



**UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO**

FACULTAD DE CIENCIAS

**Efectos paisajísticos en la distribución de la
diversidad genética en una población de *Abies
religiosa***

T E S I S

**QUE PARA OBTENER EL TÍTULO DE:
BIÓLOGO**

P R E S E N T A:

IVÁN DAVID MÉNDEZ GONZÁLEZ



**DIRECTOR DE TESIS:
DR. JUAN PABLO JARAMILLO CORREA**

CIUDAD UNIVERSITARIA, MÉXICO, D.F.

DICIEMBRE DEL 2014



Universidad Nacional
Autónoma de México



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.

Hoja de datos del jurado

1. Datos del alumno

Méndez González Iván David
Universidad Nacional Autónoma de México
Facultad de Ciencias
Biología
410087404

2. Datos del Tutor

Dr.
Juan Pablo Jaramillo Correa

3. Datos del Sinodal

Dra.
Ella Vázquez Domínguez

4. Datos del Sinodal

Dr.
Lev Orlando Jardón Barbolla

5. Datos del Sinodal

M. en C.
Gabriela Castellanos Morales

6. Datos del Sinodal

Dra.
Erika Aguirre Planter

7. Datos del trabajo escrito

Efectos paisajísticos en la distribución de la variación genética en una población de *Abies religiosa*.
pp 114
2014

AGRADECIMIENTOS

A la UNAM, por la gran experiencia que resultó estudiar en una universidad de estas dimensiones (en muchos sentidos).

A mi tutor, el Dr. Juan Pablo Jaramillo Correa, por los conocimientos compartidos, la libertad para el análisis de datos, y la insistencia en el auto aprendizaje y la constante búsqueda de información. A mis sinodales, por los valiosos comentarios que sin duda ayudaron a mejorar este trabajo.

A la M. En C. Ariadna Morales y a la Biol. Laura Figueroa, por toda la ayuda en el largo camino del trabajo de laboratorio, y en las pocas pero divertidas salidas al campo. Al Dr. Lev Jardón y a Nancy, por ayudarme a completar las colectas.

A todos los buenos maestros, tanto dentro como fuera del aula, especialmente al profe Ramón, Martínez Vences.

A mis padres, Roberto Méndez y Ana Luisa González, por todo el apoyo en las constantes dudas . A mis hermanos, Beto, Bárbara y Paco, por ser mis mejores amigos y siempre estar ahí cuando han sido necesarios. A mi familia en general, especialmente a la familia Rubí por compartirme su hogar durante el primer año de carrera y a la familia Mundo por el apoyo en múltiples formas.

A mis amigos: los de toda la vida, Lalo y Abraham; los del bacho, Beto Jaimez, Eve y Memo Vargas; los de casa, Richie y Alan; los de la Facultad de Ciencias, Tania, Memo Turner, Vanessa, Mariana y los integrantes de las aguerridas abejas del Lecce, especialmente al buen Petrone.

Esta tesis fue finalizada en medio de una terrible crisis nacional y con la preocupación por un futuro incierto.

Esperando que la organización no cese. Que el estrecho presente se transforme en un amplio futuro.

¡Vivos se los llevaron, vivos los queremos!

ÍNDICE

Resumen

1. INTRODUCCIÓN	10
1.1 Estudio de la diversidad genética	10
1.2 Estructura genética poblacional y patrones espaciales	13
1.3 Genética del paisaje	15
1.3 Genómica ecológica	18
1.4 Coníferas. Biología, evolución y ecología	22
1.5 Coníferas. Estructura genética, ambiente y evolución molecular	24
1.6 Historia natural del Ajusco	29
2. OBJETIVO GENERAL	36
2.1 Objetivos particulares	36
3. JUSTIFICACIÓN	37
4. HIPÓTESIS	39
5. MÉTODO	40
<i>Análisis moleculares</i>	
5.1 Obtención, extracción y amplificación de muestras	40
5.2 Selección de fragmentos y reproducibilidad de AFLPs	42
5.3 Amplificación de microsatélites de cloroplasto	43
<i>Análisis estadísticos</i>	
5.4 Diversidad genética	44
5.5 Estructura genética	45
5.6 Modelo causal	47

5.7 Detección de loci bajo selección natural	48
5.8 Procesos selectivos y diferenciación genética	51
6. RESULTADOS	55
<i>Análisis moleculares</i>	
6.1 Extracción de muestras, amplificación y selección de marcadores AFLPs	55
<i>Análisis estadísticos</i>	
6.2 Diversidad genética	56
6.3 Estructura genética	58
6.4 Modelo causal	63
6.5 Detección de loci bajo selección natural	65
6.6 Procesos selectivos y diferenciación genética	67
7. DISCUSIÓN	69
8. CONCLUSIÓN	85
9. ANEXOS	87
Anexo 1. Recolección de muestras	87
Anexo 2. Métodos de laboratorio para la obtención de AFLPs	91
Anexo 3. Variables ambientales presentes en la zona de estudio	96
Anexo 4. Análisis de reproducibilidad para construir la matriz binaria final	97
Anexo 5. Detección de loci bajo selección natural	98
Anexo 6. Procesos selectivos y diferenciación genética-migración	101
9. REFERENCIAS	103

RESUMEN

Méndez-González, I. D. (2014) Efectos paisajísticos en la distribución de la diversidad genética en una población de *Abies religiosa*. Tesis Profesional (Biología). Facultad de Ciencias, Universidad Nacional Autónoma de México, México, D. F. Pp – 114.

Debido a que la diversidad genética es resultado tanto de procesos contemporáneos como históricos, su estudio permite realizar inferencias sobre la historia evolutiva de las poblaciones. Las diferentes condiciones donde se distribuyen las poblaciones naturales pueden provocar que partes de éstas queden parcialmente aisladas, lo cual moldea la variación genética de manera diferencial, provocando una estructura genética. Asimismo, esta estructura influye en el papel de las diferentes fuerzas evolutivas.

Adoptando un enfoque de escala geográfica fina, utilizando dos tipos de marcadores moleculares con diferente cobertura, resolución y tipo de herencia, y combinándolos con datos ambientales y espaciales, el presente estudio tuvo como objetivos:

- i) describir la variación y estructura genética de la población de *Abies religiosa* del volcán Ajusco.
- ii) identificar las determinantes y describir el papel de las diferentes fuerzas evolutivas en dicha estructura genética.
- iii) contrastar la eficacia de la dispersión del polen y la de las semillas.
- iv) detectar loci con señales de selección natural y determinar la importancia de estos en la diferenciación genética.

La población estudiada, se caracterizó por tener una diversidad genética elevada, tanto nuclear como cloroplástica; y estar subdividida espacialmente en cuatro grupos genéticos y dos linajes de cloroplasto. Ni la distancia geográfica, ni la distancia ambiental, fueron los responsables de la estructura genética observada. En cambio, ésta fue explicada por una dinámica metapoblacional que, probablemente, ha sido provocada por la complicada historia natural del volcán Ajusco. La dispersión del polen mostró ser más efectiva que la de las semillas. Se detectaron 14 loci candidatos a ser adaptativos y relacionados, principalmente, con el tipo de suelo y el tipo de vegetación. Por último, la divergencia genética observada

entre las localidades fue resultado tanto de la baja capacidad de dispersión de la especie, como de procesos selectivos que limitan, en cierto grado, el establecimiento de los migrantes.

Palabras clave: diversidad genética, estructura genética, flujo genético, selección natural, adaptación local, metapoblación, *Abies religiosa*, Ajusco.

1. Introducción

1.1 Estudio de la diversidad genética

La diversidad genética se define como la cantidad de variabilidad genética existente en una población, especie o grupo de especies emparentadas, y constituye una fuente de biodiversidad muy importante (Hughes *et al.*, 2008). Debido a que ésta es resultado tanto de procesos contemporáneos como de eventos históricos de naturaleza diversa (por ejemplo, expansión y disminución del tamaño poblacional, eventos de migración y colonización, adaptación a nuevos ambientes, etc.), su estudio resulta de suma relevancia, ya que permite obtener información sobre la historia evolutiva de la(s) población(es). Asimismo, la cantidad y distribución de la diversidad genética dentro de una población es moldeada por diferentes fuerzas evolutivas como la selección natural, la deriva genética, la mutación, la endogamia y el flujo genético (Hedrick, 2011). Cada una de estas fuerzas puede tener efectos particulares generales; por ejemplo, la deriva genética y la endogamia reducen el nivel de variación, mientras que la mutación lo incrementa. Por otro lado, la selección natural y el flujo genético pueden tanto aumentar como reducir la variación genética, dependiendo de la situación particular específica (Hedrick, 2011).

El estudio de la variación genética emplea diversos marcadores moleculares. Las diferentes características de cada uno, como la tasa de mutación y el tipo de herencia (biparental, uniparental, genoma haploide, genoma diploide), permiten extraer información distinta, por lo que es importante incluir más de un tipo de marcador molecular dependiendo de los objetivos del estudio. Por ejemplo, incluir marcadores moleculares neutrales y no neutrales permitiría identificar los procesos que determinan los patrones de variación genética (Orsini *et al.*, 2013). Asimismo, incluir marcadores con tipos de herencia contrastantes, como pueden ser marcadores nucleares y citoplásmicos, permite analizar la contribución al flujo genético a partir de la dispersión por polen o por semillas (Heuertz *et al.*, 2003).

Los polimorfismos de tamaño de fragmentos amplificados (AFLPs por sus siglas en inglés) se encuentran entre los marcadores moleculares más informativos (Gaudel *et al.*, 2004). Esta

técnica (Vos *et al.*, 1995) se basa en la amplificación selectiva de fragmentos de restricción obtenidos a partir de la digestión del ADN genómico total. Debido a que para su desarrollo y utilización no se requiere información previa acerca del genoma de la especie de interés, es posible obtener un gran número de marcadores dominantes (no es posible distinguir individuos homocigos de heterocigos) y bialélicos (para un tamaño dado, el fragmento se codifica como presente o ausente; Gaudel *et al.*, 2004) en una sola reacción. A pesar de la naturaleza dominante de estos marcadores, el gran número de fragmentos generados permite superar el rendimiento de algunos marcadores codominantes como los microsatélites (Meudt & Clarke, 2006). Por otro lado, comparado con otros sistemas dominantes como los polimorfismos de amplificación aleatoria y las repeticiones de secuencias simples (por sus siglas en inglés RAPDs e ISSRs, respectivamente) los AFLPs tienen un mejor rendimiento, debido a su mayor reproducibilidad, robustez y calidad informativa (Meudt & Clarke, 2006). Debido a estas ventajas, estos marcadores han sido utilizados en diversas aplicaciones, por ejemplo, el mapeo genético, la identificación de híbridos, el análisis de la diversidad genética, la reconstrucción de filogenias recientes (Meudt & Clarke, 2006) y el estudio de la divergencia adaptativa (Bonin *et al.* 2006). Este último punto es de gran importancia, ya que la gran cobertura genómica de los AFLPs permite incluir regiones que podrían estar en un fuerte desequilibrio de ligamiento con otras bajo selección, y así estudiar procesos de adaptación local incluso en especies no modelo (Bonin *et al.* 2006).

Entre las desventajas de estos marcadores se encuentran la naturaleza anónima de los fragmentos amplificados y la poca especificidad inherente a la técnica, lo que puede provocar homoplasia (Caballero *et al.*, 2008), es decir, fragmentos no homólogos pueden presentar un mismo estado de carácter como resultado de un evento de comigración (estado de carácter compartido “1”) o a pérdidas independientes del fragmento (estado de carácter compartido “0”) (Meudt & Clarke, 2006).

La homoplasia en los AFLPs ocurre, principalmente de dos maneras: i) homoplasia por tamaño, que es aquella que ocurre cuando dos fragmentos de ADN pertenecientes a individuos diferentes, presentan el mismo tamaño aunque su origen genómico sea distinto (i.e. representan dos loci distintos), y son codificados como una misma banda (Gort *et al.*,

2009). La magnitud de la homoplasia por tamaño puede ser de hasta el 13% en comparaciones intraespecíficas (Caballero *et al.*, 2008). ii) La homoplasia por colisión, que ocurre cuando en un mismo individuo se tienen fragmentos de distinto origen genómico (i.e. distintos loci) pero de igual tamaño. Estos fragmentos migran hacia el mismo sitio, por lo que son considerados como un solo fragmento (Gort *et al.*, 2009). Se ha reportado hasta un 30% de bandas no detectadas debido a eventos de colisión (Vekemans *et al.*, 2002). La homoplasia por tamaño o por colisión puede provocar sesgos importantes en la estimación de la diversidad, la diferenciación y la estructura genética (Gort *et al.*, 2009). Principalmente se subestima la heterocigosis, debido a la sobre estimación de la frecuencia del alelo dominante, lo que a su vez puede afectar los cálculos de estadísticos de diferenciación poblacional, como F_{ST} (Caballero *et al.*, 2008). Debido a estas características, es necesario tomar ciertas precauciones al utilizar la técnica de AFLPs. Por ejemplo, es aconsejable evitar el uso de fragmentos de tamaño pequeño (<100 pares de bases) y revisar el número de fragmentos amplificados por cada cebador; entre mayor sea el número, mayor será la probabilidad de que haya eventos de homoplasia (Vekemans *et al.*, 2002; Caballero *et al.*, 2008).

Los microsatélites están formados por unidades de motivos cortos de nucleótidos repetidos en tándem. A diferencia de los AFLPs, los microsatélites requieren un conocimiento previo del genoma de la especie para diseñar los cebadores apropiados. La principal ventaja de estos marcadores es que son (generalmente) codominantes, neutrales y multialélicos, por lo que a menudo muestran una gran variación entre individuos (Gaudel *et al.*, 2004). La popularidad de este marcador se debe en gran medida al potencial que poseen para realizar estimaciones de migración contemporánea o el poder para distinguir entre patrones de panmixia y tasas altas de migración. También permiten estimar las relaciones entre individuos (Selkoe & Toonen, 2006) y calcular su tasa de mutación (Gaudel *et al.*, 2004), esto último a diferencia de los AFLPs.

