in Functional Genomics* 9: 416–423. <https://doi.org/10.1093/bfgp/elq031>
- Devitt TJ, Baird SJE, Moritz C (2011) Asymmetric reproductive isolation between terminal forms of the salamander ring species *Ensatina eschscholtzii* revealed by fine-scale genetic analysis of a hybrid zone. *BMC Evolutionary Biology* 11. <https://doi.org/10.1186/1471-2148-11-245>
- Dobzhansky T (1940) Speciation as a Stage in Evolutionary Divergence. *The American Naturalist* 74: 312–321. <https://doi.org/10.1086/280899>
- Dobzhansky T (1951) *Genetics and the Origin of Species*. 3rd edition. Columbia University Press, New York City, NY.
- Dobzhansky T, Spassky B (1959) *Drosophila Paulistorum*, a Cluster of Species in Statu Nascendi. *Proceedings of the National Academy of Sciences* 45: 419–428. <https://doi.org/10.1073/pnas.45.3.419>
- Dobzhansky T, Pavlovsky O (1962) A comparative study of the chromosomes in the incipient species of the *Drosophila paulistorum* complex. *Chromosoma* 13: 196–218. <https://doi.org/10.1007/BF00326571>
- Dobzhansky T, Ehrman L, Pavlovsky O, Spassky B (1964) The Superspecies *Drosophila Paulistorum*. *Proceedings of the National Academy of Sciences of the United States of America* 51: 3–9. <https://doi.org/10.1073/pnas.51.1.3>
- Dorsey BL, Haevermans T, Aubriot X, Morawetz JJ, Riina R, Steinmann VW, Berry PE (2013) Phylogenetics, morphological evolution, and classification of *Euphorbia* subgenus *Euphorbia*. *Taxon* 62: 291–315.

- <https://doi.org/10.12705/622.1>
- Draper G, Jackson TA, Donovan SK (1994) Geologic provinces of the Caribbean. In: Caribbean Geology: An Introduction. University of the West Indies Publishers' Association, Kingston, 3–12.
- Dressler RL (1957) The Genus *Pedilanthus*. The Gray Herbarium of Harvard University, Cambridge, MA, 188 pp.
- Dressler RL, Sacamano CM (1992) *Pedilanthus connatus* (Euphorbiaceae), especie nueva y llamativa de Jalisco. *Acta Botanica Mexicana*: 21. <https://doi.org/10.21829/abm18.1992.640>
- Duputié A, Salick J, Mckey D (2011) Evolutionary biogeography of *Manihot* (Euphorbiaceae), a rapidly radiating Neotropical genus restricted to dry environments. *Journal of Biogeography* 38: 1033–1043. <https://doi.org/10.1111/j.1365-2699.2011.02474.x>
- Eaton DAR (2014) PyRAD: assembly of de novo RADseq loci for phylogenetic analyses. *Bioinformatics* 30: 1844–1849. <https://doi.org/10.1093/bioinformatics/btu121>
- Eaton DAR, Ree RH (2013) Inferring phylogeny and introgression using RADseq data: An example from flowering plants (*Pedicularis*: Orobanchaceae). *Systematic Biology* 62: 689–706. <https://doi.org/10.1093/sysbio/syt032>
- Eaton DAR, Overcast I (2020) Ipyrad: Interactive assembly and analysis of RADseq datasets. *Bioinformatics* 36: 2592–2594. <https://doi.org/10.1093/bioinformatics/btz966>
- Van Ee BW, Berry PE, Riina R, Gutiérrez Amaro JE (2008) 74 Botanical Review Molecular phylogenetics and biogeography of the Caribbean-centered *Croton* subgenus *Moacrotroton* (Euphorbiaceae s.s.). 132–165 pp. <https://doi.org/10.1007/s12229-008-9003-y>
- Endler JA (1977) *Geographic Variation, Speciation, and Clines*. Princeton University Press, Princeton, NJ, 248 pp.
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology* 14: 2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>
- Felsenstein J (1981) Evolutionary Trees from DNA Sequences: A Maximum Likelihood Approach. *Journal of Molecular Evolution* 17: 368–376.
- Fisher RA (1922) On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society of London* 222: 594–604.
- Fisher RA (1930) *The Genetical Theory of Natural Selection*. Clarendon Press, London.
- Fitak RR (2021) OptM: Estimating the optimal number of migration edges on population trees using Treemix. *Biology Methods and Protocols* 6: 1–6. <https://doi.org/10.1093/biomethods/bpab017>
- Fitch WM (1970) Distinguishing homologous from analogous proteins. *Systematic Zoology* 19: 99–113. <https://doi.org/10.2307/2412448>
- Fuchs J, Ericson PGP, Bonillo C, Couloux A, Pasquet E (2015) The complex phylogeography of the Indo-Malayan *Alophoixus* bulbuls with the description of a putative new ring species complex. *Molecular Ecology* 24: 5460–5474. <https://doi.org/10.1111/mec.13337>
- Funk DJ, Omland KE (2003) Species-Level Paraphyly and Polyphyly: Frequency, Causes, and Consequences, with Insights from Animal Mitochondrial DNA. *Annual Review of Ecology, Evolution, and Systematics* 34: 397–423. <https://doi.org/10.1146/annurev.ecolsys.34.011802.132421>
- Futuyma DJ (2003) *Evolution*. Sinauer Associates, Inc., Sunderland, MA, 603 pp.
- Gautier M, Gharbi K, Cezard T, Foucaud J, Kerdelhué C, Pudlo P, Cornuet JM, Estoup A (2013) The effect of RAD allele dropout on the estimation of genetic variation within and between populations. *Molecular Ecology* 22: 3165–3178. <https://doi.org/10.1111/mec.12089>
- Gemmell AP, Marcus JM (2015) A tale of two haplotype groups: Evaluating the New World *Junonia* ring species hypothesis using the distribution of divergent COI haplotypes. *Systematic Entomology* 40: 532–546. <https://doi.org/10.1111/syen.12120.A>
- Haldane JBS (1932) *The Causes of Evolution*. Longmans, Green and Co., London, 222 pp.
- Harding RM, Fullerton SM, Griffiths RC, Bond J, Cox MJ, Schneider JA, Moulin DS, Clegg JB (1997) Archaic African and Asian lineages in the genetic ancestry of modern humans. *American Journal of Human Genetics* 60: 772–789.
- Harvey MG, Smith BT, Glenn TC, Faircloth BC, Brumfield RT (2016) Sequence Capture versus Restriction Site Associated DNA Sequencing for Shallow Systematics. *Systematic Biology* 65: 910–924. <https://doi.org/10.1093/sysbio/syw036>
- Hedrick PW (2009) *Genetics of Populations*, 4th ed. Jones & Bartlett Publishers, Boston, MA, 675 pp.
- Hennig W (1965) Phylogenetic Systematics. *Annual Review of Entomology* 10: 97–116.

- Hennig W (1966) *Phylogenetic Systematics*. University of Illinois Press, Chicago, IL, 263 pp.
- Hernández AI, Landis JB, Specht CD (2022) Phylogeography and population genetics reveal ring species patterns in a highly polymorphic California lily. *Journal of Biogeography*: 1–15. <https://doi.org/10.1111/jbi.14313>
- Highton R (1998) Is *Ensatina eschscholtzii* a ring-species? *Herpetologica* 54: 254–278.
- Hijmans RJ, van Etten J (2012) raster: Geographic Data Analysis and Modeling. Available from: <http://cran.r-project.org/package=raster>.
- Hohenlohe PA, Catchen J, Cresko WA (2012) Population Genomic Analysis of Model and Nonmodel Organisms Using Sequenced RAD Tags. In: *Methods in Molecular Biology*. Humana Press, Totowa, NJ, 235–260.
- Horvath DP, Patel S, Dođramaci M, Chao WS, Anderson J V., Foley ME, Scheffler B, Lazo G, Dorn K, Yan C, Childers A, Schatz M, Marcus S (2018) Gene Space and Transcriptome Assemblies of Leafy Spurge (*Euphorbia esula*) Identify Promoter Sequences, Repetitive Elements, High-Quality Markers, and a Full-Length Chloroplast Genome. *Weed Science* 66: 355–367. <https://doi.org/10.1017/wsc.2018.2>
- Ilut DC, Nydam ML, Hare MP (2014) Defining loci in restriction-based reduced representation genomic data from nonmodel species: Sources of bias and diagnostics for optimal clustering. *BioMed Research International* 2014. <https://doi.org/10.1155/2014/675158>
- Irwin DE (2000) Song variation in an avian ring species. *Evolution* 54: 998–1010. <https://doi.org/10.1111/j.0014-3820.2000.tb00099.x>
- Irwin DE, Irwin JH, Price TD (2001a) Ring species as bridges between microevolution and speciation. *Genetica* 112–113: 223–243. <https://doi.org/10.1023/A:1013319217703>
- Irwin DE, Bensch S, Price TD (2001b) Speciation in a ring. *Nature* 409: 333–337. <https://doi.org/10.1038/35053059>
- Irwin DE, Bensch S, Irwin JH, Price TD (2005) Speciation by distance in a ring species. *Science* 307: 414–416. <https://doi.org/10.1126/science.1105201>
- Irwin DE, Alcaide M, Delmore KE, Irwin JH, Owens GL (2016) Recurrent selection explains parallel evolution of genomic regions of high relative but low absolute differentiation in a ring species. *Molecular ecology* 25: 4488–4507. <https://doi.org/10.1111/mec.13792>
- Iturralde-Vinent M, MacPhee RD (1999) Paleogeography of the Caribbean region: implications for Cenozoic biogeography. *Bulletin of the American Museum of Natural History* No. 238.
- Iturralde-Vinent MA (2006) Meso-Cenozoic Caribbean paleogeography: Implications for the historical biogeography of the region. *International Geology Review* 48: 791–827. <https://doi.org/10.2747/0020-6814.48.9.791>
- Iturralde-Vinent MA La Paleogeografía del Caribe y sus implicaciones para la biogeografía histórica. *Revista del Jardín Botánico Nacional* 25–26: 49–78.
- Jacobo-Arteaga LE, Medina-Rodríguez MD, Hernández-Hernández B, Piña de la Rosa IA, Cacho. NI Leaf morphospace in *Euphorbia tithymaloides* (Euphorbiaceae) was likely shaped by evolutionary contingencies rather than climate. *Plant Ecology and Evolution*.
- Jordon-Thaden IE, Beck JB, Rushworth CA, Windham MD, Diaz N, Cantley JT, Martine CT, Rothfels CJ (2020) A basic ddRADseq two-enzyme protocol performs well with herbarium and silica-dried tissues across four genera. *Applications in Plant Sciences* 8: 1–7. <https://doi.org/10.1002/aps3.11344>
- Joseph L, Dolman G, Donnellan S, Saint KM, Berg ML, Bennett ATD (2008) Where and when does a ring start and end? Testing the ring-species hypothesis in a species complex of Australian parrots. *Proceedings of the Royal Society B: Biological Sciences* 275: 2431–2440. <https://doi.org/10.1098/rspb.2008.0765>
- Kapli P, Yang Z, Telford MJ (2020) Phylogenetic tree building in the genomic age. *Nature Reviews Genetics* 21: 428–444. <https://doi.org/10.1038/s41576-020-0233-0>
- Kuchta SR, Wake DB (2016) Wherefore and Whither the Ring Species? *Copeia* 104: 189–201. <https://doi.org/10.1643/OT-14-176>
- Kuchta SR, Parks DS, Mueller RL, Wake DB (2009) Closing the ring: Historical biogeography of the salamander ring species *Ensatina eschscholtzii*. *Journal of Biogeography* 36: 982–995. <https://doi.org/10.1111/j.1365-2699.2008.02052.x>
- Lack D (1973) The Numbers of Species of Hummingbirds in the West Indies. *Evolution* 27: 326. <https://doi.org/10.2307/2406972>
- Larget B, Kotha S, Dewey C, Ané C (2010) BUCKy: Gene Tree/Species Tree Reconciliation with Bayesian Concordance Analysis. *Bioinformatics (Oxford, England)* 26: 2910-1. [10.1093/bioinformatics/btq539](https://doi.org/10.1093/bioinformatics/btq539).
- Leitch IJ, Johnston E, Pellicer J, Hidalgo O, Bennett MD (2019) Angiosperm DNA C-values database (release 9.0,