La diversidad genética calculada a partir de diferentes marcadores moleculares depende de factores específicos (tasa de mutación y homoplasia) y no específicos del marcador (deriva y flujo genético; Mariette *et al.*, 2001). Por ello es de esperarse que utilizando dos tipos de

marcadores moleculares los niveles de diversidad genética difieran al hacer comparaciones interpopulacionales, aunque ambos tipos de marcadores deberían mostrar tendencias similares dentro de cada población (Marette *et al.*, 2001). Sin embargo, no todos los estudios empíricos donde se ha comparado la diversidad y la diferenciación genética utilizando AFLPs y microsatélites han corroborado dicha idea (Gaudeul *et al.*, 2004). Por ejemplo, al analizar la diversidad genética dentro y entre poblaciones de *Pinus pinaster* se observó una gran diferenciación poblacional con ambos marcadores, debida en mayor parte a la presencia de dos provincias geográficas, lo que indica que en esta especie la diferenciación poblacional se debe en mayor medida a factores no específicos del marcador (Marette *et al.*, 2002). Por otro lado, un estudio realizado con *Eryngium alpinum* reveló ciertas discrepancias entre los resultados de cada marcador (Gaudeul *et al.*, 2004). Las medidas de diversidad genética inter e intrapoblacionales mostraron poca correlación, mientras que el orden jerárquico de las poblaciones con base en su diversidad genética no coincidió entre ambos marcadores (Gaudeul *et al.*, 2004). De esta forma, las diferencias en la precisión de los patrones detectados por cada marcador y su relación con la escala espacial revelan la importancia de las características específicas del marcador (Gaudeul *et al.*, 2004).

1.2 Estructura genética poblacional y patrones espaciales

Las poblaciones naturales tienen a menudo historias complejas que consisten en fases de expansión y contracción del tamaño poblacional, mientras que los patrones de migración entre poblaciones pueden fluctuar en el tiempo y el espacio. Cuando partes de una población se ven parcialmente aisladas debido, por ejemplo, a barreras geográficas o a cambios en el medio ambiente, las fuerzas estocásticas y/o selectivas moldean la variación genética de manera diferencial, lo que a su vez afecta los patrones de flujo genético y apareamiento. Dichas poblaciones pueden representarse como grupos entre los cuales la variación genética difiere claramente, existiendo una estructura genética poblacional (Corander *et al.*, 2008). Debido a que el flujo genético cuantifica la dispersión efectiva, es decir los eventos de migración que resultan en reproducción exitosa, éste se ha considerado como la principal fuerza evolutiva en el mantenimiento o la erosión de la estructura genética. Por lo tanto, su estudio resulta de sumo interés para entender procesos microevolutivos en poblaciones

fragmentadas (Munshi-South, 2012), ya que el nivel de flujo genético determina hasta qué grado cada población de una especie constituye una unidad evolutivamente independiente (Slatkin, 1993).

Existen diversas aproximaciones para cuantificar la diferenciación al interior de las poblaciones naturales (Hedrick, 2011). Los estadísticos F funcionan como una herramienta para descomponer la diversidad genética dentro y entre poblaciones (Wright, 1951). Estos tienen especial relevancia ya que demuestran que la cantidad de diferenciación genética entre poblaciones se relaciona de manera predecible con las tasas de ciertos procesos evolutivos, como la mutación, el flujo genético y la deriva genética (Holsinger & Weir, 2009). F_{ST} es uno de los estadísticos descriptivos más utilizados en la genética de poblaciones y se define como la proporción de la diversidad genética debida a las diferencias en las frecuencias alélicas entre subpoblaciones con relación a la población total (Holsinger & Weir, 2009).

En años recientes al avance en métodos analíticos ha llevado a la definición de diversos patrones que describen la estructura genética en relación a los procesos y variables que la determinan. Diferentes procesos pueden llevar a un mismo patrón, sin embargo, la correcta interpretación del patrón observado depende de identificar la variable de mayor importancia durante el proceso subyacente (Orsini *et al.*, 2013). A continuación se describen algunos de estos patrones:

- i) aislamiento por distancia (IBD; Wright, 1943). Este patrón se define como el incremento en la diferenciación genética, como resultado de la disminución en el flujo genético a medida que aumenta la distancia geográfica (Orsini *et al.*, 2013). Para que esto ocurra se supone que las poblaciones están en equilibrio deriva-flujo genético (en las cuáles el efecto de fundador, es decir, el efecto en la diversidad genética debido al reducido tamaño poblacional inicial, ha sido erosionado) y no están afectadas por la selección natural (Wright, 1943).
- ii) aislamiento por resistencia (IBR; McRae, 2006). Este cuantifica la resistencia al flujo genético a través de diversos hábitats; a mayor resistencia a la dispersión, menor flujo genético y por lo tanto mayor diferenciación (Wang, 2013). En este patrón se considera explícitamente la heterogeneidad espacial, así como la forma del rango de distribución de la

especie; al hacer esto en lugar de considerar el hábitat como ilimitado o dividido simplemente por barreras discretas, las predicciones sobre el flujo genético pueden ser muy precisas (McRae & Beier, 2007).

iii) aislamiento por adaptación (IBA; Nossil *et al.*, 2009). Este patrón asocia la variación genética con determinados hábitats y/o variables ambientales como resultado de la selección natural en contra del migrante (Andrew *et al.*, 2012). Cuando la divergencia adaptativa promueve barreras al flujo genético entre poblaciones de sitios ecológicamente divergentes, la diferenciación genética puede darse aún en loci neutrales, debido a la deriva genética; por lo que no queda restringida a aquellos loci bajo selección (Orsini *et al.*, 2013).

iv) aislamiento por colonización (IBC; Orsini *et al.*, 2013). Este patrón es generado por eventos fundador y resulta difícil de interpretar ya que puede tener causas ecológicas y/o evolutivas. Por ejemplo, la ventaja numérica de los primeros colonizadores puede traducirse en adaptaciones locales que a la vez disminuyen la probabilidad de establecimiento de nuevos migrantes, de esta forma se refuerza el efecto de fundador debido a la reducción en el flujo genético, lo que conlleva a un escenario de monopolización (Orsini *et al.*, 2013). Otro escenario es la colonización serial (más de un evento de colonización) en este caso se espera una correlación positiva entre la diferenciación genética en loci neutrales y la distancia geográfica, dicha correlación podría o no observarse en loci bajo selección divergente, dependiendo de la fuerza de los procesos de adaptación local (Orsini *et al.*, 2013).

1.3 Genética del paisaje

Los patrones mencionados anteriormente surgen de la unificación de datos genéticos y ambientales. Recientemente este segundo punto ha tenido un gran avance, debido sobre todo al desarrollo de los sistemas de información geográfica (SIG), que permiten capturar, almacenar, manipular, analizar y desplegar en todas sus formas la información geográficamente referenciada. La genética del paisaje se basa justamente en este tipo de datos paisajísticos y busca entender los procesos microevolutivos que influyen en la distribución espacial de la variación genética intraespecífica (Manel *et al.*, 2003). El método general utilizado en la genética del paisaje consta de dos pasos importantes: i) la detección

de discontinuidades genéticas y ii) su correlación con características ambientales y/o físicas del paisaje (por ejemplo, gradientes de temperatura o la presencia de barreras físicas o ecológicas) (Manel *et al.*, 2003).

Diversas herramientas se han desarrollado para llevar a cabo cada paso. Por ejemplo, los métodos bayesianos de agrupamiento (Pritchard *et al.*, 2000) determinan la estructura genética de la población basándose en genotipos individuales (Anderson *et al.*, 2010); de esta forma permiten realizar conclusiones *a posteriori* acerca de la intensidad y/o direccionalidad del flujo genético. Algunos de estos métodos de agrupamiento tienen, además, la capacidad de incluir información sobre la distribución espacial de los individuos para mejorar la resolución de los datos (Corander *et al.*, 2008). La inferencia, hecha dentro de un marco bayesiano, acerca del número de grupos genéticos y su ubicación espacial se basa en simulaciones (Guillot *et al.*, 2005). Para llevarla a cabo se calcula la probabilidad de encontrar cierto número de grupos genéticos (k) con frecuencias alélicas específicas dados los genotipos observados. La k con la mayor probabilidad posterior será la elegida.

Geneland (Guillot *et al.*, 2005) es un paquete para la plataforma R (R Core Team, 2013) que utiliza datos geográficos y genómicos para inferir, dentro de un marco bayesiano, el número de grupos genéticos y la localización espacial de las discontinuidades genéticas entre dichas poblaciones (Guillot *et al.*, 2005). El modelo estadístico empleado considera el total de los individuos como una serie de poblaciones en equilibrio Hardy-Weinberg, cuyos loci están en equilibrio de ligamiento, organizadas espacialmente en forma de una teselación Poisson-Voronoi y con una distribución espacial azarosa de los individuos dentro de cada una de ellas (Guillot *et al.*, 2005). Otro programa bayesiano de agrupamiento que puede utilizar coordenadas geográficas es BAPS (Corander *et al.*, 2008). Este utiliza el algoritmo espacial de triangulación de Delaunay para conectar a los individuos o sitios de muestreo a partir de su localización. El modelo estadístico involucra tres parámetros: una gráfica que especifica los vecinos de cada individuo, un parámetro que brinda el peso específico a la información espacial en el esquema de inferencia y el número de grupos genéticos (Corander *et al.*, 2008). Posteriormente es posible esquematizar el flujo genético mediante flechas ponderadas, cuyo “peso” es proporcional a la cantidad de ancestría relativa entre el grupo

genético que dona los migrantes y el que los recibe (Corander *et al.*, 2008).

Varios análisis utilizando datos simulados y empíricos han demostrado la confiabilidad de los métodos espaciales para la detección de límites genéticos (Safner *et al.*, 2011). Sin embargo, la presencia de patrones fuertes de aislamiento por distancia puede provocar una sobrestimación del número de grupos genéticos (Safner *et al.*, 2011). A pesar de esta limitación, dichos métodos son muy confiables bajo distintos escenarios (Safner *et al.*, 2011) e incluso cuando se tienen niveles de diferenciación poblacional relativamente bajos (Latch *et al.*, 2006), por lo que han adquirido mucha popularidad. Por ejemplo, utilizando Geneland (Guillot *et al.*, 2005) se examinó el efecto de las barreras geográficas sobre la dispersión y la diversidad genética neutral en cuatro localidades arrecifales de los peces *Diplodus vulgaris*, *Mullus surmuletus*, *Serranus cabrilla*, *Oblada melanura*, *Tripterygion delaisi*, *Apogon imberbis* y *Symphodus tinca* (Galarza *et al.*, 2009). En *D. vulgaris*, *M. surmuletus* y *S. tinca* la ubicación del límite genético coincide con la presencia de un primer frente oceánico, mientras que en *A. imberbis* y *O. melanura* coincide con la presencia de otro frente oceánico; en *S. Cabrilla* y *T. delaisi* el límite genético no coincide con ninguna barrera física (Galarza *et al.*, 2009). Este estudio comparativo demuestra la capacidad de los métodos Bayesianos de agrupamiento para identificar la ubicación espacial de barreras geográficas y su efecto en la diversidad genética, a la vez que sugiere la existencia de otros factores ecológicos o conductuales que afectan la dispersión.

Otro método utilizado para analizar el flujo genético, está basado en la teoría de circuitos, y brinda especial relevancia a la conectividad entre poblaciones (McRae, 2006). En él, se estructura espacialmente el flujo genético de manera similar a una corriente eléctrica, donde concurren materiales de gran conductancia y de alta resistencia (McRae & Beier, 2007). Dicho enfoque permite evaluar las características físicas del paisaje que aumentan o limitan el flujo genético. Este método ha mostrado un mejor desempeño al compararse con pruebas de aislamiento por distancia o de ruta de menor costo; este último modela el costo relativo al movimiento entre dos hábitats asumiendo la existencia de una ruta óptima (Penrod *et al.*, 2006). Esto debido a que el flujo genético entre poblaciones reales puede ocurrir por más de una ruta e incluso realizarse de manera indirecta a través de varias generaciones (McRae &

Beier, 2007).

Otra forma de analizar las determinantes de la estructura genética es realizar correlaciones, utilizando la prueba de Mantel (1967) y de Mantel parcial (Bohonak 2002), entre las distancias genéticas y las geográficas y/o ambientales. Con estas pruebas se analiza la correlación lineal entre gradientes ambientales y la variación genética o geográfica. Recientemente, con el incremento del conocimiento ecológico de muchos sistemas naturales, también se han utilizado para estudiar la correlación entre distancias ecológicas y genéticas. Sin embargo, estas pruebas solo son apropiadas cuando la hipótesis bajo estudio puede ser formulada en términos de distancias (Orsini *et al.*, 2013).

1.4 Genómica ecológica

El medio ambiente natural donde se desarrollan todos los organismos presenta una multitud de desafíos bióticos y abióticos, los cuales requieren respuestas que pueden ser en forma de modificaciones en rasgos bioquímicos, fisiológicos, genéticos, morfológicos y/o conductuales, en tiempos tanto ecológicos como evolutivos (Ungerer *et al.*, 2008). Para poder adquirir un conocimiento profundo sobre éstas se requiere una aproximación multidisciplinaria, que combine análisis orgánicos, genéticos, genómicos y experimentos de laboratorio, todo dentro de un contexto ecológico (Ungerer *et al.*, 2008). Una de estas aproximaciones, la genómica ecológica, se define como un campo integrativo de conocimiento que busca entender los mecanismos genéticos que subyacen a las respuestas de los organismos hacia su ambiente natural (Ungerer *et al.*, 2008). Uno de los objetivos de este campo de conocimiento es detectar los loci responsables de las adaptaciones locales (Mosca *et al.*, 2014), es decir, de aquellas que surgen de la interacción entre diferentes tipos de selección, la cantidad de variación genética, el flujo genético y la heterogeneidad ambiental (Alberto *et al.*, 2013). Esto permite, por ejemplo, predecir la respuesta de los organismos ante el cambio global para así generar herramientas de manejo y conservación (Mosca *et al.*, 2012).

A pesar del creciente interés en esta área de investigación, aún quedan muchas preguntas respecto a las bases moleculares y genéticas que subyacen la evolución adaptativa (Jones

et al., 2012). Sin embargo, estudios recientes han comenzado a responder estas interrogantes. Por ejemplo, estudiando poblaciones naturales de *Arabidopsis thaliana* se identificaron señales de adaptación relacionadas con variables climáticas como la temperatura, la precipitación, la radiación fotosintéticamente activa, la humedad relativa, la duración de cada estación y la aridez (Hancock *et al.*, 2011). En otro estudio, utilizando diferentes ecotipos de esta especie, se detectaron variantes alélicas relacionadas con la supervivencia promedio y el tiempo de vida del fruto (Fournier-Level *et al.*, 2011). Estas variantes fueron relacionadas positivamente con la adecuación en distintos sitios experimentales, lo que demuestra su carácter adaptativo (Fournier-Level *et al.*, 2011).

Para identificar loci con señales de selección natural y evaluar el papel de ésta en la variación genética se utilizan generalmente dos aproximaciones. La primera de ellas se basa en el estadístico de diferenciación entre subpoblaciones F_{ST} . La idea subyacente es que, mientras los procesos demográficos afectan a todo el genoma, la selección solo actúa sobre aquellos loci involucrados en procesos adaptativos (Pannell & Fields, 2014). Considerando esto, es posible generar una distribución de valores de F_{ST} esperados bajo procesos demográficos determinados, para después detectar aquellos cuyos valores son *outliers* (Lewontin & Krakauer, 1973). De esta forma, loci bajo selección direccional mostrarán valores de F_{ST} significativamente elevados, mientras que aquellos bajo selección balanceadora mostrarán valores significativamente bajos (Excoffier & Lischer, 2011). Uno de los primeros métodos propuestos consiste en calcular la varianza esperada para F_{ST} bajo diferentes distribuciones de las frecuencias alélicas, posteriormente se pone a prueba si la varianza observada puede ser obtenida en un escenario de neutralidad (Lewontin & Krakauer, 1973). Otro método obtiene la distribución de F_{ST} a partir de un modelo de islas; aquellos loci ubicados en los extremos de la distribución generada serán considerados *outliers* (Beaumont & Nichols, 1996). Los métodos antes mencionados (Lewontin & Krakauer, 1973, Beaumont & Nichols, 1996) asumen que todas las poblaciones muestreadas son independientes o contribuyen en igual medida a la poza de migrantes (Excoffier *et al.*, 2009). El modelo jerárquico de islas (Slatkin & Voelm, 1991), implementado en el programa Arlequin (Excoffier *et al.*, 2005), toma en cuenta la heterogeneidad de las afinidades entre poblaciones; y asume que el flujo genético es mayor entre poblaciones dentro del mismo grupo genético (Excoffier & Lischer,

2011).

Por otro lado, también se han formulado aproximaciones basadas en estadística bayesiana (Excoffier *et al.*, 2005). Por ejemplo, Bayescan (Foll & Gaggiotti, 2008) es un programa basado en el estadístico F_{ST} que utiliza un escenario basado en un modelo de islas, en el cual las frecuencias alélicas de cada subpoblación están correlacionadas debido a la migración desde una poza génica común, de la cual difieren en distinto grado. En este método el coeficiente de F_{ST} obtenido se descompone en un componente población-específico, compartido por todos los loci, y otro componente locus-específico, compartido por las poblaciones. Por lo tanto, si para explicar el patrón de diversidad de un locus determinado es necesario utilizar el componente locus-específico, entonces dicho locus se considera un *outlier* (Foll & Gaggiotti, 2008).

La segunda aproximación para evaluar el papel de la selección natural consiste en la asociación entre variantes alélicas determinadas y hábitats particulares o variables ambientales específicas; de esta forma se resalta la posible existencia de genes implicados en procesos adaptativos (Pannell & Fields, 2014). Existen diversos métodos basados en esta aproximación. El más sencillo de ellos busca detectar la correlación entre las frecuencias alélicas entre subpoblaciones con una variable ambiental determinada. Comúnmente, el nivel de significancia es evaluado utilizando el valor de p obtenido a partir de una prueba de Student (Villemereuil *et al.*, 2014). Una extensión de este método es el uso de la regresión logística (Bonin *et al.*, 2007). Este tipo de regresión describe la asociación entre una variable de respuesta cualitativa que sigue una distribución binomial y una o más variables predictivas, en este caso variables ambientales (Bonin *et al.*, 2007); por lo tanto, resulta perfectamente adaptable para marcadores dominantes, ya que estos tienen solo dos estados posibles (presencia y ausencia) y por consiguiente presentan una distribución binomial. Posteriormente se puede utilizar el estadístico G y/o la prueba de Wald (Hosmer & Lemeshow, 2000) para examinar el nivel de significancia de la asociación (Bonin *et al.*, 2007). Otro método que además toma en cuenta la estructura poblacional es el de Modelos Mixtos de Factores Latentes (FLMM por sus siglas en inglés) (Frichot *et al.*, 2013). Este algoritmo combina análisis de componentes principales con una regresión (Villemereuil *et al.*, 2014),

de esta manera la variación genética es dividida en aquella que es explicada por variables ambientales y aquella que es resultado de la historia demográfica de la especie; esto teóricamente reduce la tasa de falsos positivos, es decir la tasa a la cual los loci neutrales son erróneamente considerados selectivos (Lotterhos & Withlock, 2014), cuando se tienen escenarios de aislamiento por distancia (Frichot *et al.*, 2013). Para más información acerca de las características y el desempeño de distintos métodos basados en ambos tipos de aproximaciones véase De Mita *et al.* (2013), Villemeuril *et al.* (2014) y Lotterhos & Withlock (2014).

Comparando tres métodos basados en F_{ST} y cinco basados en correlaciones ambientales se encontró que el segundo tipo de aproximación (correlaciones ambientales) tiene un mayor poder estadístico, aunque también presenta una mayor tasa de falsos positivos (De Mita *et al.*, 2013). Así mismo, el rendimiento de cada método depende de varios factores: i) a mayor diferenciación genética y menor intensidad de selección, el poder de todos los métodos evaluados disminuye. Sin embargo, incluso en escenarios con baja intensidad de selección los métodos basados en correlaciones ambientales muestran suficiente poder y no son afectados por el modelo demográfico adoptado (De Mita *et al.*, 2013); ii) a mayor tasa de autofertilización, mayor cantidad de falsos positivos, esto podría deberse al aumento en la varianza genética entre poblaciones debido a la disminución del tamaño poblacional efectivo (De Mita *et al.*, 2013); iii) en general el rendimiento de todos los métodos es mejor en un esquema de muestreo que abarque un gran número de poblaciones y no muchos individuos en pocas poblaciones (De Mita *et al.*, 2013).

En otra comparación de estos métodos simulando escenarios más complejos, como estructuras poblacionales jerárquicas o selección poligénica (Villemeuril *et al.*, 2014) se resalta la necesidad de utilizar métodos más flexibles, como el FLMM, que no dependan de un modelo poblacional particular. También se recomienda el uso de un marco estadístico en el que sea posible poner a prueba cada método en cuanto a su poder en relación a la tasa de falsos positivos. Esto resulta difícil debido a que cada método utiliza un criterio diferente para medir el nivel de significancia estadística, por ejemplo, el valor p utilizado en Arlequin no es equivalente al factor Bayes utilizado en Bayescan (Lotterhos & Withlock, 2014). Sin embargo,

utilizando el concepto de tasa de falso descubrimiento, es decir, la tasa a la cual loci considerados selectivos son en realidad neutrales, es posible medir la significancia estadística en este tipo de estudios (Storey & Tibshirani, 2003). Esta aproximación se basa en el valor q , similar al valor p , que brinda a cada característica su nivel individual de significancia. Mientras que el valor p es una medida de significancia en términos de la tasa de falsos positivos, el valor q es una medida en términos de la tasa de falso descubrimiento (Storey & Tibshirani, 2003). Por ejemplo, una tasa de falsos positivos de 5% significa que, en promedio, 5% de los loci realmente neutrales serán considerados outliers. Mientras que una tasa de falso descubrimiento del 5%, significa que, de los loci considerados selectivos, en promedio 5% son en realidad neutrales (Storey & Tibshirani, 2003). El uso de esta metodología permitiría obtener un balance entre la tasa de falsos positivos y la detección de loci verdaderamente selectivos, además de ser una medida de significancia que puede ser interpretada universalmente (Storey & Tibshirani, 2003).

1.5 Coníferas. Biología, evolución y ecología

Las coníferas son un grupo de plantas con semilla que descienden de un ancestro común en el Paleozoico tardío. Hoy en día, existen aproximadamente 550 especies de estas plantas, distribuidas por todo el mundo, con excepción de la Antártida. El nombre significa portador de conos, sin embargo, la inclusión de las semillas dentro de un cono no es una característica de todas las coníferas. Si bien no existe una singularidad que una a todas estas especies y las separe del resto de las plantas con semillas, existe amplia evidencia de su identidad y diferenciación como grupo natural. Por ejemplo, la estructura particular de los conos, cuando están presentes, es única de las coníferas; las características de la madera, como la composición del tejido conductor de agua (únicamente con traqueidas), que las diferencia de las plantas con flor y de las gnetales, así como la estructura particular de sus granos de polen (Eckenwalder, 2009).

Las coníferas más antiguas conocidas datan del período Carbonífero tardío (Eckenwalder, 2009). Durante el período Pérmico éstas se volvieron más abundantes, dominando al menos en parte, los bosques secos de tierras bajas característicos de este período; sin embargo,

ninguna de estas especies pertenece a alguna de las seis familias existentes hoy en día (Eckenwalder, 2009). Durante el periodo Triásico, siguiendo las extinciones del Pérmico, las coníferas continuaron diversificándose. Al final de este periodo, se encuentran fósiles pertenecientes a tres de las familias actuales: Araucariaceae, Cupressaceae y Podocarpaceae. Durante el Mesozoico, los bosques de coníferas se volvieron dominantes; sin embargo, la primer evidencia incontrovertible de miembros con las características completas de las familias existentes restantes (Pinaceae y Sciadopityaceae) solo surge hasta el Cretácico, hace unos 145 millones de años (Eckenwalder, 2009).

Alrededor del mundo las coníferas se localizan en una gran variedad de hábitats, dominando especialmente los bosques boreales del hemisferio norte y continuando hacia el sur a través de cadenas montañosas de gran altitud (Rushfort 1987). Una característica de estas regiones es un invierno largo y seco y un verano corto y frío. Las principales adaptaciones que permiten la dominancia de las coníferas en estas regiones son la tolerancia a la falta de humedad, y la capacidad para conservar agua y llevar a cabo un crecimiento acelerado durante la temporada favorable (Rushfort 1987). Por el contrario, en regiones donde la precipitación supera los 100 cm anuales, como los bosques húmedos de la costa oeste de los Estados Unidos, también pueden ser dominantes; sin embargo en estas regiones el verano puede ser muy cálido y el potencial de evapotranspiración muy alto, por lo tanto la capacidad de tolerar períodos secos brinda a las coníferas una ventaja sobre otras plantas (Rushfort 1987).

Dentro de la familia Pinaceae se encuentra el género *Abies*. Estos son árboles medianos a grandes, con hojas arregladas generalmente en espiral con forma de aguja, que alcanzan la madurez sexual a los 20 años; viven en promedio 60 años, son monoicos y sus conos masculinos se agrupan en la parte baja de la corona, mientras que los femeninos lo hacen en la parte superior. Estos son polinizados por el viento y se rompen al llegar a la madurez, liberando numerosas escamas ovulíferas con dos semillas aladas cada una, que también son diseminadas por el viento. Todos los *Abies* son diploides con un número cromosómico de $(2n) 24$ (Eckenwalder, 2009). Los diversos abetos (nombre común del género) se encuentran cercanamente emparentados y muestran una apariencia uniforme, por lo que su clasificación

resulta complicada, además existe un componente geográfico importante en su taxonomía; particularmente en zonas como China, México y Guatemala, donde existen serios problemas taxonómicos y de conservación debido al poco conocimiento de las especies y a su explotación descontrolada (Eckenwalder, 2009).

En México los bosques de *Abies* cubren cerca del 0.16% del territorio nacional, estando distribuidos de manera dispersa y localizada. Generalmente se presentan entre los 2,400 – 3,600 msnm, en la mayor parte de los casos restringidos a un cerro, una ladera o una cañada. Esta distribución se explica debido a los altos requerimientos de humedad de los abetos mexicanos (Rzedowski, 2006). Los bosques de *Abies* en México se encuentran en sitios con características climáticas muy particulares, que incluyen una estación seca bien definida y nevadas moderadas o incluso ausentes. Esto tiene consecuencias en la fenología de la comunidad, ya que las actividades de fotosíntesis, absorción y transpiración prácticamente no son interrumpidas a lo largo del año. Las áreas de mayor extensión se encuentran en el Eje Neovolcánico Transversal; en la Sierra Madre Occidental y Oriental los bosques de *Abies* son más escasos, mientras que en la Sierra Madre del Sur solo se encuentran manchones aislados, sobre todo en los estados de Guerrero y Oaxaca (Rzedowski, 2006). Dependiendo del autor, en México se reconocen de seis a diez especies (Martínez, 1948; Liu, 1971; Rushfort, 1987; Farjon, 1990) con un alto grado de endemismo (Rzedowski 2006). La especie más común a lo largo del Eje Neovolcánico Transversal, en el centro del país, es *Abies religiosa* (H. B. K.) Schl. Et Cham. Ésta tiene una distribución mayoritariamente continua entre los 2,000 a 3,500 msnm. Los suelos donde se establece el oyamel (nombre común con que se conoce esta especie en los estados del centro) se caracterizan por ser muy jóvenes, de origen volcánico y con pendientes muy pronunciadas (CONAFOR). Los bosques de esta especie tienen una gran importancia, ya que son los hábitats preferidos por la mariposa monarca (*Danaus plexippus*) durante la época invernal (Anderson & Browler 1996). Su uso comercial más importante es como “árbol de navidad”.

1.6 Coníferas: Estructura genética, ambiente y evolución molecular

Las coníferas han experimentado numerosos cambios ambientales a lo largo de su historia

ecológica y evolutiva (Hamrick, 2004). Las expansiones territoriales que han ocurrido desde el final de las últimas glaciaciones y la presencia de adaptaciones en áreas recientemente colonizadas resaltan la capacidad de estos organismos para adaptarse en respuesta a fuerzas selectivas que pueden variar geográficamente (Mosca *et al.*, 2014). La capacidad de dispersión tanto del polen como de las semillas, así como su tiempo generacional largo hacen de las coníferas un valioso modelo de estudio para investigar la diferenciación genética en presencia de flujo genético, así como el papel de la selección natural dependiente del medio ambiente en la generación de estructura genética poblacional (Mosca *et al.*, 2014). El flujo genético tiene un papel muy importante en este proceso ya que, por un lado, podría retrasar ciertos procesos de divergencia ecológica, aún aquellos que implican una fuerte selección contra el migrante, mientras que por otro, podría introducir alelos ventajosos en la población, posibilitando adaptaciones locales (Nossil *et al.*, 2009).

Diversos estudios han examinado las determinantes de la diversidad y la estructura genética en coníferas. Por ejemplo, a nivel intrapoblacional se han descrito patrones de estructura genética a escala fina, la cual se pierde al analizar árboles adultos (González-Martínez *et al.*, 2002). Este estudio es muy importante porque analiza el flujo genético *sensu stricto*; es decir, toma en cuenta eventos posteriores a la dispersión, como la adaptación a nuevos hábitats. Por otro lado, la distancia geográfica ha sido identificada como una de las principales determinantes de la estructura genética en diversas especies de coníferas (Cuenca *et al.*, 2003; Mimura & Aitken 2007; Delgado *et al.*, 2007; Mosca *et al.*, 2014); mientras que en otras especies, ciertas variables ambientales han demostrado una mayor importancia, por ejemplo: la precipitación invernal (Mosca *et al.*, 2012 y 2014), la temperatura al inicio de la primavera (Mosca *et al.* 2014), la temperatura máxima estacional (Mosca *et al.* 2012) y la aridez (Eckert *et al.*, 2010).