- Apr 2019). Available from: <https://cvalues.science.kew.org/>.
- Liebers D, De Knijff P, Helbig AJ (2004) The herring gull complex is not a ring species. *Proceedings of the Royal Society B: Biological Sciences* 271: 893–901. <https://doi.org/10.1098/rspb.2004.2679>
- Linder C, Moret B, Nakhleh L, Warnow T (2004). *Network (Reticulate) Evolution: Biology, Models, and Algorithms*. The Pacific Symposium on Biocomputing.
- Liu L (2008) BEST: Bayesian estimation of species trees under the coalescent model. *Bioinformatics (Oxford, England)* 24: 2542–3. [10.1093/bioinformatics/btn484](https://doi.org/10.1093/bioinformatics/btn484).
- Lomelí-Sención JA (2006) Revisión de *Euphorbia* subgénero *Pedilanthus*. Universidad de Guadalajara
- Martin M, Patterson M, Garg S, Fischer SO, Pisanti N, Gunnar W, Marschall T (2016) WhatsHap : fast and accurate read-based phasing. : 1–18.
- Martin SH, Dasmahapatra KK, Nadeau NJ, Salazar C, Walters JR, Simpson F, Blaxter M, Manica A, Mallet J, Jiggins CD (2013) Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies. *Genome Research* 23: 1817–1828. <https://doi.org/10.1101/gr.159426.113>
- Martínez-Gordillo M, Jiménez-Ramírez J, Cruz-Durán R, Juárez-Arriaga E, García R, Cervantes A, Mejía-Hernández R (2002) Los géneros de la familia Euphorbiaceae en México. *Anales del Instituto de Biología, Universidad Nacional Autónoma de México, Serie Botánica* 73: 155–281.
- Martins AB, De Aguiar MAM, Bar-Yam Y (2013) Evolution and stability of ring species. *Proceedings of the National Academy of Sciences of the United States of America* 110: 5080–5084. <https://doi.org/10.1073/pnas.1217034110>
- Mastretta-Yanes A, Arrigo N, Alvarez N, Jorgensen TH, Piñero D, Emerson BC (2015) Restriction site-associated DNA sequencing, genotyping error estimation and de novo assembly optimization for population genetic inference. *Molecular Ecology Resources* 15: 28–41. <https://doi.org/10.1111/1755-0998.12291>
- Maynard Smith J (1966) Sympatric Speciation. *The American naturalist* 100: 637–650.
- Mayr E (1940) Speciation Phenomena in Birds. *The American Naturalist* 74: 249–278. Available from: <https://www.jstor.org/stable/2457576>.
- Mayr E (1942) *Systematics and the origin of species from the viewpoint of a zoologist*. Columbia University Press, New York, USA, 354 pp.
- Mayr E (1963) *Animal Species and Evolution*. Harvard University Press, Cambridge, MA, 797 pp.
- Mayr E (1970) *Populations, Species and Evolution: An Abridgment of Animal Species and Evolution*. Belknap Press of Harvard University Press, Cambridge, MA, 453 pp.
- Mboumba GB, Ward D (2008) Phenotypic plasticity and local adaptation in two extreme populations of *Acacia karroo*. *African Journal of Range and Forage Science* 25: 121–130. <https://doi.org/10.2989/AJRF.2008.25.3.4.601>
- McCartney-Melstad E, Gidiş M, Shaffer HB (2019) An empirical pipeline for choosing the optimal clustering threshold in RADseq studies. *Molecular Ecology Resources* 19: 1195–1204. <https://doi.org/10.1111/1755-0998.13029>
- McCormack JE, Hird SM, Zellmer AJ, Carstens BC, Brumfield RT (2013) Applications of next-generation sequencing to phylogeography and phylogenetics. *Molecular Phylogenetics and Evolution* 66: 526–538. <https://doi.org/10.1016/j.ympev.2011.12.007>
- McGuire JA, Witt CC, Remsen J V., Corl A, Rabosky DL, Altshuler DL, Dudley R (2014) Molecular phylogenetics and the diversification of hummingbirds. *Current Biology* 24: 910–916. <https://doi.org/10.1016/j.cub.2014.03.016>
- Miller MR, Dunham JP, Amores A, Cresko WA, Johnson EA (2007) Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Research* 17: 240–248. <https://doi.org/10.1101/gr.5681207>
- Monteverde-Suárez MJ (2021) La especie-anillo *Euphorbia tithymaloides* evaluada con base en evidencia genética, ecológica, morfológica y biogeográfica. Universidad Nacional Autónoma de México
- Moritz C, Bi K (2011) Spontaneous speciation by ploidy elevation: Laboratory synthesis of a new clonal vertebrate. *Proceedings of the National Academy of Sciences of the United States of America* 108: 9733–9734. <https://doi.org/10.1073/pnas.1106455108>
- Nazareno A, Bemmels J, Dick C, Lohmann L (2017) Minimum sample sizes for population genomics: an empirical study from an Amazonian plant species. *Molecular Ecology Resources* 17: 1136–1147.
- Nguyen DD, Oguri E, Yamada A, Lin CC, Zhilin C, Nguyen AD, Suyama Y, Eguchi K (2020) Genome-wide mig-seq and morphometric data reveals heterospecificity of the *gnamptogenys taivanensis* group (Hymenoptera:

- Formicidae: Ectatomminae) in the northern mountainous region of vietnam. *Raffles Bulletin of Zoology* 68: 539–555. <https://doi.org/10.26107/RBZ-2020-0070>
- Nguyen LT, Schmidt HA, Von Haeseler A, Minh BQ (2015) IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution* 32: 268–274. <https://doi.org/10.1093/molbev/msu300>
- Niemiller ML, Fitzpatrick BM, Miller BT (2008) Recent divergence with gene flow in Tennessee cave salamanders (Plethodontidae: Gyrinophilus) inferred from gene genealogies. *Molecular Ecology* 17: 2258–2275. <https://doi.org/10.1111/j.1365-294X.2008.03750.x>
- Nosil P (2012) What is ecological speciation? In: *Ecological Speciation*. Oxford University Press, 4–21.
- Novembre J, Stephens M (2008) Interpreting principal component analyses of spatial population genetic variation. *Nature Genetics* 40: 646–649. <https://doi.org/10.1038/ng.139>
- Papadopoulos AST, Baker WJ, Crayn D, Butlin RK, Kynast RG, Hutton I, Savolainen V (2011) Speciation with gene flow on Lord Howe Island. *Proceedings of the National Academy of Sciences of the United States of America* 108: 13188–13193. <https://doi.org/10.1073/pnas.1106085108>
- Patten MA, Pruett CL (2009) The Song Sparrow, *Melospiza melodia*, as a ring species: Patterns of geographic variation, a revision of subspecies, and implications for speciation. *Systematics and Biodiversity* 7: 33–62. <https://doi.org/10.1017/S1477200008002867>
- Pearson KFRS (1901) On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 2: 559–572. Available from: <https://doi.org/10.1080/14786440109462720>.
- Pickrell JK, Pritchard JK (2012) Inference of Population Splits and Mixtures from Genome-Wide Allele Frequency Data. *PLoS Genetics* 8. <https://doi.org/10.1371/journal.pgen.1002967>
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of Population Structure Using Multilocus Genotype Data. *Genetics* 155: 945–959.
- Pulich WM (1968) The Occurrence of the Crested Hummingbird, *Orthorhynchus cristatus exilis*, in the United States. *The Auk* 85: 322. <https://doi.org/https://doi.org/10.2307/4083596>
- Qi Y, Ding L, Zhao Y, Niu C, Wang Ex, Zhao W (2020) Toad-headed lizard *phrynocephalus forsythii* (Squamata, Agamidae) as a potential ring species inferred from population genetic differentiation. *Asian Herpetological Research* 11: 312–319. <https://doi.org/10.16373/j.cnki.ahr.200001>
- de Queiroz K (1998) The General Lineage Concept of Species, Species Criteria, and the Process of Speciation. In: *Endless Forms: Species and Speciation*. Oxford University Press, New York City, NY, 57–75.
- de Queiroz K (1999) The General Lineage Concept of Species and the Defining Properties of the Species Category. In: *Species: New Interdisciplinary Essays*. The MIT Press, Cambridge, MA, 49–89.
- De Queiroz K (2005) Ernst Mayr and the modern concept of species. *Proceedings of the National Academy of Sciences of the United States of America* 102: 6600–6607. <https://doi.org/10.17226/11310>
- R Core Team (2020) R: A language and environment for statistical computing. Available from: <https://www.r-project.org/>.
- Raffaele H, Wiley J, Garrido O, Keith A, Raffaele J (2003) *Birds of the West Indies*. Princeton University Press, Princeton, NJ, 216 pp.
- Rahman AYA, Usharraj AO, Misra BB, Thottathil GP, Jayasekaran K, Feng Y, Hou S, Ong SY, Ng FL, Lee LS, Tan HS, Sakaff MKLM, Teh BS, Khoo BF, Badai SS, Aziz NA, Yuryev A, Knudsen B, Dionne-Laporte A, Mchunu NP, Yu Q, Langston BJ, Freitas TAK, Young AG, Chen R, Wang L, Najimudin N, Saito JA, Alam M (2013) Draft genome sequence of the rubber tree *Hevea brasiliensis*. *BMC Genomics* 14. <https://doi.org/10.1186/1471-2164-14-75>
- Reich D, Price AL, Patterson N (2008) Principal component analysis of genetic data. *Nature Genetics* 40: 491–492. <https://doi.org/10.1038/ng0508-491>
- Rice WR, Hostert EE (1993) Laboratory experiments on speciation: what have we learned in 40 years? *Evolution* 47: 1637–1653. <https://doi.org/10.1038/161162d0>
- Riina R, Peirson JA, Geltman D V., Molero J, Frajman B, Pahlevani A, Barres L, Morawetz JJ, Salmaki Y, Zarre S, Kryukov A, Bruyns P V., Berry PE (2013) A worldwide molecular phylogeny and classification of the leafy spurge, *Euphorbia* subgenus *Esula* (Euphorbiaceae). *Taxon* 62: 316–342. <https://doi.org/10.12705/622.3>
- Rochette NC, Catchen JM (2017) Deriving genotypes from RAD-seq short-read data using Stacks. *Nature Protocols* 12: 2640–2659. <https://doi.org/10.1038/nprot.2017.123>
- Roncal J, Nieto-Blázquez ME, Cardona A, Bacon CD (2020) Historical Biogeography of Caribbean Plants Revises