Estudiando procesos de adaptación local en coníferas se han encontrado loci con señales de selección natural correlacionados con distintas variables ambientales, por ejemplo: la temperatura global (Eckert *et al.* 2010), máxima y mínima diaria (Tsumura *et al.*, 2012) y estacional mínima (Mosca *et al.*, 2012); la aridez global y en invierno (Eckert *et al.* 2010); la cantidad de radiación solar global y la profundidad máxima de la nieve (Tsumura *et al.*, 2012);

y la precipitación estacional y en invierno y otoño (Mosca *et al.*, 2012). La mayoría de los genes detectados están relacionados con distintos mecanismos de respuesta al estrés abiótico, como la respuesta al estrés oxidativo (Eckert *et al.*, 2010) o al inicio del frío otoñal (Mosca *et al.*, 2014) e incluyen diversos procesos fisiológicos relacionados con la membrana celular (Eckert *et al.* 2010; Mosca *et al.*, 2014), el metabolismo de azúcares (Eckert *et al.* 2010), la actividad fotosintética, la reparación del ADN (Frichot *et al.*, 2013) y la respuesta a enfermedades (Frichot *et al.*, 2013; Mosca *et al.*, 2014). Estos trabajos resaltan la necesidad de incluir datos ambientales y ecológicos en el estudio de la estructura genética, haciendo posible evaluar la influencia de los procesos adaptativos en la limitación del flujo genético (Orsini *et al.*, 2013).

La complicada historia biogeográfica de México hace difícil discernir entre las distintas fuerzas evolutivas que moldean la variación genética de sus especies, en especial de aquellas con distribución montana. Se ha propuesto que entre hace 12,500 y 9,000 años, al final del Pleistoceno, el clima era más húmedo y frío que actualmente y los bosques templados alcanzaron su máximo desarrollo en el país (Lozano-García, 1993). Se sabe que el género *Abies* ya estaba presente en México durante el Plioceno medio (Graham, 1993). Debido a esto es posible que las poblaciones de *Abies* mexicanos hayan tenido distribuciones históricas más amplias en los períodos más fríos, y hayan quedado aisladas en los interglaciales más cálidos tal como ocurre actualmente (Aguirre-Planter *et al.* 2000, Jaramillo-Correa *et al.*, 2008).

Diversos estudios han examinado las consecuencias de estos eventos históricos en la cantidad y distribución de la diversidad genética (Aguirre-Planter *et al.* 2000, Jaramillo-Correa *et al.*, 2008, Mugica-Gallart, 2013) así como en las tasas de diversificación e identidad taxonómica (Aguirre-Planter *et al.*, 2012) de los *Abies* mexicanos. Utilizando diferentes marcadores moleculares se han encontrado patrones de diversidad contrastantes (Tabla 1). Sin embargo, una característica en común es una gran diferenciación genética entre poblaciones, incluso al compararse con lo reportado en otras coníferas (Aguirre-Planter *et al.*, 2000). Los tres estudios concuerdan en un fuerte efecto de la deriva genética así como una capacidad de dispersión limitada, incluso para el polen, como lo muestra el ADN de

cloroplasto (Jaramillo-Correa *et al.*, 2008). Dichas características podrían sugerir un escenario de aislamiento por distancia. Sin embargo dado que éste solo se demuestra en uno de los estudios (Aguirre-Planter *et al.*, 2000) para el caso de *Abies religiosa*, se sugiere que las poblaciones de esta y las demás especies no han alcanzado un equilibrio deriva genética-migración (Aguirre-Planter *et al.*, 2000, Mugica-Gallart, 2013). En cuanto a la identidad taxonómica, la evidencia sugiere un estado incipiente de especiación, lo que se refleja en la existencia de haplotipos compartidos por la mayoría de las especies (Jaramillo-Correa *et al.*, 2008); es posible que esto se deba a eventos de introgresión durante los períodos glaciales, al flujo genético contemporáneo, a la retención de polimorfismos ancestrales o a una combinación de estos escenarios.

La poca divergencia interespecífica, combinada con la gran diferenciación intraespecífica plantea la necesidad de estudiar más a fondo las determinantes de la distribución de la variación genética en *Abies* mexicanos, incluyendo la influencia de la heterogeneidad ambiental. Los estudios antes mencionados señalan un fuerte impacto de la deriva genética, a la vez que sugieren una capacidad de dispersión limitada. Para el caso del polen, esto contradice la suposición de que especies con polinización por viento tienen una gran cantidad de flujo genético. Lo anterior puede deberse a la escala de los estudios previos, indicando que es necesario analizar estos patrones tanto para el polen como para las semillas a escala más fina, es decir intrapoblacionalmente. Por último, el papel de la selección natural ha sido poco estudiado en los *Abies* mexicanos, la heterogeneidad ambiental dentro de su rango de distribución, la complicada historia demográfica, con fases de contracción y expansión, así como la distribución actual a manera de manchones podrían sugerir la existencia de adaptaciones locales.

La información obtenida a partir de aproximaciones como la genética del paisaje y/o la genómica ecológica es de suma relevancia, ya que permite identificar las causas del aislamiento de una población, así como la importancia de procesos microevolutivos (como las adaptaciones locales) en su supervivencia a largo plazo. Estos resultados posibilitan la formulación de planes de conservación, reforestación, restauración por introducción de diversidad genética (una vez que se conocen las afinidades evolutivas de las poblaciones), o

el reestablecimiento de la conectividad entre poblaciones (en los casos pertinentes) en especies amenazadas como los *Abies* mexicanos. Para formular planes de conservación efectivos es necesario considerar, y entender, el potencial adaptativo de las poblaciones ante los cambios naturales y antropogénicos, tanto los actuales como los históricos. Una forma de hacer esto es estudiar poblaciones en zonas geográficas de las que se tenga conocimiento acerca de su historia natural, así como de las presiones actuales que amenazan su viabilidad. El conocimiento generado a partir de este enfoque permitirá identificar las respuestas históricas que estas poblaciones han tenido, para así poder hacer predicciones sobre su comportamiento frente a las presiones actuales. Las poblaciones *de Abies* mexicanos más amenazadas se encuentran en el Eje Neovolcánico Transversal, ya que ahí se concentra la mayor diversidad humana del país. Lo anterior, aunado al gran conocimiento de la historia glacial de los volcanes del centro del país (hay evidencia de glaciaciones en tres volcanes de esta zona) hacen de estas poblaciones candidatos ideales para este fin.

Tabla 1. Niveles de diversidad en especies de *Abies* mexicanos utilizando diferentes marcadores moleculares.

Marcador	Nivel de diversidad	Explicaciones	Referencia
Isoenzimas	Bajo comparado con el nivel de diversidad registrado en otras especies de <i>Abies</i> y coníferas de zonas templadas.	Distribución geográfica restringida Fuerte efecto de la deriva genética.	(Aguirre-Planter <i>et al.</i> , 2000)
ADN mitocondrial	Bajo, sin embargo se encuentra previamente en otras coníferas mexicanas y de zonas templadas.	Fuerte efecto de la deriva genética.	(Jaramillo-Correa <i>et al.</i> , 2008)
Microsatélites de cloroplasto	Alto y en el mismo rango que lo observado en especies de <i>Abies</i> de Europa y Norteamérica.		(Jaramillo-Correa <i>et al.</i> , 2008)
Microsatélites nucleares	Alto y en el mismo rango que lo reportado en otras coníferas de Norteamérica con distribuciones continuas.	Gran longevidad, entrecruzamiento y dispersión por viento.	(Múgica Gaillard, 2013)

1.7 Historia natural del Ajusco

El volcán Ajusco se encuentra en la parte sur de la provincia fisiográfica de la Mesa Central (Fig. 1). La cuenca tiene una altura promedio de 2,300 msnm y está formada, casi en su totalidad por mantos de lava y materiales piroplásticos (White 1990).



Figura 1 – Ubicación y modelo fisiográfico del volcán Ajusco. Modificada de García-Aguirre *et al.* (2007).

La montaña conocida como Ajusco es en realidad un grupo de picos aislados a unos 1700 m sobre el nivel del piso de la cuenca; el volcán consiste en un conjunto de varias rocas volcánicas formadas por mantos andesítico-dacíticos en su parte inferior, que van cambiando gradualmente de composición hacia el este hasta que se hacen casi dacíticos (White, 1990). A mitad de la montaña se presentan mantos de andesita lamprobolita de piroxeno, la cual continúa con un agregado de las rocas de andesita hasta la cumbre. En el lado este

alrededor del Valle del Doberman hay un gran cuerpo de roca que posiblemente corresponda a rocas volcánicas del Terciario Medio de Schlaepfer (1968), éstas se presentan en los flancos sureste del volcán. Hacia el oeste y noroeste se encuentran rocas antiguas conocidas como la serie Andesítica de las Cruces (White 1990). Las cumbres del Ajusco sobresalen de los mantos de lava y sedimentos volcánicos de la Serie Basáltica Chichinautzin del Pleistoceno superior y del Holoceno (White 1990).

El volcán Xitle es un pequeño cono cinerítico de la Serie Chichinautzin que se encuentra a 2 km del flanco noreste del Ajusco (Fig. 2). Este volcán produjo varios mantos delgados de lava y una serie de erupciones más recientes, fechadas con técnicas de radiocarbón en $2,422 \pm 250$ años y $1,960 \pm 70$ años, respectivamente (White 1990). Durante estas últimas lanzó varias capas de ceniza volcánica que cubrieron las praderas del noreste del Ajusco y los depósitos glaciales que ahí se encontraban. Hoy en día, los flujos de lava del Xitle forman un depósito continuo que va desde su parte sur hasta la ceniza negra de las morrenas de la parte norte y noreste del Ajusco (Fig. 2).

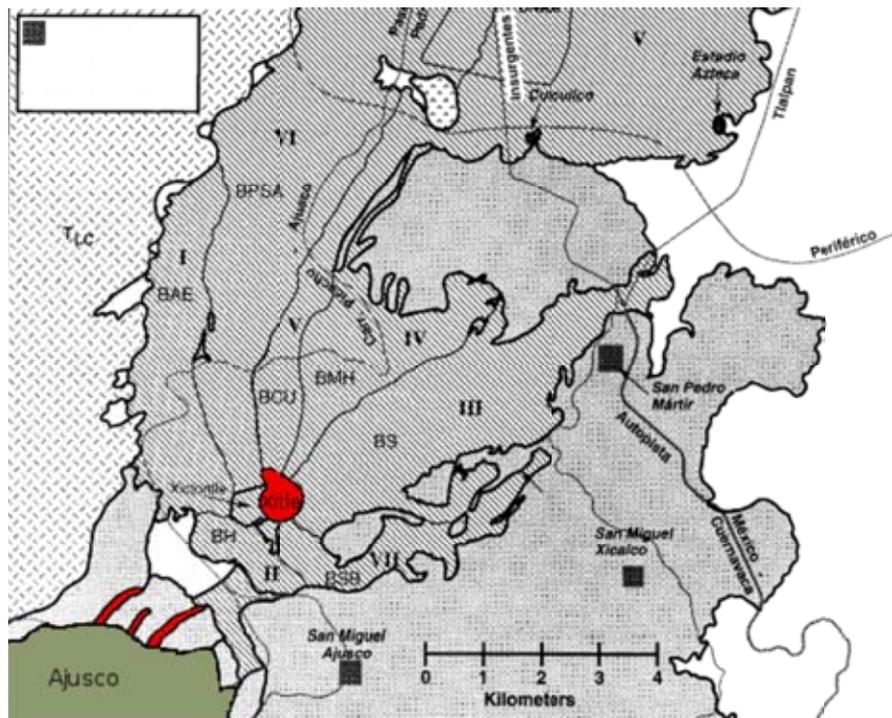


Figura 2 – Flujos de lava del volcán Xitle. En rojo se detalla la ubicación del Xitle y los flujos que impactaron el noreste del volcán Ajusco. Los números romanos indican la secuencia extrusiva de los flujos de lava: I) Agua

escondida; II) Héroes de 1910; III) Miguel Hidalgo; IV) Ciudad Universitaria; V) Pedregal de San Ángel; VI) Seminario y; VII) San Buenaventura Modificada de Delgado *et al.* (1998).

En el Ajusco se han inferido al menos cinco glaciaciones durante los últimos 28,000 años (Tabla 2, Fig. 3). Los bordes o el delineamiento de los glaciares se enmarcan por las morrenas laterales, o por las crestas cubiertas de depósito glacial.

Los primeros glaciares que, con alguna certeza, alcanzaron las laderas inferiores del Ajusco, fueron los de la glaciación Marqués hace unos 27,190 (± 930) años (White, 1990). Después, estos retrocedieron e incluso desaparecieron durante unos 2,000 años, hasta la glaciación Santo Tomás, ocurrida hace 25,077 (± 680) años, cuando volvieron a desarrollarse (White, 1990). Luego de éste período, el glacial retrocedió nuevamente y el área alrededor del Ajusco, incluyendo la Cuenca de México, estuvo sometida a una historia eruptiva volcánica complicada. En ésta se presentaron varios períodos de grandes precipitaciones de cenizas volcánicas que fueron seguidos por la erosión de las mismas en las laderas altas del Ajusco y por su acumulación sobre las morrenas y los sedimentos coluviales. Durante este prolongado tiempo de erupciones volcánicas tuvo lugar la glaciación Albergue (hace 15,088 ± 150), que fue un acontecimiento corto y tardío (White 1990). Sobre los depósitos glaciares producto de las tres glaciaciones mayores (glaciación Marqués, glaciación Santo Tomás y glaciación Albergue) se formaron los suelos inceptisoles, cuya diferencia radica en el grado de desarrollo; siendo los más desarrollados aquellos formados sobre el depósito producto de la glaciación Marqués (White 1990).

Tabla 2. Longitud total y edad de los glaciares en los cuatro valles del Ajusco durante las glaciaciones del Pleistoceno superior. La edad está calculada a partir de técnicas de radiocarbono (White 1990).

Glaciación	Edad (años)	Longitud (km)			
		Valle del Águila	Valle de Santo Tomás	Valle de Pipixaca	Valle del Doberman
Marqués	27,190 \pm 930	4.5	3.5	1.5	2.3
Santo Tomás	25,077 \pm 680	2.1	2.6		1.6
Albergue	15,088 \pm 150	1.2	2.8		
Neoglaciación I	11,420 \pm 120	0.4			

El clima del Ajusco depende de la latitud y de la altitud de los diferentes picos (White 1990).

La temperatura y la precipitación media anuales son de 8.1 °C y 1,341 mm, respectivamente (García-Aguirre *et al.* 2007). En las alturas menores y en las zonas que rodean el volcán por el sur, existe un clima cálido de meseta, con inviernos secos y veranos frescos, según la clasificación de Koeppen (White 1990). Alrededor de las laderas superiores hay un clima lluvioso templado-cálido sin estación seca notable. Esta parte está sometida a la influencia del viento portador de humedad de los alisios del noreste en el verano. Las laderas superiores tienen un clima de tundra (White, 1990). La parte norte es más húmeda que la parte sur debido a una menor radiación solar y a que el mismo volcán actúa como una barrera orográfica contra los vientos prevalecientes de dirección noreste (García-Aguirre *et al.* 2007). Los veranos pueden ser más prolongados y calurosos en la parte sur de la montaña que en la norte, debido al clima tropical presente a corta distancia del sur de la Meseta Central (White 1990).

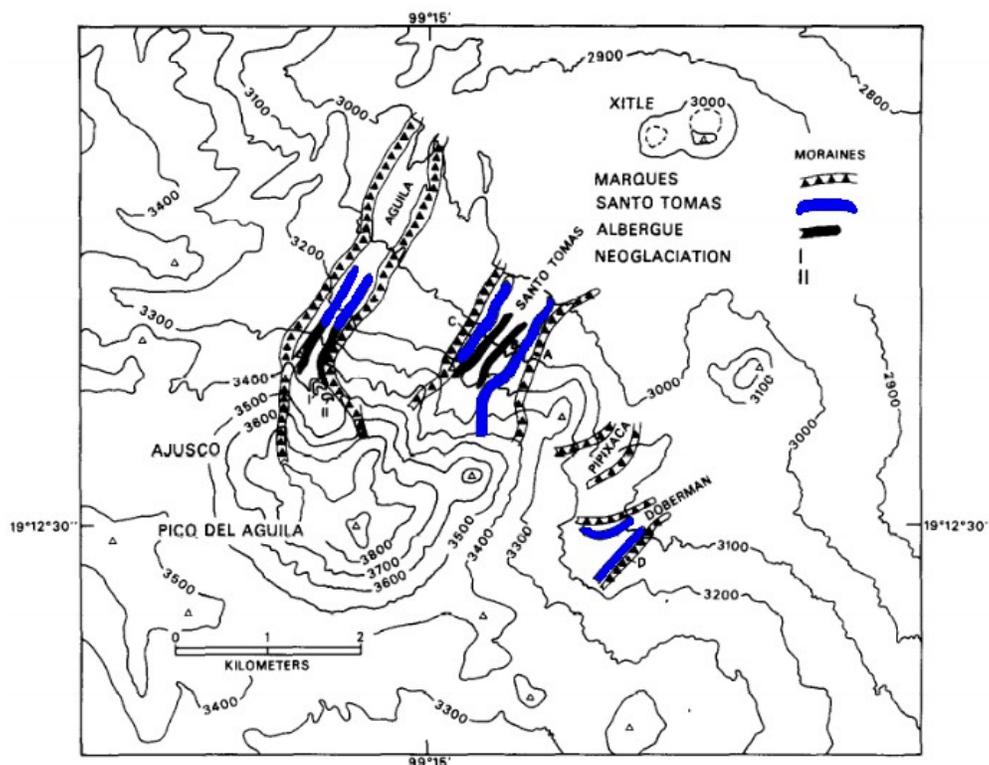


Figura 3 – Mapa que ilustra la ubicación de las morrenas producto de las cinco glaciaciones ocurridas en el volcán Ajusco. Modificada de White & Valastro (1984).

En los lados norte y noreste se encuentran varios tipos de suelo inceptisoles (ocreptos y

andptos), de acuerdo a la clasificación de suelos de la Séptima Aproximación (1960), mientras que las laderas más secas y rocosas del sureste, sur y suroeste están formadas por entiosoles (ortentos) (White, 1990).

Existe una diferencia sustancial entre la pendiente norte y la sur. En la primera se encuentran remanentes glaciales como morrenas (Fig 3), mientras que en la segunda se desarrollaron fenómenos periglaciales como fluctuaciones frecuentes de temperatura y procesos cryoclásticos (White & Valastro 1990), es decir producción de clastos debido a la fragmentación de las rocas por el frío. El movimiento de estos clastos resultó en la formación de abanicos aluviales en la base de la pendiente. Estas diferencias entre las pendientes provocan una asimetría en el volcán e influyen fuertemente en la distribución de la vegetación (García-Aguirre *et al.* 2007).

Tabla 3 – Comunidades vegetales en el volcán del Ajusco (Tomado de García-Aguirre *et al.* 2007).

Comunidad vegetal	Altitud	Localización	Cobertura	Pendiente (°)
(1) <i>Festuca amplissima</i> , <i>Muhlenbergia macroura</i>	3370 – 3910	S	20 – 80	20 – 45
		SO	1 – 60	
(2) <i>Pinus hartwegii</i> , <i>Festuca amplissima</i> , <i>Muhlenbergia macroura</i>	3260 – 3800	S	10 – 60	15 – 45
		SE	1 – 40	
(3) <i>Pinus hartwegii</i> , <i>Festuca amplissima</i> , <i>Muhlenbergia macroura</i> , <i>Calamagrostis tolucensis</i>	3430 – 3570	SO	10 – 30	10 – 40
		SE	2 – 70	
(4) <i>Festuca amplissima</i> , <i>Muhlenbergia macroura</i> , <i>Pinus hartwegii</i>	3200 – 3820	SE	10 – 40	12 – 40
		S	1 – 70	
(5) <i>Abies religiosa</i> , <i>Senecio angulifolius</i> , <i>Senecio barba-johannis</i>	3050 – 3500	NO	10 – 50	15 – 45
		O	2 – 50	
(6) <i>Calamagrostis tolucensis</i> , <i>Muhlenbergia macroura</i> , <i>Pinus hartwegii</i>	3600 – 3800	NO	5 – 10	8 – 35
		SE	1 – 20	
(7) <i>Abies religiosa</i> , <i>Senecio angulifolius</i> , <i>Senecio barba-johannis</i>	3210 – 3560	NE	10 – 70	10 – 35
		NE	5 – 70	
(8) <i>Pinus montezumae</i> , <i>Abies religiosa</i> , <i>Solanum cernitessi</i>	2980 – 3200	NE	20 – 80	5 – 18
		NO	5- 30	

En el Ajusco se encuentran definidas ocho comunidades vegetales según un estudio realizado bajo un enfoque fito-geomorfológico (García-Aguirre *et al.* 2007) (Tabla 3). Cinco de ellas corresponden al grupo de pinos y pastos, dos al de abetos y una es una comunidad de pastos (Fig. 4). Las comunidades de *Festuca*, *Pinus* y *Muhlenbergia* crecen en la pendiente

sur entre los 3,200 y 3,800 msnm. En todas las pendientes, entre los 3,400 y 3,900 msnm se distribuyen comunidades de *Festuca* y *Muhlenbergia*, mientras que los bosques de *Abies* están mejor representados en la pendiente norte (2,900 – 3,200 msnm) y la noreste (3,200 – 3,560 msnm). En las pendientes sur y sureste (3,050 – 3500 msnm) solo se encuentran en las barrancas más profundas. Esta distribución puede ser explicada por las condiciones de humedad, fertilidad del suelo y protección del viento, requerimientos que son cumplidos en la mayoría de las pendientes cóncavas del norte, así como en algunas de estas pendientes en el sur (García-Aguirre *et al.* 2007).

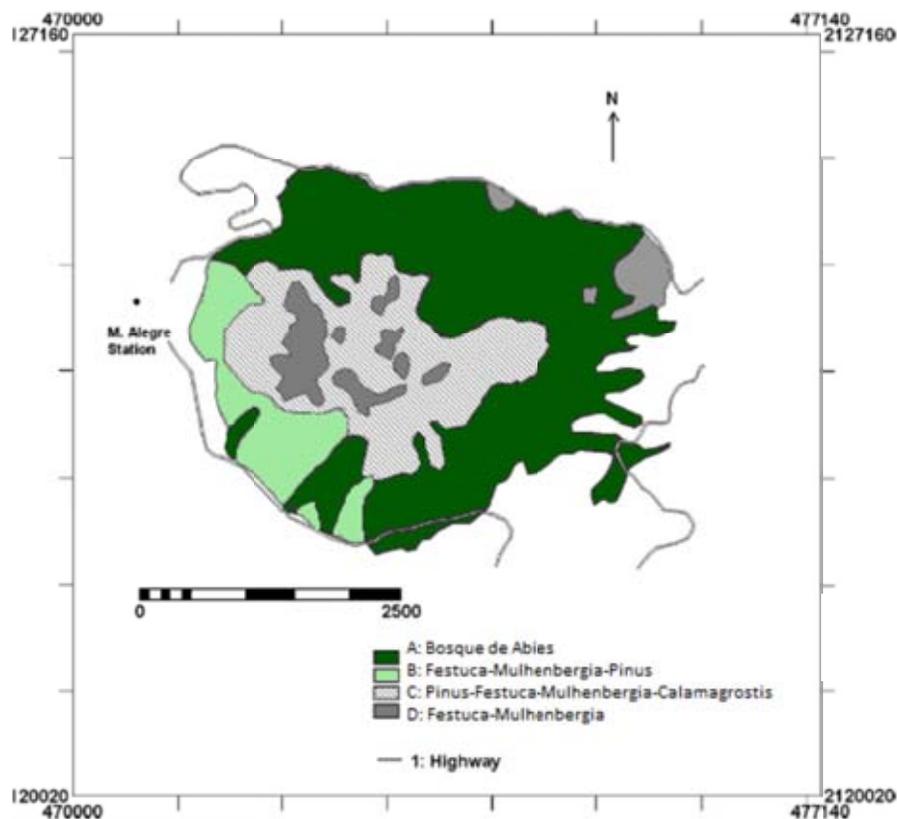


Figura 4 – Mapa de la vegetación del volcán Ajusco. (A) Los bosques de *Abies* se distribuyen en su mayoría sobre la pendiente norte y noreste (comunidades 5 y 7 de la Tabla 3). (B-D) Comunidades de pinos y pastos se encuentran sobre todo en la pendiente sur y a grandes alturas (comunidades 1, 2, 3, 4, 6 y 8). Modificada de García-Aguirre *et al.* (2007).

Si bien las condiciones ambientales donde se distribuye *Abies religiosa* dentro del Ajusco son muy similares, existen factores que podrían sugerir la existencia de micro ambientes; esto explicaría la distribución tan localizada de esta especie y las diferencias en su abundancia entre la pendiente norte y sur. Lo anterior podría sugerir la ocurrencia de procesos de diferenciación genética (adaptativa o neutral) a nivel intrapoblacional; así como procesos de adaptación local.

2. Objetivos generales

Identificar los elementos ecológicos y/o históricos que influyen en la distribución espacial de la variación genética de la población de *Abies religiosa* del volcán Ajusco.

2.1 Objetivos particulares

Determinar la diversidad y estructura genética de la población de *Abies religiosa* del volcán Ajusco.

Comparar la intensidad del flujo genético producto de la dispersión a través del polen y de las semillas, así como su efecto en la estructura genética intrapoblacional.

Identificar loci posiblemente afectados por selección natural y su relación con determinadas variables ambientales, así como su importancia en la diferenciación genética.

3. Justificación

La complicada historia biogeográfica de los *Abies* mexicanos hace difícil discernir entre las distintas fuerzas evolutivas que han moldeado su variación genética. Es posible que las poblaciones de estas especies hayan tenido fluctuaciones en cuanto a su tamaño y distribución, estos eventos demográficos sin lugar a dudas han tenido un fuerte impacto en la cantidad y distribución de la diversidad genética (Aguirre-Planter *et al.* 2000, Jaramillo-Correa *et al.*, 2008). Entender las consecuencias genéticas de dichos procesos demográficos, a distintas escalas geográficas, es importante para conocer bajo qué condiciones los cambios en las distribuciones pueden amenazar la existencia de determinadas poblaciones y/o especies o, por el contrario, tener un rol activo en la generación de la diversidad de especies (Knowles & Alvarado – Serrano, 2010). La poca divergencia interespecífica, combinada con la gran diferenciación intraespecífica de los *Abies* mexicanos plantea la necesidad de estudiar más a fondo las determinantes de la distribución espacial de la variación genética. La deriva genética y el flujo genético han sido señaladas como las principales fuerzas evolutivas que actúan sobre estas especies (Aguirre-Planter *et al.* 2000, Jaramillo-Correa *et al.*, 2008, Mugica-Gallart, 2013). Es necesario analizar el papel de dichas fuerzas a escala microespacial e intrapoblacional, ya que estas tienen una gran influencia en poblaciones pequeñas y localizadas, como las de las especies endémicas (Hamrick & Loveless, 1984); además, la importancia relativa de cada proceso depende (en gran medida) de la capacidad de dispersión de la especie focal, que a su vez esta correlacionada con la escala espacial (Orsini *et al.*, 2014).

La información aquí obtenida complementa el conocimiento previo acerca de los procesos evolutivos macroespaciales (Aguirre-Planter *et al.*, 2000; Jaramillo-Correa *et al.*, 2008; Mugica-Gallart, 2013), además de que permite estudiar un espectro más amplio del flujo genético al utilizar marcadores moleculares con diferentes patrones de herencia. En este estudio se analizó explícitamente, y por primera vez para los *Abies* mexicanos, el papel de la selección natural en la generación de la estructura genética intrapoblacional; esto no solo brinda un panorama más amplio sobre la historia evolutiva de este género en México, sino que permite identificar los factores ambientales que han provocado el surgimiento de

adaptaciones locales. El conocimiento de las presiones que afectan la dinámica evolutiva de las poblaciones a un nivel microespacial, así como la información obtenida a partir de las distintas aproximaciones empleadas en este estudio, permiten generar estrategias de conservación y manejo de poblaciones fragmentadas debido a presiones como la urbanización, la agricultura o los incendios forestales, todas situaciones particularmente relevantes para el Eje Neovolcánico Transversal donde se concentra la mayor densidad poblacional (humana) en el país (INEGI).

4. Hipótesis

La distribución espacial de la variación genética depende de factores paisajísticos, como la distancia geográfica, o de similitudes ecológicas entre localidades.

Dado que la capacidad de dispersión del polen es mayor que la de las semillas, no se encontrará estructura genética cuando se estudien los marcadores moleculares de cloroplasto.

Las diferencias ambientales presentes en la población estudiada, generan adaptaciones locales intrapoblacionales y estas se reflejan en unos pocos loci afectados por selección natural.

5. Método

Análisis moleculares

5.1 Obtención, extracción y amplificación de muestras

Para el muestreo, realizado durante diciembre del 2012, se localizaron 12 puntos sobre la carretera Picacho – Ajusco que rodea el volcán (Fig. 5); allí se colectaron acículas de 20 individuos separados por 10 metros o más dependiendo de la abundancia de la especie en cada localidad. Cada árbol colectado fue georeferenciado utilizando un GPS (Anexo 1a).