- Regional Paleogeography. In: Rull V, Carnaval AC (Eds), Neotropical Diversification: Patterns and Processes. Springer Nature Switzerland, 521–546. https://doi.org/10.1007/978-3-030-31167-4_20
- Rubin BER, Ree RH, Moreau CS (2012) Inferring phylogenies from RAD sequence data. *PLoS ONE* 7: 1–12. <https://doi.org/10.1371/journal.pone.0033394>
- Sabeti PC, Varilly P, Fry B, Lohmueller J, Hostetter E, Cotsapas C, Xie X, Byrne EH, McCarroll SA, Gaudet R, Schaffner SF, Lander ES, Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA, Belmont JW, Boudreau A, Hardenbol P, Leal SM, Pasternak S, Wheeler DA, Willis TD, Yu F, Yang H, Zeng C, Gao Y, Hu H, Hu W, Li C, Lin W, Liu S, Pan H, Tang X, Wang J, Wang W, Yu J, Zhang B, Zhang Q, Zhao H, Zhao H, Zhou J, Gabriel SB, Barry R, Blumenstiel B, Camargo A, Defelice M, Faggart M, Goyette M, Gupta S, Moore J, Nguyen H, Onofrio RC, Parkin M, Roy J, Stahl E, Winchester E, Ziaugra L, Altshuler D, Shen Y, Yao Z, Huang W, Chu X, He Y, Jin L, Liu Y, Shen Y, Sun W, Wang H, Wang Y, Wang Y, Xiong X, Xu L, Wayne MMY, Tsui SKW, Xue H, Wong JTF, Galver LM, Fan JB, Gunderson K, Murray SS, Oliphant AR, Chee MS, Montpetit A, Chagnon F, Ferretti V, Leboeuf M, Olivier JF, Phillips MS, Roumy S, Sallée C, Verner A, Hudson TJ, Kwok PY, Cai D, Koboldt DC, Miller RD, Pawlikowska L, Taillon-Miller P, Xiao M, Tsui LC, Mak W, You QS, Tam PKH, Nakamura Y, Kawaguchi T, Kitamoto T, Morizono T, Nagashima A, Ohnishi Y, Sekine A, Tanaka T, Tsunoda T, Deloukas P, Bird CP, Delgado M, Dermitzakis ET, Gwilliam R, Hunt S, Morrison J, Powell D, Stranger BE, Whittaker P, Bentley DR, Daly MJ, De Bakker PIW, Barrett J, Chretien YR, Maller J, McCarroll S, Patterson N, Pe'er I, Price A, Purcell S, Richter DJ, Saxena R, Sham PC, Stein LD, Krishnan L, Smith AV, Tello-Ruiz MK, Thorisson GA, Chakravarti A, Chen PE, Cutler DJ, Kashuk CS, Lin S, Abecasis GR, Guan W, Li Y, Munro HM, Qin ZS, Thomas DJ, McVean G, Auton A, Bottolo L, Cardin N, Eyheramendy S, Freeman C, Marchini J, Myers S, Spencer C, Stephens M, Donnelly P, Cardon LR, Clarke G, Evans DM, Morris AP, Weir BS, Johnson TA, Mullikin JC, Sherry ST, Feolo M, Skol A, Zhang H, Matsuda I, Fukushima Y, MacEr DR, Suda E, Rotimi CN, Adebamowo CA, Ajayi I, Aniagwu T, Marshall PA, Nkwodimmah C, Royal CDM, Leppert MF, Dixon M, Peiffer A, Qiu R, Kent A, Kato K, Niikawa N, Adewole IF, Knoppers BM, Foster MW, Clayton EW, Watkin J, Muzny D, Nazareth L, Sodergren E, Weinstock GM, Yakub I, Birren BW, Wilson RK, Fulton LL, Rogers J, Burton J, Carter NP, Clee CM, Griffiths M, Jones MC, McLay K, Plumb RW, Ross MT, Sims SK, Willey DL, Chen Z, Han H, Kang L, Godbout M, Wallenburg JC, L'Archevêque P, Bellemare G, Saeki K, Wang H, An D, Fu H, Li Q, Wang Z, Wang R, Holden AL, Brooks LD, McEwen JE, Guyer MS, Wang VO, Peterson JL, Shi M, Spiegel J, Sung LM, Zacharia LF, Collins FS, Kennedy K, Jamieson R, Stewart J (2007) Genome-wide detection and characterization of positive selection in human populations. *Nature* 449: 913–918. <https://doi.org/10.1038/nature06250>
- Sahagún-Godínez E, Lomelí-Senci6n JA (1997) *Pedilanthus diazlananus* (Euphorbiaceae): Pollination by hymenopterans in a bird-pollinated genus. *American Journal of Botany* 84: 1584–1587. <https://doi.org/10.2307/2446620>
- Santiago-Valentin E, Olmstead RG (2004) Historical Biogeography of Caribbean Plants: Introduction to Current Knowledge and Possibilities from a Phylogenetic Perspective. *Taxon* 53: 299–319.
- Shaw KL, Mullen SP (2014) Speciation continuum. *Journal of Heredity* 105: 741–742. <https://doi.org/10.1093/jhered/esu060>
- Shead, T. M. 2014. toyplot. Technical Report TOYPLOT; 003233MLTPL00. Sandia National Laboratory. Sandia, NM, USA.
- Shepherdson JP (2018) Observations of territorial behavior of the Antillean Crested Hummingbird (*Orthorhynchus cristatus*) on St. Eustatius. *Journal of Caribbean Ornithology* 31: 48–50.
- Simpson GG (1961) Principles of animal taxonomy. Columbia University Press, New York City, NY, 247 pp.
- Smith TB, Wayne RK, Girman DJ, Bruford MW (1997) A Role for Ecotones in Generating Rainforest Biodiversity. *Science* 276: 1855–1857.
- Smyth JF, Patten MA, Pruett CL (2015) The evolutionary ecology of a species ring: a test of alternative models. *Folia Zoologica* 64: 233–244. <https://doi.org/10.25225/fozo.v64.i3.a5.2015>
- Soltani GG, B6non D, Alvarez N, Praz CJ (2017) When different contact zones tell different stories: Putative ring species in the *Megachile concinna* species complex (Hymenoptera: Megachilidae). *Biological Journal of the Linnean Society* 121: 815–832. <https://doi.org/10.1093/biolinnean/blx023>
- Stamatakis A (2014) RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30: 1312–1313. <https://doi.org/10.1093/bioinformatics/btu033>
- Stankowski S, Ravinet M (2021) Defining the speciation continuum. *Evolution* 75: 1256–1273. <https://doi.org/10.1111/evo.14215>

- Stebbins RC (1957) Intraspecific Sympatry in the Lungless Salamander *Ensatina eschscholtzi*. *Evolution* 11: 265–270.
- Steinmann VW (2002) Diversidad y endemismo de la familia Euphorbiaceae en México. *Acta Botanica Mexicana*: 61. <https://doi.org/10.21829/abm61.2002.909>
- Steinmann VW (2003) The submersion of *Pedilanthus* into *Euphorbia* (Euphorbiaceae). *Acta Botánica Mexicana*: 45–50.
- Stimson LA (1944) *Riccordia riccordii* Seen at Miami, Florida. *The Auk* 61: 300–301.
- Svensson EI, Abbott JK, Gosden TP, Coreau A (2009) Female polymorphisms, sexual conflict and limits to speciation processes in animals. *Evolutionary Ecology* 23: 93–108. <https://doi.org/10.1007/s10682-007-9208-2>
- Torres-Vanegas F, Hadley AS, Kormann UG, Jones FA, Betts MG, Wagner HH (2019) The Landscape Genetic Signature of Pollination by Trapliners: Evidence From the Tropical Herb, *Heliconia tortuosa*. *Frontiers in Genetics* 10: 1–12. <https://doi.org/10.3389/fgene.2019.01206>
- Tregenza T (2002) Divergence and reproductive isolation in the early stages of speciation. *Genetica* 116: 291–300. <https://doi.org/10.1023/A:1021257114996>
- Tsvuura Z, Ward D (2022) Does a reciprocal transplant experiment of neighboring *Vachellia karroo* populations demonstrate local adaptation? *South African Journal of Botany* 144: 316–324. <https://doi.org/10.1016/j.sajb.2021.09.012>
- Veiga Blanco T, Galetto L, Machado IC (2013) Nectar regulation in *Euphorbia tithymaloides* L., a hummingbird-pollinated Euphorbiaceae. *Plant Biology* 15: 910–918. <https://doi.org/10.1111/j.1438-8677.2012.00695.x>
- Volpe NL, Hadley AS, Robinson WD, Betts MG (2014) Functional connectivity experiments reflect routine movement behavior of a tropical hummingbird species. *Ecological Applications* 24: 2122–2131. <https://doi.org/10.1890/13-2168.1>
- Volz EM, Wiuf C, Grad YH, Frost SDW, Dennis AM, Didelot X (2020) Identification of Hidden Population Structure in Time-Scaled Phylogenies. *Systematic Biology* 69: 884–896. <https://doi.org/10.1093/sysbio/syaa009>
- Wadge G (1994) The Lesser Antilles. In: Donovan SK, Jackson TA (Eds), *Caribbean Geology: An Introduction*. The University of the West Indies Publishers' Association (UWIPA), Kingston, 167–177.
- Wake DB, Yanev KP (1986) Geographic variation in allozymes in a “ring species”, the plethodontid salamander *Ensatina eschscholtzii* of western North America. *Evolution* 40: 702–715. <https://doi.org/10.1111/j.1558-5646.1986.tb00532.x>
- Wake DB, Yanev KP, Brown CW (1986) Intraspecific Sympatry in a “Ring Species,” the Plethodontid Salamander *Ensatina eschscholtzii*, in Southern California. *Evolution* 40: 866–868.
- Wang W, Feng B, Xiao J, Xia Z, Zhou X, Li P, Zhang W, Wang Y, Møller BL, Zhang P, Luo MC, Xiao G, Liu J, Yang J, Chen S, Rabinowicz PD, Chen X, Zhang H Bin, Ceballos H, Lou Q, Zou M, Carvalho LJC, Zeng C, Xia J, Sun S, Fu Y, Wang H, Lu C, Ruan M, Zhou S, Wu Z, Liu H, Kannangara RM, Jørgensen K, Neale RL, Bonde M, Heinz N, Zhu W, Wang S, Zhang Y, Pan K, Wen M, Ma PA, Li Z, Hu M, Liao W, Hu W, Zhang S, Pei J, Guo A, Guo J, Zhang J, Zhang Z, Ye J, Ou W, Ma Y, Liu X, Tallon LJ, Galens K, Ott S, Huang J, Xue J, An F, Yao Q, Lu X, Fregene M, López-Lavalle LAB, Wu J, You FM, Chen M, Hu S, Wu G, Zhong S, Ling P, Chen Y, Wang Q, Liu G, Liu B, Li K, Peng M (2014) Cassava genome from a wild ancestor to cultivated varieties. *Nature Communications* 5. <https://doi.org/10.1038/ncomms6110>
- Wang X, Ye X, Zhao L, Li D, Guo Z, Zhuang H (2017) Genome-wide RAD sequencing data provide unprecedented resolution of the phylogeny of temperate bamboos (Poaceae: Bambusoideae). *Scientific Reports* 7: 1–11. <https://doi.org/10.1038/s41598-017-11367-x>
- Ward D (2011) Population differentiation in a purported ring species, *Acacia karroo* (Mimosoideae). *Biological Journal of the Linnean Society* 104: 748–755. <https://doi.org/10.1111/j.1095-8312.2011.01757.x>
- Webster GL (1994) Synopsis of the Genera and Suprageneric Taxa of Euphorbiaceae. *Annals of the Missouri Botanical Garden*: 81: 33–144.
- Wen G, Fu J (2021) Isolation and re-connection: the formation of a ring-shaped speciation continuum in an odorous frog (*Odorrana margaretae*). *Authorea*: 1–40.
- Wenger AM, Peluso P, Rowell WJ, Chang PC, Hall RJ, Concepcion GT, Ebler J, Functammasan A, Kolesnikov A, Olson ND, Töpfer A, Alonge M, Mahmoud M, Qian Y, Chin CS, Phillippy AM, Schatz MC, Myers G, DePristo MA, Ruan J, Marschall T, Sedlazeck FJ, Zook JM, Li H, Koren S, Carroll A, Rank DR, Hunkapiller MW (2019) Accurate circular consensus long-read sequencing improves variant detection and assembly of a human genome. *Nature Biotechnology* 37: 1155–1162. <https://doi.org/10.1038/s41587-019-0217-9>

- Willing EM, Dreyer C, van Oosterhout C (2012) Estimates of genetic differentiation measured by *f_{st}* do not necessarily require large sample sizes when using many snp markers. *PLoS ONE* 7: 1–7. <https://doi.org/10.1371/journal.pone.0042649>
- Wright S (1931) Evolution in mendelian populations. *Genetics* 16: 97–159.
- Wright S (1940) Breeding Structure of Populations in Relation to Speciation. *The American Naturalist* 74: 232–248. <https://doi.org/10.1086/280891>
- Wright S (1951) The genetical structure of populations. *Annals of Eugenics* 15: 323–354. <https://doi.org/10.1111/j.1469-1809.1949.tb02451.x>.
- Wright S (1978) *Evolution and the Genetics of Populations, Volume 4: Variability within and among natural populations*. University of Chicago Press, Chicago, IL, 590 pp.
- Wu P, Zhou C, Cheng S, Wu Z, Lu W, Han J, Chen Y, Ni P, Wang Y, Xu X, Huang Y, Song C, Wang Z, Shi N, Zhang X, Fang X, Yang Q, Jiang H, Chen Y, Li M, Chen F, Wang J, Wu G (2015) Integrated genome sequence and linkage map of physic nut (*Jatropha curcas* L.), a biodiesel plant. *Plant Journal* 81: 810–821. <https://doi.org/10.1111/tpj.12761>
- Wurdack KJ, Hoffmann P, Chase MW (2005) Molecular phylogenetic analysis of uniovulate Euphorbiaceae (Euphorbiaceae sensu stricto) using plastid *rbcl* and *trnL-F* DNA sequences. *American Journal of Botany* 92: 1397–1420. <https://doi.org/10.3732/ajb.92.8.1397>
- Yunusbaev U, Valeev A, Yunusbaeva M, Kwon HW, Mägi R, Metspalu M, Yunusbayev B (2019) Reconstructing recent population history while mapping rare variants using haplotypes. *Scientific Reports* 9: 1–9. <https://doi.org/10.1038/s41598-019-42385-6>
- Zhang C, Scornavacca C, Molloy EK, Mirarab S (2020) ASTRAL-pro: Quartet-based species-tree inference despite paralogy. *Molecular Biology and Evolution* 37: 3292–3307. <https://doi.org/10.1093/molbev/msaa139>

Anexo I. Algunos conceptos, métodos y parámetros utilizados en este trabajo

Especies-anillo y el concepto de especie

El presente trabajo pretende evaluar evidencia que apoye la existencia de un patrón histórico de divergencia a través de dos frentes geográficos que dan lugar a dos “formas” (en este caso concreto, formas terminales de una cadena de poblaciones) que se comportan como especies diferentes, por lo que se adopta la definición de especie propuesta por Mayr (1940): “una especie se compone de un grupo de poblaciones [...] en el que las poblaciones adyacentes hibridan cuando están en contacto o son potencialmente capaces de hacerlo (con una o más poblaciones) en aquellos casos en los que el contacto es prevenido por barreras ecológicas o geográficas”. Asimismo, se utiliza su definición de subespecie: “la subespecie, o raza geográfica, es una subdivisión geográficamente localizada de la especie, que difiere genética y taxonómicamente de otras subdivisiones de la especie” (Mayr, 1940). Históricamente a este concepto de especie se le ha llamado el concepto biológico de especie o el concepto de entrecruzamiento [*inbreeding*] de especie, aunque también puede ser referido como el concepto de poza génica de especie o el concepto sexual de especie (de Queiroz, 1998). El criterio que va de la mano de este concepto es la capacidad potencial de entrecruzamiento entre los organismos de poblaciones diferentes (de Queiroz, 1998). Es importante resaltar que Mayr (1940) no consideraba que los sistemas de especie-anillo (o como él los llamaba, traslapes circulares) se adecuaban al concepto biológico de especie, sino que más bien representaban una excepción, pues ejemplificaban sistemas en el punto transitorio entre subespecies y especies. Aun así, uno de los criterios que se utilizan para identificar estos sistemas es el criterio de entrecruzamiento entre poblaciones terminales, por lo que es posible considerar que se relaciona más con el concepto biológico que con otros conceptos de especie.