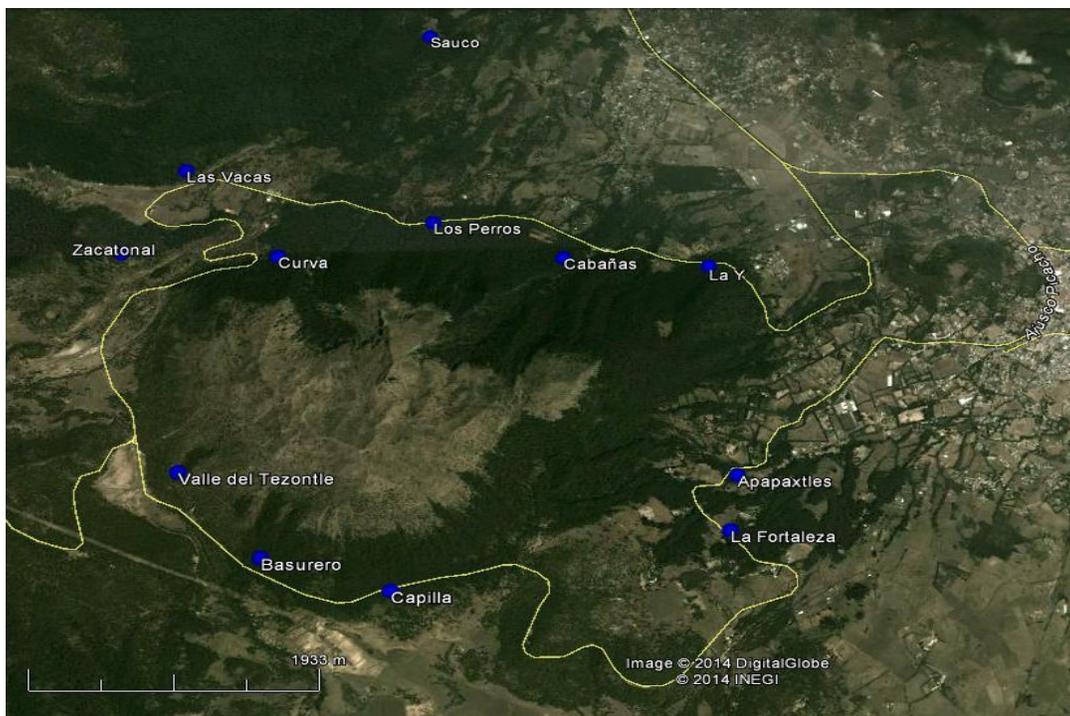


Figura 5. Ubicación geográfica y nombre de las doce localidades ubicadas sobre de la carretera Picacho – Ajusco que rodea el volcán Ajusco donde se realizó el muestreo. En cada localidad se tomaron acículas de 20 individuos y la geo referencia individual.

El ADN se extrajo utilizando el kit “DNeasy plant mini kit” (QIAGEN) siguiendo las instrucciones del fabricante. Para comprobar que la extracción fue exitosa se corrieron 5 μ L del ADN en un gel de agarosa al 1% en TBE y teñido con bromuro de etidio, con una corriente de 100 V durante 30 minutos, para después observarlo bajo luz ultravioleta.

El protocolo utilizado para la obtención de polimorfismos de tamaño de fragmentos

amplificados (AFLPs) constó de cuatro pasos:

1. El ADN genómico (1 μ L) fue digerido utilizando cinco unidades de la enzima de restricción *Eco RI* (Invitrogen) y tres de *Tru 11 (Mse I)* (Thermo Scientific), en el buffer Tango 2X (Thermo Scientific) a un volumen final de 15 μ L. La mezcla se incubó por tres horas a 37° C y luego por otras tres horas a 65° C.
2. Se ligaron los adaptadores EcoRI-adapter (F_CTCGTAGACTGCGTACC y R_AATTGGTACGCAGTCTAC) y MseI-adapter (F_GACGATGAGTCCTGAG y R_TACTCAGGACTCAT) con concentraciones de 3.8 mM y 0.38 mM, respectivamente. Esto se hizo en un volumen total de 6.47 μ L que contenía *bufferT4* (1.53X), NaCl₂ (0.003 mM), BSA (0.001 mg/mL) y una unidad de Ligasa (Invitrogen). Posteriormente se añadió 6.47 μ L de ADN-digerido. La reacción se incubó por 3 horas a 37° C y el producto final se diluyó 1:10 con agua bidestilada.
3. Se realizó un PCR pre-selectivo con la siguiente mezcla: 1X de Green go Taq Flexi Buffer, 1.5 mM de MgCl₂, 0.5 mM de dNTPs, 0.2 mM de cebador Eco+0, 0.2 mM de cebador Mse+0, una unidad de Go Taq Flexi DNA Polymerase (PROMEGA), y 3 μ L del producto diluido de la ligación, para un volumen total de 50 μ L.
El programa de amplificación fue el siguiente: 2 minutos de desnaturalización a 94° C, seguido de 25 ciclos de 20 segundos de desnaturalización a 94° C, 1.5 minutos de alineación a 57° C y 3 minutos de extensión a 72° C.
4. Se realizó un PCR selectivo con la siguiente mezcla: 1X de *Green go Taq Flexi Buffer*, 1.5 mM de MgCl₂, 0.5 mM de dNTPs, 0.5 mM de cebador *Eco+3*, 0.5 mM de cebador *Mse+4*, una unidad de *Go Taq Flexi DNA Polymerase* (PROMEGA), y 2.5 μ L del producto del PCR preselectivo, para un volumen total de 12.5 μ L. (Véase Anexo 2a para información sobre la secuencia y combinaciones de cebadores usados).
El programa de amplificación para este último paso fue el siguiente: 2 minutos de desnaturalización a 94° C, seguido de 25 ciclos de 20 segundos de desnaturalización a 94° C, 1.5 minutos de alineación a 57° C y 3 minutos de extensión a 72° C, por último un ciclo de extensión de 40 minutos a 60° C.

La intención de realizar una segunda ronda de amplificación (PCR selectivo), utilizando cebadores con cierto número extra de oligonucleótidos, es tener un submuestreo del total de fragmentos de restricción obtenidos en el primer PCR, de esta forma tanto la especificidad como la reproducibilidad del método aumentan.

Para comprobar el éxito de las amplificaciones, preselectiva y selectiva, estas se analizaron mediante una electroforesis, utilizando 3µL del producto final en un gel de agarosa al 1.5% en buffer de corrida TBE, teñido con bromuro de etidio, utilizando una carga de 100 V durante 30 minutos, para después observarlo bajo luz ultravioleta.

En las reacciones de PCR selectivo el cebador con la secuencia *Eco* se encontraba marcado con un fluoróforo; utilizando los servicios de secuenciación de la Universidad de Illinois en Urbana-Champaign se obtuvo un electroforegrama de cada reacción. En este electroforegrama (o perfil genético) cada fragmento amplificado es reportado con su tamaño (en pares de bases, pb) e intensidad (medida en unidades de fluorescencia relativa, rfu).

5.2 Selección de fragmentos y reproducibilidad de AFLPs.

Debido a la naturaleza anónima e inespecífica de los AFLPs fue necesario realizar una serie de pruebas para verificar la confiabilidad y seleccionar únicamente los fragmentos reproducibles. Para ello se eligieron ocho muestras al azar con las que se realizó el protocolo completo por duplicado (réplicas).

Los perfiles genéticos de todas las muestras, incluyendo las réplicas, fueron analizados utilizando dos programas. Primero se empleó el programa Peak Scanner (Applied Biosystems) para detectar los fragmentos amplificados contenidos en el electroforegrama así como su tamaño e intensidad. Posteriormente se utilizó el paquete para R v3.0.2 (R core team, 2009) RawGeno v2.02 (Arrigo *et al.*, 2009) para agrupar y marcar los electroforegramas, con base en esto se seleccionaron los fragmentos amplificados que cumplieron con una serie de condiciones. Cada combinación de cebadores se analizó por separado, posteriormente se unieron los resultados. La primera fue una intensidad mínima de 100 rfu, a partir de esto se eliminó el 5% de los individuos con menor número de fragmentos amplificados (es decir, aquellos para los cuales la reacción de PCR no fue exitosa). La segunda condición evaluó la reproducibilidad del fragmento; para lo cual se eliminaron los fragmentos cuyo estado (presencia/ausencia) no coincidiera en al menos siete de los ocho individuos replicados (87.5 % de reproducibilidad mínima). Por último, se eliminaron aquellos

fragmentos cuya presencia se encontrara en menos del 5% del total de la población (para evitar sesgos por homoplasia). Para más detalle del procedimiento realizado véase el Anexo 2b.

Con los marcadores retenidos se generó una matriz binaria (presencia/ausencia) para cada combinación de cebadores. Posteriormente los individuos eliminados en el primer paso (aquellos que para una, o varias combinaciones de cebadores en específico tuvieron un número de fragmentos amplificados menor al 5%) se identificaron y concatenaron; aquellos con reacciones no exitosas de PCR mayores del 20% del total de fragmentos amplificados (utilizando las ocho combinaciones de cebadores) fueron eliminados de la base de datos. Una vez excluidos estos individuos se repitió el análisis en RawGeno, pero sin eliminar el 5% de los individuos con menor número de fragmentos amplificados. Por último se combinó la información de cada combinación de cebadores para obtener una matriz binaria final.

5.3 Amplificación de microsatélites de cloroplasto

Se genotiparon tres microsatélites de cloroplasto (*Pt30204*, *Pt71936* y *Pt63718*) descritos como polimórficos para *Abies religiosa* (Jaramillo-Correa *et al.*, 2008), utilizando los oligonucleótidos descritos en Vendramin *et al.* (1996). Las condiciones de PCR para los tres loci fueron las siguientes: 2.5X de buffer para PCR, 1.25 µL de MgCL₂, 0.2 µL de dNTPs (10 mM), 0.5 µL de cada oligonucleótido, una unidad de Taq polimerasa (Invitrogen) y 1.5 µL de ADN. El programa de amplificación fue el siguiente:

2 minutos de desnaturalización a 95° C, seguido de 35 ciclos de 1 minuto de desnaturalización a 95° C, 1 minuto de alineación a 55° C y 1 minuto de extensión a 72° C, y por último un ciclo de extensión de 7 minutos a 72° C. Para comprobar que la amplificación fue exitosa se corrieron 3 uL del producto en un gel de agarosa al 1.5% en TBE, teñido con bromuro de etidio, con una carga de 100 V durante 30 minutos, para después observarlo bajo luz ultravioleta. Las muestras se analizaron utilizando un secuenciador automático ABI Prism 3730xl (Applied Biosystems) del servicio de análisis de fragmentos en el Roy. J. Carver Biotechnology Center de la Universidad de Illinois en Urbana-Champaign. Posteriormente, utilizando el programa Peak Scanner (Applied Biosystems) se determinó el tamaño de fragmentos en cada locus para cada individuo.

Análisis estadísticos

5.4 Diversidad genética.

AFLPs

Utilizando el software Arlequin v3.5.1.2 (Excoffier *et al.*, 2005) se calculó el número de loci útiles por localidad, es decir aquellos que no presentaron más del 5% de datos faltantes. A partir del número de estos se calculó el porcentaje de loci útiles respecto al máximo posible, es decir al número total de loci incluidos en la matriz binaria final, y el porcentaje de loci polimórficos (*%Pol*) por localidad. Además, se calculó el valor del marcador ponderado por su frecuencia (índice $D - W$) para cada localidad; dividiendo la frecuencia de cada loci por localidad entre la frecuencia total en la población, posteriormente se sumaron estos valores. Para controlar las diferencias en el tamaño de muestra, este índice se calculó utilizando cinco individuos por localidad tomados al azar. En poblaciones aisladas por mucho tiempo se espera que este índice tenga un valor alto, debido a la acumulación de fragmentos raros producto de mutaciones; mientras que en poblaciones establecidas recientemente se esperan valores bajos, lo que permitiría diferenciar entre vicarianza histórica y dispersión reciente (Schönswetter & Tribsch 2005). Para cada medida se calculó la media y la desviación estándar.

Para analizar las relaciones entre el tamaño de muestra, la cantidad de información disponible y los valores de diversidad (*%Pol* e índice $D - W$) se calculó de manera independiente el coeficiente de Pearson entre pares de variables utilizando R v3.0.2 (R core team, 2009). Al probar simultáneamente múltiples hipótesis, la probabilidad de error tipo I (es decir, la probabilidad de rechazar al menos una de las hipótesis probadas, siendo ésta verdadera) aumenta de manera proporcional al número de hipótesis probadas (Joost *et al.*, 2007). En este caso, las combinaciones entre pares de variables calculadas resultan en un total de seis hipótesis, por lo que la probabilidad de error tipo I a un nivel de confianza del 0.05 es solo del 26.5%, debido a ello no se utilizó ningún tipo de corrección para comparaciones múltiples.

Microsatélites de cloroplasto

Los haplotipos de cloroplasto (clorotipos) fueron definidos ensamblando los genotipos

obtenidos para cada locus. Utilizando los clorotipos se calculó la diversidad genética con el software Arlequin v3.5.1.2 (Excoffier *et al.*, 2005). Además se obtuvo el número de clorotipos y su frecuencia dentro de cada localidad y para toda la población.

5.5 Estructura genética

AFLPs

Para cuantificar la variación genética contenida dentro y entre las localidades se realizó un análisis de varianza molecular (AMOVA) utilizando el índice de fijación F_{ST} . Para medir la distancia genética entre localidades se calculó este índice pero de manera pareada entre todas las localidades. Ambos análisis se realizaron en el software Arlequin v3.5.1.2 (Excoffier *et al.*, 2005). Para todos los casos, el nivel de significancia se evaluó a partir de la distribución nula de los valores pareados de estos estadísticos en 1,000 permutaciones bajo la hipótesis de cero diferenciación entre localidades. De esta manera se calcula el valor p como la proporción de las permutaciones que llevan a un valor del estadístico mayor o igual al observado. Al calcular el índice de fijación F_{ST} de manera pareada se presenta el problema de múltiples hipótesis simultáneas. Para lidiar con esto, los valores p de este índice calculados con microsatélites fueron corregidos utilizando el valor q (basado en la tasa de falso descubrimiento) calculado en el paquete $qvalue$ (Dabney *et al.*, 2013) para R v3.0.2 (R Core Team, 2013). En el caso de los AFLPs se utilizó la corrección de Bonferroni (Shaffer, 1995), ya que muchas de las comparaciones hechas (es decir, F_{ST} pareada) tuvieron valores p igual a cero, lo que imposibilitó el uso de la corrección basada en el valor q .

Para analizar la estructura genética poblacional tomando en cuenta la localización espacial de los individuos se utilizaron dos métodos Bayesianos de agrupamiento. Primero se utilizó Geneland v4.0.3 (Guillot *et al.*, 2005) para R v3.0.2 (R Core Team, 2013), en el que se efectuaron cinco corridas independientes de 1,000,000 de iteraciones cada una y efectuando un muestreo cada 1,000 iteraciones; el rango de grupos genéticos posibles (k) fue de 1 a 12 (ya que éste fue el número de localidades muestreadas). Para elegir la corrida con mayor probabilidad se utilizó el valor de la densidad posterior promedio y luego se descartaron las primeras 200 iteraciones guardadas para dicha corrida. Finalmente se obtuvieron el número de grupos genéticos, la distribución espacial de cada grupo y la probabilidad de pertenencia a cada grupo sobre el espacio muestreado. El segundo método fue implementado en el

programa BAPS v6.0 (Corander *et al.*, 2008), en éste se agruparon espacialmente los individuos muestreados (para hacer comparables los resultados con los de Geneland v4.0.3; Guillot *et al.*, 2005) utilizando un modelo sin mezcla. Para ello se realizaron 10 corridas independientes probando valores de k de 1 a 12. Con la finalidad de evaluar la intensidad y direccionalidad del flujo genético, entre los grupos identificados, posteriormente se realizó un análisis bajo un modelo de mezcla. En este análisis se esquematizó el flujo genético utilizando flechas ponderadas; el “peso” de cada flecha es igual a la cantidad de ancestría relativa entre el grupo genético que dona los migrantes y aquel que los recibe (Corander *et al.*, 2008). De esta forma se obtuvo el número de grupos genéticos, la distribución espacial de cada grupo, el nivel de mezcla por individuo y estimados de la dirección e intensidad del flujo genético entre los grupos identificados.

A partir de la estructura genética obtenida con Geneland v4.0.3 (Guillot *et al.*, 2005) se realizó otro AMOVA utilizando el programa Arlequin v3.5.1.2 (Excoffier *et al.*, 2005), para cuantificar la variación genética contenida entre grupos genéticos, entre localidades dentro del mismo grupo genético y dentro de las localidades.

Microsatélites de cloroplasto

Para medir la distancia genética entre localidades se calculó el estadístico F_{ST} de manera pareada entre todas las localidades con Arlequin v3.5.1.2 (Excoffier *et al.*, 2005). El método utilizado se basó en la suma de las diferencias cuadradas de el tamaño de los alelos, por lo que resulta análogo a R_{ST} de Slatkin (Slatkin, 1995), de esta manera se controló el efecto de la mutación, ya que este estadístico está basado explícitamente en el modelo de mutación paso a paso. El nivel de significancia se evaluó de la misma manera que para los AFLPs.

Para determinar el número de linajes de cloroplasto presentes en la población, así como su ubicación espacial se utilizó el programa BAPS v6.0 (Corander *et al.*, 2008). En éste se realizó un análisis de agrupamiento espacial basado en individuos bajo un modelo sin mezcla. Se probaron valores de k de 1 a 12 en 10 corridas independientes para cada valor y posteriormente se utilizó la corrida con la mayor probabilidad posterior para obtener el número de grupos genéticos y su distribución espacial.

5.6 Modelo causal

Para explorar las determinantes de la estructura genética se pusieron a prueba dos hipótesis bajo el marco de un modelo causal. En este marco conceptual se busca identificar la hipótesis que explicaba mejor la diferenciación genética utilizando pruebas de Mantel y de Mantel parcial (Wang & Summers, 2009). La hipótesis seleccionada fue aquella que mostró una correlación significativa con la diferenciación genética, aún cuando el efecto de la o las hipótesis alternativas haya sido removido mediante las pruebas de Mantel parcial (Wang & Summers, 2009).

Se utilizaron dos matrices de diferenciación genética, una construida con AFLPs (**G**) utilizando el estadístico F_{ST} y la otra con microsatélites de cloroplasto (**g**) utilizando el estadístico F_{ST} análogo a R_{ST} . Ambas matrices fueron las obtenidas previamente con el software Arlequin v3.5.1.2 (Excoffier *et al.*, 2005). Las dos hipótesis probadas fueron la distancia geográfica (IBD) y la distancia ecológica (IBA). La matriz de distancia geográfica (**D**) fue construida utilizando el programa Geographic Distance Matrix Generator v1.2 (http://biodiversityinformatics.amnh.org/open_source/gdmg/). Para construir la matriz de distancia ambiental (**A**) se realizó el siguiente procedimiento: i) utilizando las capas de información geográfica disponibles en el Instituto Nacional de Estadística y Geografía (<http://www.inegi.org.mx/>) se seleccionaron las variables ambientales que presentaban al menos dos valores en la zona de estudio (esto debido a la resolución de los datos disponibles comparado con la escala espacial del estudio), las cuales fueron: “Tipo de roca”, “Tipo de suelo”, “Región hidrológica” y “Tipo de vegetación” (Anexo 3a); ii) a partir de estas variables se construyó una tabla en la que se especifica el valor de cada variable ambiental en cada localidad; iii) para reducir la variación en los datos se realizó un análisis de correspondencias múltiples (MCA) con el paquete FactoMineR v1.26 (Husson *et al.*, 2013) para R v3.0.2 (R Core Team, 2013). Este análisis es análogo al análisis de componentes principales, para el caso de variables cualitativas; iv) a partir de las dimensiones obtenidas con el MCA se calcularon las distancias euclidianas entre localidades utilizando el paquete vegan v2.0 (Oksanen *et al.*, 2013) para R v3.0.2 (R Core Team, 2013); la matriz **A** se construyó a partir de estas distancias.

Una vez construidas las matrices, se realizaron las pruebas de Mantel para todas las

combinaciones de variables, posteriormente se realizaron las pruebas de Mantel parcial para controlar el efecto de una variable sobre otra, ambas pruebas se realizaron con el paquete *vegan* v2.0 (Oksanen *et al.*, 2013) para R v3.0.2 (R Core Team, 2013). De esta forma se evaluaron los factores que influyen en la diferenciación poblacional y su relación con las dos formas de dispersión: la de las semillas+polen (AFLPs) y la del polen (microsatélites de cloroplasto); a la vez que se puso a prueba la relación entre la diferenciación genética dada por ambos tipos de dispersión.

5.7 Detección de loci bajo selección natural

Para detectar loci con señales de selección natural se utilizaron tres aproximaciones:

- Bayescan v2.1 (Foll & Gaggiotti, 2008): este análisis se realizó utilizando dos modelos distintos cuya diferencia radica en el coeficiente de endogamia utilizado (F_{IS}). El algoritmo utilizado en este programa requiere información *a priori* acerca de la probabilidad de la hipótesis neutral, una probabilidad de 1 significa que, para cada loci se asume que tanto la hipótesis selectiva como la neutral tienen la misma probabilidad de ser correctas. Una probabilidad *a priori* alta disminuye el número de falsos positivos, pero también reduce el poder para detectar loci bajo selección (Foll & Gaggiotti, 2008). El valor utilizado fue 10, que es el recomendado en el manual del programa para bases de datos con unos pocos cientos de marcadores. Para estimar las frecuencias alélicas es posible equilibrio de Hardy-Weinberg o bien considerar ciertas desviaciones de éste, esto se hace a través del coeficiente de endogamia F_{IS} . Cuando no se tiene información previa sobre este coeficiente, se utiliza un intervalo entre 0.0 y 1.0. En el primer modelo se utilizó este intervalo ya que permite explorar distintos valores para F_{IS} . Para el segundo modelo el intervalo utilizado fue de 0.192 a 0.216, estos valores fueron calculados empíricamente para *Abies religiosa*, utilizando microsatélites nucleares (Múgica-Gallart, 2013) e isoenzimas (Aguirre-Planter *et al.*, 2000), respectivamente. En ambos análisis se realizaron 5000 iteraciones con un intervalo de corte cada 10 iteraciones, 20 corridas piloto de una duración de 5,000 iteraciones y un período de calentamiento de 50,000 iteraciones. solo fueron escogidos loci con señales de selección divergente, es decir aquellos con valores de F_{ST} inusualmente altos.
- Modelo jerárquico de islas: este análisis, implementado en el programa Arlequin

v3.5.1.2 (Excoffier *et al.*, 2005), toma en cuenta la estructura poblacional subyacente. Tres modelos fueron puestos a prueba. El primero de ellos (MJI-G) se basó en la estructura genética poblacional obtenida con el análisis realizado en Geneland v4.0.3 (Guillot *et al.*, 2005). Mientras que para el segundo (MJI-S) y tercero (MJI-V) se utilizó una estructura genética poblacional definida con base en las variables ambientales “Tipo de suelo” y “Vegetación”, respectivamente. Las variables ambientales “Tipo de roca” y “Cuenca hidrológica” no fueron utilizadas, ya que al agrupar con base en ellas se rompía uno de los supuestos del método; el flujo genético entre localidades pertenecientes a distintos grupos era mayor que entre localidades dentro del mismo grupo. solo fueron escogidos aquellos loci con señales de selección divergente.

- Método de análisis espacial (SAM) implementado en el programa matsam v2.0 (Joost *et al.*, 2007). Este método se basa en regresiones logísticas y busca probar la asociación entre las frecuencias alélicas y las variables ambientales (Joost *et al.*, 2007). Este tipo de regresión resulta ideal para marcadores con información binomial, como los AFLPs. Las variables ambientales utilizadas fueron las primeras dos dimensiones obtenidas con el MCA. Los individuos pertenecientes a la misma localidad fueron asignados con el mismo valor para cada dimensión ambiental, ya que el MCA fue realizado a nivel de localidad debido a la resolución espacial de los datos ambientales. El nivel de significancia de los coeficientes calculados por la regresión logística se evaluó mediante el estadístico G y/o la prueba de Wald (Hosmer & Lemeshow, 2000); estas pruebas buscan determinar si el modelo que incluye la variable ambiental es más informativo sobre la variable de respuesta que un modelo que no incluya dicha variable (Joost *et al.*, 2007). Este análisis permite utilizar individuos con falta de información en algunos loci, sin embargo, en estos casos el estadístico G no puede ser calculado. Debido a que esta situación esta presente en la base de datos utilizada, el nivel de significancia para algunos loci solo fue evaluado utilizando la prueba de Wald. Para solucionar este problema y darle robustez a los resultados el análisis se realizó en dos etapas. Primero se utilizaron todos los loci, sin importar si tenían información faltante o no, y se seleccionaron aquellos con asociaciones significativas para ambos estadísticos (o solo para la prueba de Wald, en el caso de los loci con falta de información). Posteriormente se retiraron los individuos con falta de información en aquellos loci detectados solo con la prueba de Wald y el análisis fue repetido. En este segundo análisis no se seleccionaron nuevos loci, solo

sirvió para calcular el estadístico G y confirmar el posible carácter selectivo de los loci detectados en el primer análisis y que presentaron falta de información. Realizar este análisis en dos etapas permite detectar el patrón global de selección (ya que utiliza todos los individuos sin importar si presentan falta de información) y, por otra parte, resulta sumamente conservador, ya que evalúa la significancia de cada correlación mediante dos estadísticos. El número de loci y de variables ambientales hacen necesario correr un gran número de modelos univariados para poder detectar señales de selección natural, lo que aumenta la probabilidad de error tipo I (Joost *et al.*, 2007). Para lidiar con este problema se utilizó la corrección de Bonferroni (Shaffer, 1995). Aquellos loci cuyas asociaciones hayan permanecido estadísticamente significativas para ambas pruebas, después de esta corrección, fueron considerados putativamente selectivos.

Cada uno de los métodos utilizados tiene distintos supuestos (véase la Introducción); así mismo, bajo diferentes escenarios ecológicos o demográficos, el poder para detectar señales reales de selección natural puede variar entre ellos (Villemeuril *et al.*, 2014). Debido a esto solo se mantuvieron algunos de los loci detectados por todos los métodos. Los criterios utilizados para conservar estos loci fueron los siguientes: todos los loci detectados con SAM fueron conservados, esto debido al uso de dos pruebas para medir el nivel de significancia que junto con el carácter conservador de la corrección de Bonferroni (García 2004; Narum 2006) hacen de SAM una aproximación suficientemente conservadora (Joost *et al.*, 2007); en el caso de Bayescan solo se mantuvieron aquellos loci detectados por este método y al menos uno de los tres modelos jerárquicos de islas; aquellos loci detectados solo por los modelos jerárquicos de islas fueron conservados siempre y cuando aparecieran en los tres modelos utilizados.

Para evaluar la relación entre los loci putativamente selectivos y la diferenciación genética, se calculó el índice de diferenciación genética F_{ST} de manera pareada entre localidades utilizando Arlequin v3.5.1.2 (Excoffier *et al.*, 2005) con base en las frecuencias alélicas de los loci que cumplieron con los criterios antes descritos. Posteriormente se realizaron pruebas de Mantel entre esta matriz de F_{ST} pareada y las matrices de distancia ambiental y geográfica, de manera independiente, utilizando el paquete vegan v2.0 (Oksanen *et al.*, 2013) para R v3.0.2 (R Core Team, 2013). Las matrices de distancia ambiental y geográfica fueron las

mismas que las utilizadas en las pruebas de Mantel realizadas en el modelo causal.

5.8 Procesos selectivos y diferenciación genética

Para examinar el papel de la selección natural en la diferenciación genética y contrastarlo con el papel de procesos neutrales, se construyó un modelo basado en el algoritmo *Random Forest* (Breiman, 2001). Este algoritmo es una extensión de los árboles de clasificación y regresión; la idea detrás de estos métodos es dividir de manera recurrente la base de datos en estratos cada vez más pequeños, con el fin de obtener el mejor ajuste posible. Para ello el espacio de muestra se divide en dos regiones, dentro de las cuales se ajusta un modelo, posteriormente se repite este proceso varias veces. El mayor componente de estos métodos son las reglas de selección y de finalización. La primera determina cómo será la división en cada estrato, mientras que la segunda determina el estrato final a ser formado, es decir, el término de las divisiones. Los árboles de clasificación y de regresión han demostrado una gran precisión y capacidad para caracterizar interacciones complejas entre variables, superando a técnicas paramétricas más tradicionales, por lo que su uso en modelación ecológica ha aumentado (Cutler *et al.*, 2007). A diferencia de los árboles de regresión y clasificación, la aproximación empleada por *Random Forest* no brinda una estructura clara para el modelo de asociación final, es decir, no genera un árbol final que puede ser interpretado como un modelo de asociación; en vez de esto, estima una medida de la importancia de cada variable predictiva con respecto a la variabilidad observada en la variable de respuesta (Foulkes, 2009).

Antes de explicar el algoritmo empleado en *Random Forest* es necesario definir la notación utilizada, así como algunos términos. Se define $X = (x_1, \dots, x_p)$ como el set de variables predictivas potenciales p , y “ y ” como el rasgo bajo investigación, donde n es el número de individuos en la muestra. La impureza del árbol se define como la suma de la medida de impureza en cada nodo terminal del árbol (Foulkes, 2009). Comenzando con $b = 1$, el algoritmo empleado (Foulkes, 2009) funciona de la siguiente manera:

- 1 – Realiza un muestreo aleatorio con remplazo n_1 de aproximadamente $2/3$ de n , denomina éste como la *muestra de entrenamiento* (LS). Luego deja que los datos restantes $n_2 = n - n_1$ conformen los datos *fuera de la bolsa* (OOB).
- 2 – Utilizando solo los datos LS, genera un árbol “no podado” muestreando al azar un

subconjunto de variables predictivas p para utilizarlas en cada nodo como variables potenciales para la división.