El concepto biológico de especie de Mayr (1940) considera que el flujo génico solamente debe existir a nivel intraespecífico y que entre dos especies diferentes no puede haber flujo génico. Sin embargo, otros autores, entre ellos Dobzhansky (1951) y Wright (1978), han sugerido una definición menos estricta de especie en la que sí puede existir cierto nivel de flujo génico entre especies diferentes. Para

Coyne y Orr (2004), las especies se caracterizan por un aislamiento reproductivo sustancial, pero no necesariamente completo. Según estos autores, la hibridación entre dos formas consideradas especies no es poco común, sino que suele suceder con relativa frecuencia. Por otro lado, Coyne y Orr (2004) consideran que el concepto biológico de especie tiene diversas ventajas sobre otros conceptos de especie. Quizás entre las ventajas más importantes del concepto biológico de especie es que se puede abordar desde una perspectiva teórico-práctica, teniendo un marco de referencia (el aislamiento reproductivo) relativamente fácil de detectar y comprobar, y sobre todo, que es cuantificable (medible). A partir de este concepto de especie es posible desarrollar un programa de investigación que permita evaluar las causas y los mecanismos que operan en la evolución del aislamiento reproductivo y en última instancia, cómo sucede la especiación. Otros conceptos de especie suelen ser difíciles de contrastar de manera práctica o cuantificable, y en su mayoría sólo aventajan al concepto biológico en que sí se pueden aplicar a especies que son capaces de reproducirse asexualmente.

Para de Queiroz (1998, 1999) todos los conceptos de especie se pueden considerar únicamente como diferentes criterios para demarcar a las especies o variantes de un único concepto

de especie al que el autor llama “concepto general de linajes” [*the general lineage concept of species*]. Bajo este concepto las especies se pueden definir como segmentos de linajes de poblaciones, definiendo a los linajes como una serie de entidades formando una sola línea de ancestría y descendencia directa. Este concepto de especie es similar al concepto evolutivo de especie, según el cual una especie es un linaje (una secuencia ancestro-descendiente de poblaciones) que evoluciona de forma separada de otros con sus propios roles y tendencias evolutivas (Simpson, 1961). Para Simpson (1961) la monofilia es la derivación de un taxón a través de uno o más linajes (sucesiones temporales de ancestro-descendencia de poblaciones) a partir de un taxón ancestral inmediato. Por su parte, de Queiroz (1998, 1999) enfatiza que dentro de su concepto de especie (de linajes) una especie no necesariamente debe ser monofilética sino que puede ser parafilética. De igual forma, Hennig (1966) considera que las especies no necesariamente deben ser monofiléticas, como es el caso de aquellas que se originan por hibridación.

Filogenética, filogeografía y genética de poblaciones: enfoques y limitaciones

De forma amplia, la filogenética, la filogeografía y la genética de poblaciones buscan entender y caracterizar el cambio evolutivo y la divergencia de los genomas y de los organismos que los portan, y difieren principalmente en el énfasis de sus enfoques y herramientas (Cutter, 2013). También es importante notar que para cada una de estas disciplinas, en mayor o menor medida, es importante el nivel de flujo génico que existe entre las poblaciones (o unidades operacionales) y la diferenciación (o falta de diferenciación) genética que produce.

Filogenética

La sistemática filogenética investiga las relaciones de parentesco evolutivo entre las especies (Hennig, 1965). La definición del concepto de “relación filogenética”, según Hennig (1965), se basa en la reproducción “bisexual” de los organismos y la idea de que esta reproducción sólo puede llevarse a cabo en el marco de comunidades reproductivamente confinadas que están aisladas genéticamente unas de otras. Como las relaciones filogenéticas entre los taxa dependen en gran medida de la presencia y/o ausencia de flujo génico se han utilizado filogenias o métodos filogenéticos a nivel infraespecífico con éxito para inferir estructura genética poblacional (p. ej. Pickrell y Pritchard, 2012; Volz et al., 2020). Estos enfoques han sido utilizados para estudiar especies-anillo y otros sistemas en los que existe diversidad infraespecífica (p. ej. Kuchta et al., 2009; Fuchs et al., 2015).

Las relaciones filogenéticas son representadas de manera gráfica a través de lo que conocemos como árbol filogenético (Hennig, 1965). Los árboles filogenéticos tienen que ser inferidos, para lo cual se utilizan métodos que se pueden dividir en métodos basados en distancia y métodos basados en caracteres, los cuales pueden ser paramétricos (Máxima Verosimilitud e Inferencia Bayesiana) o no paramétricos (Máxima Parsimonia). El método de máxima verosimilitud fue propuesto y desarrollado por R. A. Fisher entre los años 1912 y 1922 (Aldrich, 1997). Para Fisher (1922), la verosimilitud de que cualquier parámetro tenga cualquier valor asignado es proporcional a la probabilidad de que si esto fuera cierto, los datos serían los observados. Posteriormente, Felsenstein (1981) aplicó este método estadístico para inferir árboles filogenéticos, definiendo a la estimación de máxima verosimilitud como la búsqueda del “árbol evolutivo” que tenga la mayor probabilidad de haber producido los datos que observamos. Algunos de los programas más populares en la actualidad para realizar análisis filogenéticos de máxima verosimilitud son RAXML (*Randomized Accelerated Maximum Likelihood*; Stamatakis, 2014) e IQ-TREE (Nguyen et al., 2015), que destacan

por sus algoritmos de búsqueda rápida y eficacia al proporcionar árboles con altos valores de verosimilitud.

En estudios sobre especies-anillo, los métodos filogenéticos han sido ampliamente utilizados pues estimar las relaciones de parentesco de los individuos que conforman el anillo es una manera eficiente e informativa de evaluar estos sistemas desde un punto de vista evolutivo (p. ej.: Irwin et al., 2001; Kuchta et al., 2009; Cacho y Baum, 2012; Fuchs et al., 2015; Soltani et al., 2017; Qi et al., 2020).

Genética de poblaciones

La disciplina conocida como genética de poblaciones nació a partir de los escritos de Fisher (1930), Wright (1931) y Haldane (1932), los cuales buscaban un acercamiento a los procesos evolutivos descritos por Darwin y Wallace desde la perspectiva de la genética mendeliana redescubierta a principios del siglo XX. La genética de poblaciones se puede definir, en su forma más general, como la rama de la genética que se encarga de describir los rasgos heredables observables o inferidos en las poblaciones en el espacio y el tiempo (Amorim, 2013). Dentro de este campo de conocimiento se definen a las poblaciones como un grupo de individuos que coexisten en tiempo y espacio y que son capaces de entrecruzamiento (Hedrick, 2009).

Desde los inicios de esta disciplina uno de sus objetivos más importantes ha sido el estudio de la estructura genética de las poblaciones, definida como el número de subpoblaciones que contiene, las frecuencias genéticas de dichas subpoblaciones y el grado de aislamiento genético que tiene cada una de ellas (Chakraborty, 1993). Si los individuos de una población no se reproducen al azar, entonces se puede decir que esa población posee cierto grado de estructura genética poblacional (Wright, 1951). Asimismo, se dice que una población está sub-estructurada cuando sus componentes (subpoblaciones) tienen un nivel de aislamiento reproductivo y, por lo tanto, el flujo génico entre ellas está de cierta manera restringido (Chakraborty, 1993). De esto se puede deducir que el concepto de estructura poblacional va de la mano con el concepto de flujo genético.

Además de los estudios de estructura poblacional, la genética de poblaciones permite estudiar y cuantificar el nivel de diferenciación entre poblaciones, así como estimar niveles de flujo génico entre las mismas. Estos enfoques han sido utilizados para estudiar sistemas como las especies-anillo, donde hay divergencia incipiente y una estructura poblacional marcada (Kuchta et al., 2009; Alcaide et al., 2014). Recientemente, diversos estudios de genética de poblaciones realizados tanto con simulaciones como con datos empíricos han mostrado que el uso de datos genómicos para un número reducido de individuos por muestra poblacional es suficiente para obtener resultados congruentes de variabilidad y diferenciación genética. Esto se logra siempre y cuando se utilice una cantidad suficientemente representativa de SNPs y poblaciones muestreadas (Willing et al., 2012; Nazareno et al., 2017; Aguirre-Liguori et al., 2020).

La genética de poblaciones se ha utilizado para evaluar otros sistemas de especie-anillo putativos. En este tipo de casos es frecuente estimar índices de diversidad genética (como la heterocigosis observada y esperada) así como calcular el índice de fijación o F_{ST} pareada [*pairwise*] con el propósito de evaluar la diferenciación que existe entre las poblaciones (p. ej.: Joseph et al., 2008; Qi et al., 2020; Wen y Fu, 2021). El cálculo de la diferenciación entre poblaciones puede ser una herramienta importante para determinar el grado de divergencia entre poblaciones o bien si existe o no flujo génico entre las formas del anillo, en particular las formas terminales. En los estudios sobre especies-anillo también es usual visualizar los datos con análisis de componentes principales (PCA; p. ej.: Alcaide et al., 2014; Fuchs et al., 2015; Wen y Fu, 2021) y realizar análisis de estructura genética

usando métodos de agrupamiento bayesiano como el usado por el programa STRUCTURE (Pritchard et al., 2000).

STRUCTURE es un programa que utiliza un método de agrupamiento Bayesiano para poder asignar individuos a poblaciones a partir de sus genotipos y calcular frecuencias alélicas de las poblaciones, todo eso bajo la asunción de que existe un número K de poblaciones, que puede ser desconocido (Pritchard et al., 2000). En el caso de que se desconozca el número de *clusters* o grupos (K) se puede realizar una prueba de Evanno para el cálculo del estadístico ΔK , el cual permite estimar el número más adecuado dado un set de datos (Evanno et al., 2005).

Filogeografía

La filogeografía es un término utilizado por primera vez por Avise y colaboradores (1987), originalmente denominado “filogeografía intraespecífica” y definido como la disciplina que estudia conjuntamente las relaciones filogenéticas de las moléculas de ADN (originalmente de ADN mitocondrial) y la distribución geográfica de los agrupamientos filogenéticos. Esta definición fue ampliada y desarrollada más extensivamente por el mismo Avise (2000): “La filogeografía es el campo de estudio que se interesa por los procesos que gobiernan las distribuciones geográficas de los linajes genealógicos, [...] el tiempo y el espacio son los ejes conjuntamente considerados de la filogeografía sobre los que (idealmente) se mapean genealogías de genes particulares de interés”. En teoría, y desde el punto de vista más simplista, la distribución espacial de los linajes debe mostrar patrones similares a su variación a lo largo del tiempo. Es decir, las secuencias de linajes más alejados espacialmente deben ser más diferentes, así como las secuencias de linajes más separados temporalmente (Hedrick, 2009). Para Avise y colaboradores (1987) la filogeografía intraespecífica se fundamenta en tres hipótesis principales: a) la mayoría de las especies se componen de poblaciones geográficas cuyos miembros ocupan ramas distintas de un árbol filogenético intraespecífico; b) las especies con poca estructura filogeográfica poblacional se dispersan más y ocupan grandes distribuciones con pocos impedimentos para el flujo génico; c) los grupos monofiléticos son resultado de barreras que impiden el flujo génico a largo plazo.

Para Avise (2000) la filogeografía busca servir como un puente teórico y analítico entre las disciplinas que se encargan de estudiar la microevolución (como la genética de poblaciones) y aquellas que estudian los procesos macroevolutivos (como la filogenética). Se sugiere que son los niveles de flujo génico poblacional los que determinan la estructura o patrones filogeográficos, de modo que secuencias más alejadas (filogenéticamente) o con niveles de flujo génico bajos o nulos tenderán a ser más distintas y distantes. De esta manera, la filogeografía integra aspectos del marco teórico y analítico de la genética de poblaciones para explicar la estructura filogeográfica y la divergencia filogenética.

En diversos estudios sobre sistemas de especies-anillo se han utilizado aproximaciones y métodos filogeográficos. Uno de los métodos más usados para investigar si los sistemas evaluados presentan patrones consistentes con el modelo de especie-anillo ha sido el análisis de redes de haplotipos con ADN mitocondrial (p. ej.: Bowen et al., 2001; Liebers et al., 2004; Joseph et al., 2008; Fuchs et al., 2015; Gemmell y Marcus, 2015).

Datos de RADseq

RADseq (*Restriction site Associated DNA sequencing*) es un tipo de secuenciación de ADN de alto rendimiento basada en regiones genómicas flanqueantes a sitios de restricción, que permite la obtención de datos genómicos de alta resolución para cualquier especie a costos razonables (Baird et al., 2008; Davey y Blaxter, 2010; Harvey et al., 2016). Esta técnica es aplicable para cualquier grupo taxonómico sin necesidad de información genómica previa (Baxter et al., 2011), a diferencia de otras técnicas como la captura de secuencias [*sequence capture*] o secuenciación de genomas completos, para las cuales tener un genoma de referencia es muy importante. En comparación con técnicas de *target enrichment* (UCE, exon-cap, etc.), la secuenciación de RADseq puede muestrear una mayor proporción del genoma con menos esfuerzo, por lo que proporciona una mayor cantidad de información en general y es una mejor opción para resolver relaciones a nivel intraespecífico (McCormack et al., 2013; Harvey et al., 2016; Collins y Hrbek, 2018). Estudios han mostrado que los resultados de análisis basados en datos de RADseq son congruentes con aquellos derivados de técnicas de *target enrichment* (Harvey et al., 2016; Collins y Hrbek, 2018).

Los datos de RADseq se han utilizado en varios estudios con el propósito de dilucidar la historia evolutiva de taxones a niveles intra- e interespecíficos (Eaton and Ree, 2013; Combosch y Vollmer, 2015; Darwell et al., 2016). Además, se han utilizado datos de RADseq, así como otros tipos de datos producidos por métodos de secuenciación reducida del genoma (p. ej. GBS), para estudiar la estructura poblacional de complejos de especies-anillo (Alcaide et al., 2014; Bouzid et al., 2021).

La secuenciación de RADseq se lleva a cabo digiriendo ADN genómico usando enzimas de restricción. A los fragmentos obtenidos se les colocan enlazadores [*linkers*] en los extremos para después ser cortados en fragmentos más pequeños. Después se lleva a cabo un proceso de purificación para obtener solamente los fragmentos adyacentes a los sitios de corte, es decir, los fragmentos que tienen los *linkers* y que finalmente se secuencian (Miller et al., 2007; Figura 14). Las regiones secuenciadas son las adyacentes a los sitios de corte de las enzimas de restricción, las cuales varían entre 100 y 500 pb, y se definen como RAD loci, RADseq loci o RAD tags; éstas son las unidades sobre las que se basa el ensamblaje (Baird et al., 2008; Rochette y Catchen, 2017). En teoría, estas secuencias adyacentes a sitios de restricción son ortólogas, y el agrupamiento de las secuencias ortólogas es la principal tarea del ensamblaje y procesamiento bioinformático de las secuencias. En otras palabras, los métodos de secuenciación a partir de sitios de restricción obtienen una representación reducida del genoma en la que se obtiene una gran cantidad de secuencias ortólogas que posteriormente se agrupan o “apilan” para obtener información de homología y hacer inferencias sobre relaciones evolutivas (Rubin et al., 2012; Catchen et al., 2013; Eaton, 2014; Ilut et al., 2014).

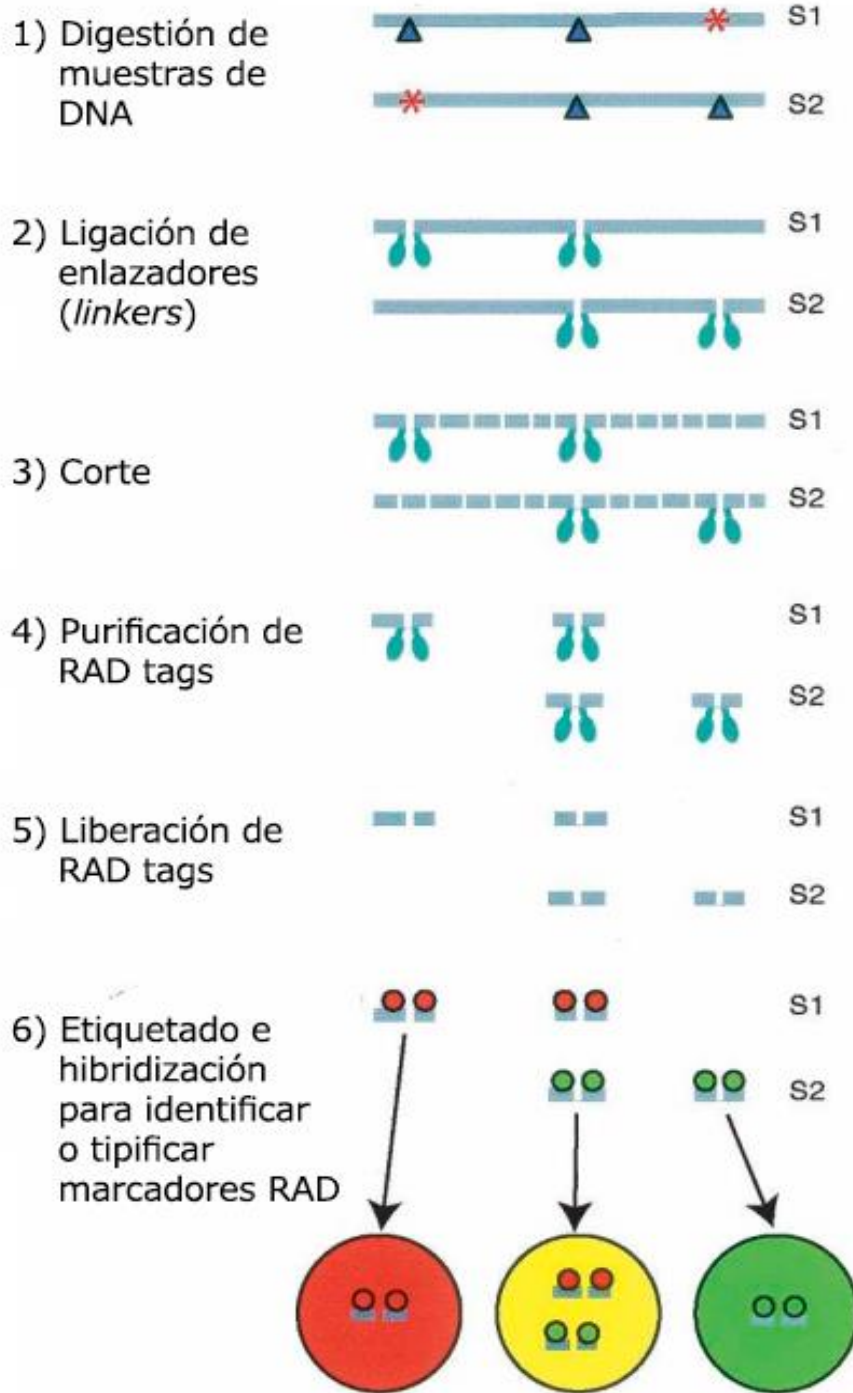


Figura 14. Proceso de identificación y tipificación de marcadores RAD (ADN asociado a enzimas de restricción). Las muestras de ADN genómico S1 y S2 tienen la secuencia de reconocimiento de varias enzimas de restricción en distintos lugares a lo largo del genoma. Los triángulos azules representan los sitios de restricción de una enzima particular. Algunos de estos sitios de restricción no están presentes en una muestra debido a polimorfismos que causan diferencias en las secuencias (asteriscos rojos). Las dos muestras son digeridas con una enzima de restricción particular y luego ligadas a los enlazadores biotinilados (elipses azules). El ADN es cortado al azar dejando sólo los fragmentos flanqueantes al sitio de restricción ligados a los enlazadores de biotina, los cuales luego son purificados (Modificado de Miller et al., 2007).

A partir de los RAD loci se pueden realizar alineamientos genómicos usando programas como *ipyRAD* (Eaton, 2014) o *Stacks* (Catchen et al., 2013), los cuales pueden ser la base para estimar árboles de genes o de especies. *ipyRAD* usa algoritmos globales de agrupamiento para realizar los alineamientos, lo cual es muy importante para realizar análisis filogenéticos (Eaton, 2014; Eaton y Overcast, 2020). Cada RAD locus puede tener uno o más “RAD alleles” (o “SNP alleles”) dependiendo de si se deriva de un locus heterocigoto u homocigoto. Esta genotipación se basa en los polimorfismos de un solo nucleótido (SNPs) que se detectan en los RAD loci (Baxter et al., 2011; Gautier et al., 2013), y es a partir de estos marcadores moleculares que se pueden realizar estudios de genómica de poblaciones (Hohenlohe et al., 2012).

Los ensamblajes que se realizan en *ipyRAD* utilizando datos de RADseq se llevan a cabo a partir de una serie de parámetros (28 en total) con los que se definen: características inherentes del conjunto de datos o de la especie de la que se obtuvieron las muestras (como la ploidía); especificaciones técnicas que no tienen que ver con las características de los datos ni con la forma en la que se procesan los datos (como el directorio donde están las lecturas, el nombre del ensamblaje, el formato de los archivos de salida, etc.); y variables que sí influyen en la forma en la que se lleva a cabo el procesamiento de datos (como el parámetro *clust_threshold*). Es importante señalar que esta clasificación de los parámetros en tres categorías puede llegar a ser simplista ya que un solo parámetro puede entrar dentro de más de una categoría. Aun así, se puede decir que los parámetros más importantes son aquellos que afectan la forma en la que se procesan los RADloci, ya que éstos pueden determinar si se obtiene un ensamblaje de mala calidad o de buena calidad independientemente de los datos crudos con los que se trabaje.

Parámetros importantes

La homología se refiere a similitud por ancestría, de modo que caracteres homólogos son aquellos que se comparten entre taxa, y que han sido heredados (y a veces modificados) de un ancestro común (Futuyma, 2003). Cuando un gen pasa por una duplicación y ambas copias del gen permanecen de forma paralela durante la historia de un taxón, a estos se les llama genes parálogos (pues su origen no es por ancestro-descendencia, sino por un evento de duplicación). Por otro lado, a copias de genes que fueron resultado de un proceso de especiación se les llama genes ortólogos; su historia refleja la historia de los taxa en cuestión. Cuando el objetivo es inferir las relaciones evolutivas de los taxa focales, únicamente los genes ortólogos, no los parálogos, son apropiados para inferir filogenias (Fitch, 1970). Seleccionar los genes (o regiones genéticas) focales es por lo tanto un paso esencial en los análisis filogenéticos –incluidos los filogenómicos– pues deben ser marcadores capaces de reflejar adecuadamente las relaciones entre las especies o poblaciones objeto de estudio (Kapli et al., 2020).

Uno de los parámetros más importantes para hacer ensamblajes con datos de secuenciación con base en sitios de restricción es el umbral de similitud (conocido en *ipyRAD* como “*clustering threshold*” y en *Stacks* como “*distance allowed between stacks*”), que se define como el porcentaje de similitud que deben tener las lecturas (*reads*) para considerar que pertenecen al mismo loci. Si se utiliza un valor de *clust_threshold* de 0.95, por ejemplo, entonces las lecturas que sean similares en un 95% o más se considerarán como pertenecientes al mismo locus y se agruparán (se “apilarán” en el lenguaje del programa *Stacks*), mientras que las lecturas que sean similares en un porcentaje menor se considerarán pertenecientes a loci diferentes, es decir, se considerarán parálogas. Si se utiliza un umbral de agrupamiento demasiado bajo, entonces se estarán agrupando lecturas que en realidad son parálogas [*under-splitting*], mientras que si se usa un valor demasiado alto se estarán sobre-separando [*over-splitting*] las lecturas (Harvey et al., 2016). Es importante recalcar que el uso de un

valor inadecuado en este parámetro puede llevar a errores en la genotipación y, por lo tanto, puede causar subestimación o sobreestimación en medidas de variación genética, como la heterocigosidad estimada (la cual es la proporción de sitios heterocigóticos), que es susceptible de variar dependiendo de los valores del umbral de agrupamiento (Rubin et al., 2012; Ilut et al., 2014).

Determinar el porcentaje de similitud u ortología de las lecturas es uno de los mayores retos en el uso de secuencias RAD para estudios filogenéticos (Rubin et al., 2012). Es difícil conocer *a priori* el valor óptimo de *clust_threshold* para un conjunto de datos ya que puede variar incluso entre poblaciones. Esto último se debe a varios factores, como el grado de diferenciación genética de alelos neutrales, el tamaño efectivo poblacional, o el nivel de heterocigosidad inherente de cada población (Ilut et al., 2014). Afortunadamente, se han ideado formas de estimar el valor adecuado de este y otros parámetros a partir de realizar diferentes ensambles con diferentes valores de parámetro y observar estadísticos o resultados de análisis posteriores, como la maximización del número total de SNPs obtenidos o la mayor suma de valores de *bootstrap* en filogenias hechas con los ensambles a partir de RAxML, entre otros (Mastretta-Yanes et al., 2015; McCartney-Melstad et al., 2019).

Otro parámetro importante que ha sido menos estudiado es el *min_samples_locus*, definido como el número mínimo de muestras que deben tener datos en un determinado locus. Si a este parámetro se le asigna un valor igual al número total de muestras, se obtiene un ensamble compuesto únicamente por loci que tienen datos en todas las muestras. Mientras mayor sea este valor, más densa será la matriz del ensamble final, y mientras menor sea, menos densa será la matriz. Al igual que el parámetro *clust_threshold*, el parámetro de número mínimo de muestras por locus suele ser optimizado realizando ensambles con distintos valores asignados para encontrar el más adecuado (p. ej. Jordon-Thaden et al., 2020; Nguyen et al., 2020), pero a diferencia de aquel parámetro, la optimización de *min_samples_locus* se basa en la mejor reconstrucción del árbol filogenético más que en la maximización del número de loci.

Existen otros parámetros a partir de los cuales se puede ajustar la proporción de bases ambiguas y bases heterocígas en las secuencias consenso. El parámetro *max_Ns_consens* sirve para definir el porcentaje máximo de bases ambiguas permitidas en las secuencias consenso. En este caso, si una base no se puede identificar con confianza estadística entonces se considera ambigua (N). Asimismo, el parámetro *max_Hs_consens* sirve para definir el porcentaje máximo de bases heterocígas permitidas en las secuencias consenso. Ambos parámetros tienen como finalidad no permitir una cantidad excesiva de “diversidad” en el ensamble final, debido a que el exceso de bases ambiguas y heterocígas puede ser indicativo de que el ensamble no se realizó adecuadamente, y se termina sobreestimando la cantidad de secuencias que presentan bases diferentes al consenso.

Anexo II. *Euphorbia tithymaloides*

Euphorbia tithymaloides es un arbusto moderadamente suculento, generalmente de 1 a 2 m de altura, que se ramifica desde la base. Pertenece a la tribu Euphorbieae, que se caracteriza por inflorescencias, que asemejan y funcionan como flores individuales; es decir, son pseudantios. Éstos se componen de una flor pistilada rodeada de flores estaminadas, las cuales están envueltas en un conjunto de brácteas fusionadas en un involucre; al pseudantio de *Euphorbia* se le denomina ciatio (Figura 15). *Euphorbia tithymaloides* pertenece al clado Pedilanthus, que se caracteriza por ciatios marcadamente bilaterales que asemejan a una zapatilla o un pajarito. La distribución geográfica de esta especie en el Caribe abarca el Este de México, Centroamérica, Florida, las Antillas Mayores, las Antillas Menores y el área costera del norte de Sudamérica (Figura 16).



Figura 15. Fotografía de un individuo de la especie *Euphorbia tithymaloides* con diversos ciatios, cortesía de la Dra. N. Ivalú Cacho.

Dressler (1957) reconoce seis subespecies en las islas del Caribe (*angustifolia*, *bahamensis*, *jamaicensis*, *padifolia*, *parasitica* y *smalli*) basándose en caracteres vegetativos, de morfología floral, y en la distribución geográfica de los individuos. A nivel intraespecífico, Dressler describe la existencia de intergradación y casos de hibridación en zonas que han sido alteradas por causas antropogénicas. Dressler reconoce a las subespecies *angustifolia* y *padifolia* como entidades “cercanas” debido a su morfología floral compartida, y sugiere que las subespecies simpátricas *angustifolia* y *parasitica* actúan como especies distintas, cada una conectada a las poblaciones continentales por medio de una cadena de poblaciones isleñas.

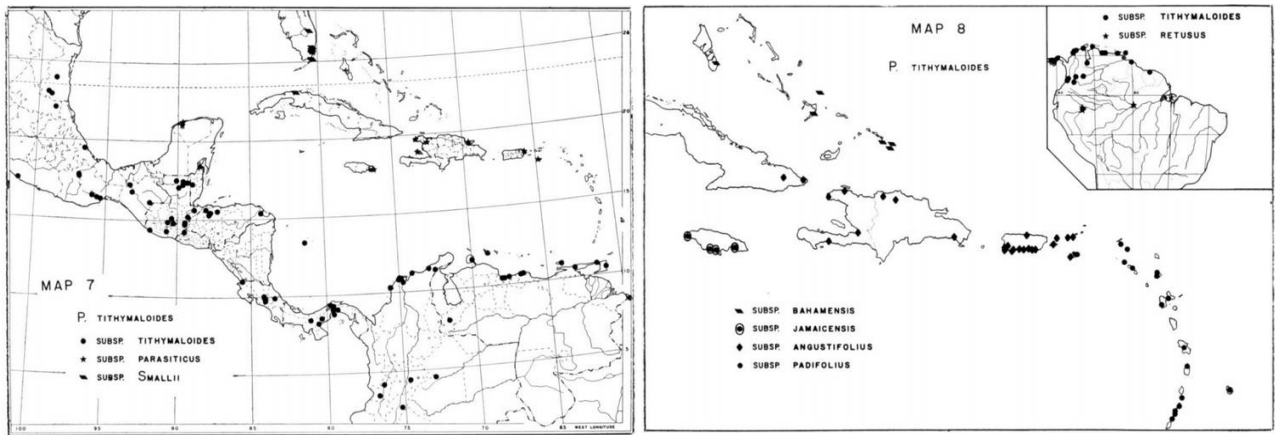


Figura 16. Distribución geográfica de *Euphorbia tithymaloides* (los puntos representan localidades; tomado de Dressler, 1957).

Los modos de dispersión por fruto o semilla que presenta *E. tithymaloides* son, al igual que la mayoría de las especies de *Euphorbia*, relativamente limitados: sus frutos son cápsulas explosivas, que al abrir expulsan las semillas no más allá de 2-3 metros de la planta madre. Asimismo, no existen adaptaciones evidentes en las semillas de *E. tithymaloides* para la dispersión por animales, viento, o agua (Dressler, 1957). Por otro lado, los individuos de *E. tithymaloides*, de forma similar a otras suculentas y otros miembros del género *Euphorbia*, son capaces de sobrevivir incluso semanas fuera de la tierra y enraízan con relativa facilidad a partir de esquejes (Dressler, 1957; N.I. Cacho, observaciones personales).

La evidencia disponible apunta a que los miembros del clado *Pedilanthus*, al que pertenece *E. tithymaloides*, son polinizados principalmente por colibríes, y esta relación pudo haber tenido un papel muy importante en la evolución de este clado (Dressler, 1957; Sahagún-Godínez y Lomelí-Sención, 1997; Cacho et al., 2010; Veiga Blanco et al., 2013). Además *E. tithymaloides* posee características clásicas del síndrome de polinización por colibríes, con ciatios relativamente alargados, bilaterales, de coloración rojiza, y con producción de néctar. Lack (1973) reporta que existen un total de 15 especies de colibríes en las islas del Caribe: cinco en las Antillas Menores, nueve en las Antillas Mayores y una perteneciente a las Bahamas. Otros autores reportan un total de 18 especies de colibríes en las islas del Caribe (Raffaele et al., 2003), agregando dos especies cuya distribución incluye a las Antillas Mayores y a las Bahamas, y una especie que habita las islas de Providencia y San Andrés, en Colombia. Las áreas de distribución de las especies de colibríes que habitan las Antillas Mayores y las Antillas Menores no se traslapan, excepto por dos especies típicamente de las Antillas Menores que también habitan el noreste de Puerto Rico: *Orthorhyncus cristatus* y *Eulampis holosericeus* (Lack, 1973; Raffaele et al., 2003). Por otro lado, el movimiento de colibríes de isla en isla o desde las islas hasta el continente no parece ser raro, pues hay varios registros de avistamientos de colibríes en islas fuera de su distribución normal (Lack, 1973). Por ejemplo, existen registros de la especie cubana *Chlorostilbon ricordii* y la especie de las Bahamas *Calliphlox evelynae* en Florida, Estados Unidos, aproximadamente a unos 200 km de su distribución normal (Stimson, 1944; Lack, 1973). Un caso particularmente notable es el avistamiento de *Orthorhyncus cristatus*, una especie endémica de las Antillas Menores y Puerto Rico en Texas, Estados Unidos, por lo menos a unos 3,300 km de su área de distribución normal (Pulich, 1968). Se sabe que esta última especie consume el néctar de *E. tithymaloides* en las Antillas Menores (Shepherdson, 2018).

Se ha observado que en otras especies de plantas del Caribe que son polinizadas por colibríes, como el género *Heliconia*, los colibríes suelen dividirse entre aquellos que son territoriales y aquellos denominados *trapliners*. Los colibríes territoriales suelen llevar a cabo polinización a nivel local. Por

otro lado, los colibríes *trapliners* suelen tener un patrón de forrajeo de distancias mayores, lo que podría facilitar el flujo genético entre poblaciones (Volpe et al., 2014; Torres-Vanegas et al., 2019).

Euphorbia tithymaloides tiene un genoma nuclear de aproximadamente 1.85 Gb (Cacho y McIntyre, no publicado), lo cual es menor al tamaño promedio reportado para *Euphorbia*: 4.36 Gb (Leitch et al., 2019). Los parientes más cercanos de *E. tithymaloides* para los que existen esfuerzos de secuenciación genómica son los siguientes: *Euphorbia esula*, *Hevea brasiliensis*, *Jatropha curcas*, *Ricinus communis* y *Manihot esculenta* (Chan et al., 2010; Rahman et al., 2013; Wang et al., 2014; Wu et al., 2015; Horvath et al., 2018). Sin embargo, la lejanía filogenética de estas especies en relación con *E. tithymaloides* limita la utilidad de dichos genomas como referencia para realizar el ensamble del genoma de esta especie. La mayoría de las especies mencionadas (excepto *E. esula*) pertenecen a subfamilias distintas a la de *E. tithymaloides* dentro de la familia Euphorbiaceae. El género *Euphorbia* se divide en cuatro subgéneros: *Esula*, *Athymalus*, *Chamaesyce* y *Euphorbia*. *Euphorbia esula* pertenece al subgénero *Esula*, mientras que *E. tithymaloides* pertenece al subgénero *Euphorbia*. El subgénero *Euphorbia* y el subgénero *Chamaesyce* forman un clado que es el grupo hermano del subgénero *Athymalus*. El clado formado por estos tres subgéneros a su vez es el grupo hermano del subgénero *Esula* (Dorsey et al., 2013; Riina et al., 2013). Por lo tanto, *E. tithymaloides* no está tan cercanamente emparentada a *E. esula*, en comparación con otras especies dentro del género. Por esa razón, usar el genoma de *E. esula* como referencia al momento de realizar los ensamblajes genómicos de *E. tithymaloides* no es viable.

El Caribe y *Euphorbia tithymaloides*

El Caribe es una región geográfica altamente compleja, compuesta por más de 100 provincias geológicas (Case et al., 1984; Draper et al., 1994). La formación del Caribe comenzó aproximadamente entre el periodo Triásico tardío y el periodo Jurásico yendo de la mano con la desaparición del supercontinente Pangea, aunque esta formación geológica aún era muy distinta a la cuenca del Caribe actual (Coney, 1982; Iturralde-Vinent and MacPhee, 1999; Iturralde-Vinent, 2004-2005). La provincia geológica conocida como las Antillas Mayores es un cinturón orogénico cuyas islas se componen de una base oceánica Jurásica superpuesta por depósitos volcánicos y epiclásticos que datan desde el Cretácico tardío hasta el Paleógeno. Las Antillas Menores son un arco volcánico que va desde la isla de Granada en el sur, hasta el Pasaje de Anegada en el norte, y tienen un origen independiente al de las Antillas Mayores: se empezaron a formar por actividad volcánica posterior al término del vulcanismo en las Antillas Mayores, durante el Paleógeno (Coney, 1982; Draper et al., 1994; Wadge, 1994). La región de la cuenca del Caribe ha sufrido una serie de cambios y transformaciones a lo largo del tiempo asociados al aumento y disminución del nivel del mar, el movimiento de capas tectónicas, y la aparición y desaparición de islas (Coney, 1982; Iturralde-Vinent and MacPhee, 1999; Iturralde-Vinent, 2004-2005).

Los estudios que se han encargado de evaluar cómo distintas especies llegaron a habitar el Caribe han sugerido diversas respuestas. En el caso de las plantas, distintas especies y géneros parecen haber llegado y haberse expandido por el Caribe de maneras muy diversas, ya sea por dispersión por agua o vicarianza, y presentando mayor parentesco a taxones de distintas zonas continentales (Santiago-Valentin y Olmstead, 2004). En casos de especies que no son capaces de volar ni nadar (como los mamíferos terrestres y las plantas), una hipótesis que ha tenido aceptación es la de utilización de GAARlandia, una proyección de tierra que pudo haber conectado brevemente a las Antillas Mayores con el noroeste de Sudamérica durante el Eoceno-Oligoceno, dando lugar a la distribución actual de estas especies, que debió originarse por una migración y eventos posteriores de vicarianza (Iturralde-Vinent and MacPhee, 1999).

En el caso particular de *Euphorbia tithymaloides*, esta especie es capaz de desplazarse de unas islas a otras [*island-hop*] de forma ocasional (Dressler, 1957), por lo que el agua no parece ser una barrera total para la dispersión, sino únicamente parcial. Se ha planteado que *E. tithymaloides* se originó en el continente y de ahí debió haber colonizado las islas del Caribe aproximadamente en el Plioceno tardío y Pleistoceno (Dressler, 1957). Esta datación coincide con estimaciones recientes de reloj molecular que sugieren que los colibríes (los polinizadores de *E. tithymaloides*) colonizaron el Caribe aproximadamente hace 5 millones de años (McGuire et al., 2014). Respecto al modo de dispersión operante, éste pudo haber sido con base en esquejes de partes vegetativas de la planta, las cuales son capaces de flotar y establecerse hasta en condiciones relativamente desfavorables (Dressler, 1957, Cacho, observaciones personales), pues las semillas no parecen presentar adaptaciones para un modo de dispersión por agua. La hipótesis de colonización y dispersión de *E. tithymaloides* a partir de "*island-hopping*" es consistente con resultados de estudios posteriores (Cacho and Baum, 2012; Cacho et al., 2019).

Si las hipótesis anteriormente expuestas son ciertas se puede concluir que: (a) la colonización de *E. tithymaloides* a las islas del Caribe es un evento relativamente reciente (Plioceno-Pleistoceno) que ocurrió una vez que las islas ya estaban formadas; (b) la distribución actual de *E. tithymaloides* en el Caribe es resultado (al menos mayoritariamente) de eventos de dispersión, pues su colonización del Caribe ocurrió tiempo después de los eventos de vicarianza que han sido documentados para otras especies (Eoceno-Oligoceno), y; (c) la principal forma de dispersión de esta especie parece ser por medio de las partes vegetativas de la planta y no semillas o frutos.

***Euphorbia tithymaloides* en México**

Euphorbia tithymaloides pertenece al subgénero *Pedilanthus*, el cual anteriormente poseía la categoría de género. Citando literalmente a Dressler (1957), menciona que el género es "esencialmente mexicano", pues "no hay razón para dudar de que el género haya evolucionado en México tropical, su centro de diversidad en el presente". A pesar de que para dicho autor un centro de origen no puede ser determinado, se destaca que el mayor número de especies de este subgénero se encuentran en la región centro-occidental de México (Dressler and Sacamano, 1992; Lomelí-Sención, 2006).

En el estudio filogeográfico de Cacho y Baum (2012) se realizó una filogenia con individuos de *Euphorbia tithymaloides* de distintas zonas de México, Centroamérica, el norte de Sudamérica y las islas del Caribe. La topología de dicha filogenia sugiere que el centro de expansión geográfica de esta especie se encuentra entre México y Guatemala, lo cual coincide con la hipótesis de que lo que hoy es el centro-sur de México pudo haber sido la principal zona en la que se originó *E. tithymaloides*. Consistente con estas observaciones, las poblaciones pertenecientes a México y América Central una mayor heterocigosidad que las poblaciones insulares.

Anexo III. Tabla detallada de los ensambles realizados

Tabla 6. Ensamblados con los valores usados de los cuatro parámetros modificados, el número de loci resultante, el número de SNPs resultante, el porcentaje de sitios faltantes (*missing sites*) en la matriz de SNPs y la topología resultante a partir de análisis RAxML.

Nombre del ensamble	Clustering threshold	Min. samples per locus	Maximum N's consens	Maximum H's consens	Número de loci retenidos	Número de SNPs	Porcentaje de sitios faltantes en la matriz de SNPs	Topología resultante (RAxML)
alldefault	0.85	4	0.05	0.05	26,176	204,971	89.50%	1-origen
min_sample_locus 02	0.85	2	0.05	0.05	113,220	593,172	95.03%	1-origen
euphorbia	0.88	4	0.05	0.05	20,753	166,016	88.08%	1-origen
min12	0.88	12	0.05	0.05	2,473	18,953	38.75%	1-origen
min25	0.88	25	0.05	0.05	1,677	13,096	18.52%	1-origen
min50	0.88	50	0.05	0.05	1,338	11,012	9.78%	1-origen
min75	0.88	75	0.05	0.05	1,107	9,191	4.98%	1-origen
min85	0.88	85	0.05	0.05	1,012	8,333	3.68%	2-orígenes
min95	0.88	95	0.05	0.05	87	720	0.00%	1-origen
min50_maxN01_maxH01	0.88	50	0.01	0.01	1,299	4,674	22.87%	1-origen
min50_maxN01_maxH10	0.88	50	0.01	0.10	1,331	10,669	10.27%	2-orígenes
min50_maxN10_maxH01	0.88	50	0.10	0.01	1,301	4,702	22.63%	1-origen
min50_maxN10_maxH10	0.88	50	0.10	0.10	1,313	10,478	9.78%	2-orígenes
min85_maxN01_maxH01	0.88	85	0.01	0.01	302	517	8.59%	2-orígenes
min85_maxN01_maxH10	0.88	85	0.01	0.10	991	7,852	3.93%	2-orígenes

min85_maxN10_maxH01	0.88	85	0.10	0.01	535	535	8.54%	2-orígenes
min85_maxN10_maxH10	0.88	85	0.10	0.10	990	7,870	3.85%	2-orígenes
min85_maxN20_maxH20	0.88	85	0.20	0.20	989	7,865	3.78%	2-orígenes
maxH01	0.85	4	0.05	0.01	17,679	132,998	91.31%	1-origen
maxH10	0.85	4	0.05	0.10	19,808	161,147	88.23%	1-origen
maxH25	0.85	4	0.05	0.25	19,372	158,671	88.26%	1-origen
maxH50	0.85	4	0.05	0.50	19,333	158,294	88.24%	1-origen
maxH75	0.85	4	0.05	0.75	19,333	158,294	88.24%	1-origen

Referencias

- Aguirre-Liguori, J. A., J. A. Luna-Sánchez, J. Gasca-Pineda, and L. E. Eguiarte. 2020. Evaluation of the Minimum Sampling Design for Population Genomic and Microsatellite Studies: An Analysis Based on Wild Maize. *Frontiers in Genetics* 11: 1–16.
- Alcaide, M., E. S. C. Scordato, T. D. Price, and D. E. Irwin. 2014. Genomic divergence in a ring species complex. *Nature* 511: 83–85.
- Aldrich, J. 1997. R. A. Fisher and the making of maximum likelihood 1912 - 1922. *Statistical Science* 12: 162–176.
- Amorim, A. 2013. Population Genetics. Elsevier Inc.
- Avice, J. C. 2000. Phylogeography: The History and Formation of Species. Harvard University Press, Cambridge, MASS., USA.
- Avice, J. C., J. Arnold, R. M. Ball, E. Bermingham, T. Lamb, J. E. Neigel, C. A. Reeb, and N. C. Saunders. 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual review of ecology and systematics* 18: 489–522.
- Baird, N. A., P. D. Etter, T. S. Atwood, M. C. Currey, A. L. Shiver, Z. A. Lewis, E. U. Selker, et al. 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE* 3: 1–7.
- Baxter, S. W., J. W. Davey, J. S. Johnston, A. M. Shelton, D. G. Heckel, C. D. Jiggins, and M. L. Blaxter. 2011. Linkage mapping and comparative genomics using next-generation rad sequencing of a non-model organism. *PLoS ONE* 6.
- Bouzig, N. M., J. W. Archie, R. A. Anderson, J. A. Grummer, and A. D. Leaché. 2021. Evidence for ephemeral ring species formation during the diversification history of western fence lizards (*Sceloporus occidentalis*). *Molecular Ecology*: 1–12.
- Bowen, B. W., A. L. Bass, L. A. Rocha, W. S. Grant, and D. R. Robertson. 2001. Phylogeography of the trumpetfishes (*Aulostomus*): Ring species complex on a global scale. *Evolution* 55: 1029–1039.
- Cacho, N. I., and D. A. Baum. 2012. The Caribbean slipper spurge *Euphorbia tithymaloides*: The first example of a ring species in plants. *Proceedings of the Royal Society B: Biological Sciences* 279: 3377–3383.
- Cacho, N. I., P. E. Berry, M. E. Olson, V. W. Steinmann, and D. A. Baum. 2010. Are spurred cyathia a key innovation? Molecular systematics and trait evolution in the slipper spurge (*Pedilanthus* clade: *Euphorbia*, *Euphorbiaceae*). *American Journal of Botany* 97: 493–510.

- Cacho, N. I., M. J. Monteverde-Suárez, and P. J. McIntyre. 2019. Convergent evolution in floral morphology in a plant ring species, the Caribbean *Euphorbia tithymaloides*. *American Journal of Botany* 106: 1032–1045.
- Case, J. E., T. L. Holcombe, and R. G. Martin. 1984. Map of geologic provinces in the Caribbean region. *Geological Society of America Memoir* 162: 1–30.
- Catchen, J., P. A. Hohenlohe, S. Bassham, A. Amores, and W. A. Cresko. 2013. Stacks: An analysis tool set for population genomics. *Molecular Ecology* 22: 3124–3140.
- Chakraborty, R. 1993. Analysis of Genetic Structure of Populations: Meaning, Methods, and Implications. *Human Population Genetics*, 189–206. Springer, Boston, MA.
- Chan, A. P., J. Crabtree, Q. Zhao, H. Lorenzi, J. Orvis, D. Puiu, A. Melake-berhan, et al. 2010. Draft genome sequence of the ricin-producing oilseed castor bean. *Nature biotechnology* 28: 951–956.
- Collins, R. A., and T. Hrbek. 2018. An in silico comparison of protocols for dated phylogenomics. *Systematic Biology* 67: 633–650.
- Combosch, D. J., and S. v. Vollmer. 2015. Trans-Pacific RAD-Seq population genomics confirms introgressive hybridization in Eastern Pacific *Pocillopora* corals. *Molecular Phylogenetics and Evolution* 88: 154–162.
- Coney, P. J. 1982. Plate tectonic constraints on the biogeography of middle America and the Caribbean region (Central America). *Annals - Missouri Botanical Garden* 69: 432–443.
- Coyne, J. A., and H. A. Orr. 2004. *Speciation*. Sinauer Associates, Inc., Sunderland, MA.
- Cutter, A. D. 2013. Integrating phylogenetics, phylogeography and population genetics through genomes and evolutionary theory. *Molecular Phylogenetics and Evolution* 69: 1172–1185.
- Darwell, C. T., D. M. Rivers, and D. M. Althoff. 2016. RAD-seq phylogenomics recovers a well-resolved phylogeny of a rapid radiation of mutualistic and antagonistic yucca moths. *Systematic Entomology* 41: 672–682.
- Davey, J. L., and M. W. Blaxter. 2010. RADseq: Next-generation population genetics. *Briefings in Functional Genomics* 9: 416–423.
- Dobzhansky, Th. 1951. *Genetics and the Origin of Species*. 3rd edition. Columbia University Press, New York City, NY.
- Dorsey, B. L., T. Haeuermans, X. Aubriot, J. J. Morawetz, R. Riina, V. W. Steinmann, and P. E. Berry. 2013. Phylogenetics, morphological evolution, and classification of *Euphorbia* subgenus *Euphorbia*. *Taxon* 62: 291–315.
- Draper, G., T. A. Jackson, and S. K. Donovan. 1994. Geologic provinces of the Caribbean. *Caribbean Geology: An Introduction*, 3–12. University of the West Indies Publishers' Association, Kingston.
- Dressler, R. L. 1957. *The Genus Pedilanthus*. Contributions from the Gray Herbarium of Harvard University, Cambridge, MA.
- Dressler, R. L., and C. M. Sacamano. 1992. *Pedilanthus connatus* (Euphorbiaceae), especie nueva y llamativa de Jalisco. *Acta Botanica Mexicana*: 21.
- Eaton, D. A. R. 2014. PyRAD: assembly of de novo RADseq loci for phylogenetic analyses. *Bioinformatics* 30: 1844–1849.
- Eaton, D. A. R., and I. Overcast. 2020. Ipyrad: Interactive assembly and analysis of RADseq datasets. *Bioinformatics* 36: 2592–2594.
- Eaton, D. A. R., and R. H. Ree. 2013. Inferring phylogeny and introgression using RADseq data: An example from flowering plants (Pedicularis: Orobanchaceae). *Systematic Biology* 62: 689–706.
- Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology* 14: 2611–2620.
- Felsenstein, J. 1981. Evolutionary Trees from DNA Sequences: A Maximum Likelihood Approach. *Journal of Molecular Evolution* 17: 368–376.
- Fisher, R. A. 1922. On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society of London* 222: 594–604.
- Fisher, R. A. 1930. *The Genetical Theory of Natural Selection*. Clarendon Press, London.
- Fitch, W. M. 1970. Distinguishing homologous from analogous proteins. *Systematic Zoology* 19: 99–113.
- Fuchs, J., P. G. P. Ericson, C. Bonillo, A. Couloux, and E. Pasquet. 2015. The complex phylogeography of the Indo-Malayan *Alophoixus* bulbuls with the description of a putative new ring species complex. *Molecular Ecology* 24: 5460–5474.
- Futuyma, D. J. 2003. *Evolution*. Sinauer Associates, Inc., Sunderland, MA.
- Gautier, M., K. Gharbi, T. Cezard, J. Foucaud, C. Kerdelhué, P. Pudlo, J. M. Cornuet, and A. Estoup. 2013. The effect of RAD allele dropout on the estimation of genetic variation within and between populations. *Molecular Ecology* 22: 3165–3178.

- Gemmell, A. P., and J. M. Marcus. 2015. A tale of two haplotype groups: Evaluating the New World *Junonia* ring species hypothesis using the distribution of divergent COI haplotypes. *Systematic Entomology* 40: 532–546.
- Haldane, J. B. S. 1932. *The Causes of Evolution*. Longmans, Green and Co., London.
- Harvey, M. G., B. T. Smith, T. C. Glenn, B. C. Faircloth, and R. T. Brumfield. 2016. Sequence Capture versus Restriction Site Associated DNA Sequencing for Shallow Systematics. *Systematic Biology* 65: 910–924.
- Hedrick, P. W. 2009. *Genetics of Populations*, 4th ed. Jones & Bartlett Publishers, Boston, MA.
- Hennig, W. 1966. *Phylogenetic Systematics*. University of Illinois Press, Chicago, IL.
- Hennig, W. 1965. *Phylogenetic Systematics*. *Annual Review of Entomology* 10: 97–116.
- Hohenlohe, P. A., J. Catchen, and W. A. Cresko. 2012. Population Genomic Analysis of Model and Nonmodel Organisms Using Sequenced RAD Tags. *Methods in Molecular Biology*, 235–260. Humana Press, Totowa, NJ.
- Horvath, D. P., S. Patel, M. Dođramaci, W. S. Chao, J. v. Anderson, M. E. Foley, B. Scheffler, et al. 2018. Gene Space and Transcriptome Assemblies of Leafy Spurge (*Euphorbia esula*) Identify Promoter Sequences, Repetitive Elements, High-Quality Markers, and a Full-Length Chloroplast Genome. *Weed Science* 66: 355–367.
- Ilut, D. C., M. L. Nydam, and M. P. Hare. 2014. Defining loci in restriction-based reduced representation genomic data from nonmodel species: Sources of bias and diagnostics for optimal clustering. *BioMed Research International* 2014.
- Irwin, D. E., J. H. Irwin, and T. D. Price. 2001. Ring species as bridges between microevolution and speciation. *Genetica* 112–113: 223–243.
- Iturralde-Vinent, M. A. La Paleogeografía del Caribe y sus implicaciones para la biogeografía histórica. *Revista del Jardín Botánico Nacional* 25–26: 49–78.
- Iturralde-Vinent, M., and R. D. MacPhee. 1999. Paleogeography of the Caribbean region: implications for Cenozoic biogeography. *Bulletin of the American Museum of Natural History* No. 238.
- Jordon-Thaden, I. E., J. B. Beck, C. A. Rushworth, M. D. Windham, N. Diaz, J. T. Cantley, C. T. Martine, and C. J. Rothfels. 2020. A basic ddRADseq two-enzyme protocol performs well with herbarium and silica-dried tissues across four genera. *Applications in Plant Sciences* 8: 1–7.
- Joseph, L., G. Dolman, S. Donnellan, K. M. Saint, M. L. Berg, and A. T. D. Bennett. 2008. Where and when does a ring start and end? Testing the ring-species hypothesis in a species complex of Australian parrots. *Proceedings of the Royal Society B: Biological Sciences* 275: 2431–2440.
- Kapli, P., Z. Yang, and M. J. Telford. 2020. Phylogenetic tree building in the genomic age. *Nature Reviews Genetics* 21: 428–444.
- Kuchta, S. R., D. S. Parks, R. L. Mueller, and D. B. Wake. 2009. Closing the ring: Historical biogeography of the salamander ring species *Ensatina eschscholtzii*. *Journal of Biogeography* 36: 982–995.
- Lack, D. 1973. The Numbers of Species of Hummingbirds in the West Indies. *Evolution* 27: 326.
- Leitch, I. J., E. Johnston, J. Pellicer, O. Hidalgo, and M. D. Bennett. 2019. Angiosperm DNA C-values database (release 9.0, Apr 2019). Website <https://cvalues.science.kew.org/>.
- Liebers, D., P. de Knijff, and A. J. Helbig. 2004. The herring gull complex is not a ring species.
- Lomelí-Senci3n, J. A. 2006. Revisi3n de *Euphorbia* subg3nero *Pedilanthus*. Tesis de doctorado. Universidad de Guadalajara.
- Mastretta-Yanes, A., N. Arrigo, N. Alvarez, T. H. Jorgensen, D. Piñero, and B. C. Emerson. 2015. Restriction site-associated DNA sequencing, genotyping error estimation and de novo assembly optimization for population genetic inference. *Molecular Ecology Resources* 15: 28–41.
- Mayr, E. 1940. Speciation Phenomena in Birds. *The American Naturalist* 74: 249–278.
- McCartney-Melstad, E., M. Gidiş, and H. B. Shaffer. 2019. An empirical pipeline for choosing the optimal clustering threshold in RADseq studies. *Molecular Ecology Resources* 19: 1195–1204.
- McCormack, J. E., S. M. Hird, A. J. Zellmer, B. C. Carstens, and R. T. Brumfield. 2013. Applications of next-generation sequencing to phylogeography and phylogenetics. *Molecular Phylogenetics and Evolution* 66: 526–538.
- McGuire, J. A., C. C. Witt, J. v. Remsen, A. Corl, D. L. Rabosky, D. L. Altshuler, and R. Dudley. 2014. Molecular phylogenetics and the diversification of hummingbirds. *Current Biology* 24: 910–916.
- Miller, M. R., J. P. Dunham, A. Amores, W. A. Cresko, and E. A. Johnson. 2007. Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Research* 17: 240–248.
- Nazareno, A., J. Bemmels, C. Dick, and L. Lohmann. 2017. Minimum sample sizes for population genomics: an empirical study from an Amazonian plant species. *Molecular Ecology Resources* 17: 1136–1147.

- Nguyen, D. D., E. Oguri, A. Yamada, C. C. Lin, C. Zhilin, A. D. Nguyen, Y. Suyama, and K. Eguchi. 2020. Genome-wide mig-seq and morphometric data reveals heterospecificity of the gnamptogenys taivanensis group (Hymenoptera: Formicidae: Ectatomminae) in the northern mountainous region of vietnam. *Raffles Bulletin of Zoology* 68: 539–555.
- Nguyen, L. T., H. A. Schmidt, A. von Haeseler, and B. Q. Minh. 2015. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution* 32: 268–274.
- Pickrell, J. K., and J. K. Pritchard. 2012. Inference of Population Splits and Mixtures from Genome-Wide Allele Frequency Data. *PLoS Genetics* 8.
- Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of Population Structure Using Multilocus Genotype Data. *Genetics* 155: 945–959.
- Pulich, W. M. 1968. The Occurrence of the Crested Hummingbird, *Orthorhyncus cristatus exilis*, in the United States. *The Auk* 85: 322.
- Qi, Y., L. Ding, Y. Zhao, C. Niu, Ex. Wang, and W. Zhao. 2020. Toad-headed lizard phrynocephalus forsythii (Squamata, Agamidae) as a potential ring species inferred from population genetic differentiation. *Asian Herpetological Research* 11: 312–319.
- de Queiroz, K. 1999. The General Lineage Concept of Species and the Defining Properties of the Species Category. *Species: New Interdisciplinary Essays*, 49–89. The MIT Press, Cambridge, MA.
- de Queiroz, K. 1998. The General Lineage Concept of Species, Species Criteria, and the Process of Speciation. *Endless Forms: Species and Speciation*, 57–75. Oxford University Press, New York City, NY.
- Raffaele, H., J. Wiley, O. Garrido, A. Keith, and J. Raffaele. 2003. *Birds of the West Indies*. Princeton University Press, Princeton, NJ.
- Rahman, A. Y. A., A. O. Usharraj, B. B. Misra, G. P. Thottathil, K. Jayasekaran, Y. Feng, S. Hou, et al. 2013. Draft genome sequence of the rubber tree *Hevea brasiliensis*. *BMC Genomics* 14.
- Riina, R., J. A. Peirson, D. v. Geltman, J. Molero, B. Frajman, A. Pahlevani, L. Barres, et al. 2013. A worldwide molecular phylogeny and classification of the leafy spurges, *Euphorbia* subgenus *Esula* (Euphorbiaceae). *Taxon* 62: 316–342.
- Rochette, N. C., and J. M. Catchen. 2017. Deriving genotypes from RAD-seq short-read data using Stacks. *Nature Protocols* 12: 2640–2659.
- Rubin, B. E. R., R. H. Ree, and C. S. Moreau. 2012. Inferring phylogenies from RAD sequence data. *PLoS ONE* 7: 1–12.
- Sahagún-Godínez, E., and J. A. Lomelí-Senci3n. 1997. *Pedilanthus diazluanus* (Euphorbiaceae): Pollination by hymenopterans in a bird-pollinated genus. *American Journal of Botany* 84: 1584–1587.
- Santiago-Valentin, E., and R. G. Olmstead. 2004. Historical Biogeography of Caribbean Plants: Introduction to Current Knowledge and Possibilities from a Phylogenetic Perspective. *Taxon* 53: 299–319.
- Shepherdson, J. P. 2018. Observations of territorial behavior of the Antillean Crested Hummingbird (*Orthorhyncus cristatus*) on St. Eustatius. *Journal of Caribbean Ornithology* 31: 48–50.
- Simpson, G. G. 1961. *Principles of animal taxonomy*. Columbia University Press, New York City, NY.
- Soltani, G. G., D. B3non, N. Alvarez, and C. J. Praz. 2017. When different contact zones tell different stories: Putative ring species in the *Megachile concinna* species complex (Hymenoptera: Megachilidae). *Biological Journal of the Linnean Society* 121: 815–832.
- Stamatakis, A. 2014. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30: 1312–1313.
- Stimson, L. A. 1944. *Riccordia ricordii* Seen at Miami, Florida. *The Auk* 61: 300–301.
- Torres-Vanegas, F., A. S. Hadley, U. G. Kormann, F. A. Jones, M. G. Betts, and H. H. Wagner. 2019. The Landscape Genetic Signature of Pollination by Trappliners: Evidence From the Tropical Herb, *Heliconia tortuosa*. *Frontiers in Genetics* 10: 1–12.
- Veiga Blanco, T., L. Galetto, and I. C. Machado. 2013. Nectar regulation in *Euphorbia tithymaloides* L., a hummingbird-pollinated Euphorbiaceae. *Plant Biology* 15: 910–918.
- Volpe, N. L., A. S. Hadley, W. D. Robinson, and M. G. Betts. 2014. Functional connectivity experiments reflect routine movement behavior of a tropical hummingbird species. *Ecological Applications* 24: 2122–2131.
- Volz, E. M., C. Wiuf, Y. H. Grad, S. D. W. Frost, A. M. Dennis, and X. Didelot. 2020. Identification of Hidden Population Structure in Time-Scaled Phylogenies. *Systematic Biology* 69: 884–896.
- Wadge, G. 1994. The Lesser Antilles. In S. K. Donovan, and T. A. Jackson [eds.], *Caribbean Geology: An Introduction*, 167–177. The University of the West Indies Publishers' Association (UWIPA), Kingston.

- Wang, W., B. Feng, J. Xiao, Z. Xia, X. Zhou, P. Li, W. Zhang, et al. 2014. Cassava genome from a wild ancestor to cultivated varieties. *Nature Communications* 5.
- Wen, G., and J. Fu. 2021. Isolation and re-connection: the formation of a ring-shaped speciation continuum in an odorous frog (*Odorrana margaretae*). *Authorea*: 1–40.
- Willing, E. M., C. Dreyer, and C. van Oosterhout. 2012. Estimates of genetic differentiation measured by f_{st} do not necessarily require large sample sizes when using many snp markers. *PLoS ONE* 7: 1–7.
- Wright, S. 1978. *Evolution and the Genetics of Populations, Volume 4: Variability within and among natural populations*. University of Chicago Press, Chicago, IL.
- Wright, S. 1931. Evolution in mendelian populations. *Genetics* 16: 97–159.
- Wright, S. 1951. The genetical structure of populations. *Annals of Eugenics* 15: 323–354.
- Wu, P., C. Zhou, S. Cheng, Z. Wu, W. Lu, J. Han, Y. Chen, et al. 2015. Integrated genome sequence and linkage map of physic nut (*Jatropha curcas* L.), a biodiesel plant. *Plant Journal* 81: 810–821.