3 – Utilizando solo los datos OOB:

a) Registra la impureza del árbol, denótalo como π_b .

b) Permuta el valor de cada variable predictiva x_j , registra la impureza del árbol para cada valor permutado con cada variable predictiva $j = 1, \dots, p$. Denota este valor como π_{bj} y define la importancia de cada variable predictiva como $\delta_{bj} = \pi_{bj} - \pi_b$.

4. Repite los pasos (1) – (3) para el total de árboles a construir $b = 2, \dots, B$, calcula la importancia de cada variable predictiva en cada árbol construido $\delta_{1j}, \dots, \delta_{Bj}$.

5. Calcula la importancia total de cada variable predictiva (θ_j) definida como la suma de la importancia de cada variable predictiva en cada árbol construido ($\delta_{1j}, \dots, \delta_{Bj}$) entre el número total de árboles (B).

A partir del algoritmo explicado anteriormente, es necesario señalar dos diferencias con respecto a la aproximación tradicional de los árboles de regresión y clasificación. Primero, los árboles construidos son “no podados”, la idea de podar el árbol es implementar un paso extra en el algoritmo que busque cuáles nodos/ramas pueden ser removidos sin afectar en exceso el desempeño del árbol. La aproximación utilizada en *Random Forest* no requiere de este paso, ya que la importancia relativa de cada variable predictiva es evaluada de manera independiente en el paso tres del algoritmo. Segundo, para la división de cada nodo *Random Forest* solo utiliza un subconjunto de las variables como predictores potenciales; es decir, en lugar de determinar la mejor división entre todos los predictores potenciales, se considera una muestra aleatoria de estas variables (típicamente cerca de un tercio de las variables) como variables divisoras potenciales. Una ventaja de utilizar un subconjunto aleatorio de variables predictoras potenciales es que brinda una aproximación natural para manejar colinearidades en los datos. Esto permite que, en teoría, los valores de importancia de cada variable no estén influenciados por la colinearidad entre variables, como podría ser el caso de dos loci en desequilibrio de ligamiento (Foulkes, 2009).

La relevancia de la aproximación empleada por *Random Forest* es que permite evaluar la importancia de cada variable predictiva tomando en cuenta posibles colinearidades o interacciones complejas entre las variables predictivas. El concepto de *importancia de cada variable* se basa en la idea de que permutar el valor de una variable sin importancia no

provoca cambios en la impureza total del árbol, mientras que permutar los valores de una variable muy influyente lleva a un incremento en la medida de la impureza total (Foulkes, 2009). De esta manera es posible identificar variables importantes altamente relacionadas con la variable de respuesta, esto con fines interpretativos; y/o detectar el número mínimo de variables necesarias para una buena predicción sobre la variable de respuesta (Genuer *et al.*, 2010).

Combinando los genotipos obtenidos a partir de AFLPs y microsatélites de cloroplasto se construyó una matriz de distancia entre localidades (Cavalli-Sforza & Edwards, 1967) seguido de un análisis de coordenadas principales (PCoA); éste método permite visualizar y explorar las disimilitudes dentro de una base de datos, por lo que en este contexto se interpretó como una medida de la diferenciación genética entre localidades. Asimismo, se calcularon las frecuencias alélicas de todos los loci por localidad, ambos análisis se realizaron en el programa FAMD v1.31 (Schlüter & Harris, 2006). Para evaluar la importancia en la diferenciación genética intrapoblacional de los loci neutrales y los loci putativamente selectivos, se construyó un modelo; en éste se realizaron regresiones de tipo *Random Forest* utilizando como variables predictivas las frecuencias alélicas de todos los loci por localidad (hipótesis nula), las frecuencias alélicas de los loci putativamente selectivos detectados en los análisis previos (hipótesis selectiva) y las frecuencias alélicas de los loci putativamente neutrales, es decir todos los loci menos los putativamente selectivos (hipótesis neutral). Como variable de respuesta se utilizaron la primera y la segunda coordenada resultado del PCoA, cada una de manera independiente, por lo que se tienen dos regresiones para cada hipótesis. Para medir la bondad de ajuste de cada hipótesis se utilizó el porcentaje de la varianza explicada por las variables predictivas (análogo a una pseudo R cuadrada). La interpretación de este modelo es que si la variación genética neutral tiene una mayor importancia, entonces se puede concluir que el flujo genético observado está determinado, en su mayoría, por la capacidad de dispersión de la especie; ya que no existen procesos adaptativos suficientemente importantes como para imponer restricciones al establecimiento de los migrantes. Por el contrario, si la variación genética selectiva tiene una mayor importancia, entonces se puede concluir que el flujo genético observado está fuertemente influenciado por procesos adaptativos, que imponen restricciones al establecimiento de los migrantes.

Con el fin de identificar los loci de mayor relevancia en la diferenciación genética se calculó la importancia de cada locus en la hipótesis que mostró el mayor porcentaje de la varianza explicada. A partir de esta información se buscó detallar dicha hipótesis bajo un modelo basado en la mayor parsimonia posible, para ello se eliminaron una a una las variables de menor a mayor importancia, hasta tener un modelo que explique el mayor porcentaje de la varianza (genética) con el menor número posible de variables predictivas. El nivel de significancia de las hipótesis selectivas se evaluó a partir de la distribución nula de los valores de la varianza explicada por modelos aleatorios. Esta distribución se obtuvo realizando regresiones tipo *Random Forests* utilizando como variables predictivas las frecuencias alélicas de 14 loci tomados al azar a partir de la base de datos completa, y como variable de respuesta la primera y la segunda coordenada resultado del PCoA, de manera independiente. De esta manera se calculó el valor p como la proporción de las regresiones que llevan a un valor de la varianza explicada mayor o igual al observado en las hipótesis selectivas, el número de regresiones aleatorias utilizado fue 550. Para obtener un estadístico que refleje el desempeño de cada una de las hipótesis, se calculó el coeficiente de Pearson entre los eigenvectores predichos y los observados utilizando R v3.0.2 (R Core Team, 2013).

6. Resultados

Análisis moleculares

6.1 Extracción de muestras, amplificación y selección de marcadores AFLPs

Debido probablemente a la condición del tejido recolectado, la extracción de 11 muestras (de las 240 disponibles) no resultó en ADN de calidad aceptable, estas muestras fueron descartadas del estudio incluyendo la amplificación de AFLPs y microsatélites del cloroplasto, por lo tanto se retuvieron 229 individuos.

Se probaron trece combinaciones de cebadores, de los cuales ocho tuvieron una alta reproducibilidad (Anexo 2a).

Tabla 4. Resultados del análisis de reproducibilidad de AFLPs realizado en RawGeno (Arrigo *et al.*, 2009) para ocho combinaciones de cebadores. Número y porcentaje de fragmentos polimórficos por combinación de cebadores utilizando 229 individuos de *A. religiosa* distribuidos en doce localidades ubicadas en el Vocán Ajusco. Número de fragmentos finales (# Fragmentos Final) una vez que se eliminaron aquellos individuos cuyas reacciones negativas de PCR provocaron un exceso de falta de información (>20%). De esta manera se construyó la matriz binaria final con 213 individuos y 394 fragmentos polimórficos.

Combinación	1a	2c	3a	3c	3f	3g	3i	3j	TOTAL
# Fragmentos	26	57	12	104	30	72	39	78	418
Porcentaje	6.22	13.64	2.87	24.88	7.18	17.22	9.33	18.66	100
# Fragmentos Final	28	47	10	94	26	71	43	75	394

El análisis de reproducibilidad realizado en RawGeno (Arrigo *et al.*, 2009) constó de tres filtros y fue realizado en dos etapas. En la primera etapa el número final de fragmentos polimórficos fue de 418 (Tabla 4), posteriormente se eliminaron aquellos individuos con reacciones de PCR no exitosas (16 individuos en total). Con los individuos restantes (213) se repitió el análisis de reproducibilidad. Del total de fragmentos reportados en los electroferogramas para las ocho combinaciones de cebadores, el 53.03% fueron eliminados debido a que no superaron el umbral de intensidad, el 17.55% fueron eliminados debido a que no cumplieron los criterios de reproducibilidad entre las muestras genotipadas por duplicado y el 17.11% fueron eliminados debido a su baja frecuencia en la población. En total

se conservaron 394 fragmentos con los que se construyó la matriz binaria final (Tabla 4). La cantidad de fragmentos polimórficos por combinación de cebadores varió desde 10 (combinación 3a) hasta 94 (combinación 3c) (Tabla 4). Se encontró una correlación negativa y significativa entre el tamaño de los fragmentos y su frecuencia en las combinaciones 1a, 3f y 3g ($p = 0.011, 0.034, 0.002$, respectivamente; más información sobre los resultados del análisis de reproducibilidad Anexo 4a).

Análisis estadísticos

6.2 Diversidad genética

Un porcentaje elevado de los loci retenidos fueron polimórficos dentro de cada localidad de *Abies religiosa* (74.208%) muestreada en el volcán Ajusco. Sin embargo, la localidad “Las Vacas” mostró un porcentaje considerablemente bajo (32.083%) con respecto a las demás (Tabla 5). Como era de esperarse, todos los haplotipos fueron de copia única. Todas las localidades, a excepción nuevamente de “Las Vacas”, mostraron un valor alto para el índice $D - W$ (Tabla 5). El tamaño de muestra por localidad no tuvo una correlación significativa con el número de loci útiles ($p > 0.05$); sin embargo, sí la tuvo con el %*Pol* y con el índice $D - W$ ($R = 0.87, p < 0.001$ y $R = 0.86, p < 0.001$, respectivamente). Dicha significancia se perdió al excluir de la correlación a la localidad “Las Vacas” ($p > 0.05$). Ni el %*Pol* ni el índice $D - W$ se correlacionaron con el número de loci útiles ($p > 0.05$). El %*Pol* se correlacionó significativamente con el índice $D - W$ ($R = 0.793, p < 0.01$), pero esta correlación perdió significancia al excluir la localidad “Las Vacas” ($p > 0.05$).

Tabla 5. Diversidad genética de la población de *A. religiosa* del volcán Ajusco muestreada en 12 localidades. Se analizaron 213 individuos genotipados para 394 AFLPs. El número de loci útiles se refiere a aquellos loci que fueron informativos por encima del 95% para cada localidad. El porcentaje de loci útiles fue calculado en relación al número máximo posible de loci útiles (394). El porcentaje de loci polimórficos por localidad fue determinado en relación al número de loci útiles. Mientras que el índice $W - D$, relaciona la frecuencia del marcador en cada localidad con la frecuencia total en la población. Este índice se calculó utilizando cinco individuos tomados al azar en cada localidad.

Localidad	N	# Loci útiles	% Loci útiles	% Loci Polimórficos	Índice D – W
Apapaxtles	19	319	80.964	69.592	107.000
Basurero	17	285	72.335	76.140	116.250
Cabañas	18	366	92.893	72.131	90.500
Capilla	20	394	100.000	81.980	154.500
La Curva	18	309	78.426	81.553	118.000
La Fortaleza	18	276	70.051	69.565	124.500
Los Perros	19	276	70.051	86.594	128.250
Las Vacas	10	240	60.914	32.083	65.750
Sauco	18	212	53.807	81.132	93.000
Valle del Tezontle	19	304	77.157	68.750	120.000
La Y	19	175	44.416	88.000	102.750
Zacatonal	18	323	81.980	82.972	124.750
Media	17.750	289.917	73.583	74.208	112.104

Las localidades con mayor y menor diversidad de clorotipos fueron, respectivamente, “Sauco” y “Las Vacas” (Tabla 6). Se encontraron 65 clorotipos, de los cuales solo el 6% tuvieron una frecuencia por encima del 5%; de estos, 35 (53%) fueron exclusivos de un individuo, la localidad con mayor número de clorotipos exclusivos fue “Valle del Tezontle” (Tabla 6).

Tabla 6. Diversidad genética del ADN de cloroplasto de la población de *Abies religiosa* perteneciente al volcán Ajusco muestreada en doce localidades. Se analizaron 229 individuos genotipados para tres microsatélites.

Localidad	Número de Clorotipos	Diversidad de clorotipos	Clorotipos exclusivos de un individuo
Apapaxtles	14	0.963	4
Basurero	10	0.883	2
Cabañas	13	0.959	2
Capilla	9	0.858	3
La Curva	13	0.961	4
La Fortaleza	14	0.963	2
Los Perros	15	0.971	4
Las Vacas	8	0.850	2
Sauco	16	0.977	4
Valle del Tezontle	14	0.958	6
La Y	12	0.941	1
Zacatonal	13	0.948	1
TOTAL	65	0.956	35

6.3 Estructura genética

Los resultados del AMOVA, calculado con AFLPs, indicaron que la mayor parte de la variación genética se encontró dentro de las localidades (86.74%), el restante (13.26%) se debió a las diferencias entre localidades. El índice de fijación a nivel poblacional (F_{ST}) fue de 0.134.

Las diferenciación genética entre localidades y mediada por la dispersión de las semillas+polen (calculada con AFLPs) fue significativa en 63 de las 66 comparaciones (95.45%; Fig. 6, A); mientras que aquella mediada por la dispersión del polen (calculada con cpADN) solo fue significativa en 27 de ellas (40.9%; Fig. 6, B). El rango de valores de F_{ST} fue menor con AFLPs (0.04876 a 0.26741) que con cpADN (- 0.05347 a 0.50918); a pesar de encontrar algunos valores negativos con este marcador, ninguno de ellos fue significativo.

Ninguno de los dos tipos de marcadores mostró un patrón claro en la diferenciación genética entre localidades y su ubicación espacial. En la mayoría de las comparaciones, realizadas con AFLPs, se observaron diferencias significativas sin importar la ubicación de una localidad con respecto a otra. La localidad “Los Perros” mostró el mayor número de comparaciones no

significativas (3), mientras que “Apapaxtles fue la localidad más diferenciada, mostrando valores altos de F_{ST} con respecto a la mayoría de las otras (Fig. 6, A). Con cpADN la localidad más divergente fue “Capilla”, siendo la única que presentó diferencias significativas con todas las demás (Fig. 6, B). Con este marcador, la diferenciación genética, cuando fue significativa, no tuvo relación con la ubicación espacial; por ejemplo, la localidad más divergente (“Capilla”) mostró un valor muy elevado (0.433) al compararse con su vecina “Basurero”, este valor solo es superado por el obtenido en la comparación con “La Curva” (0.509) (Fig. 6, B).

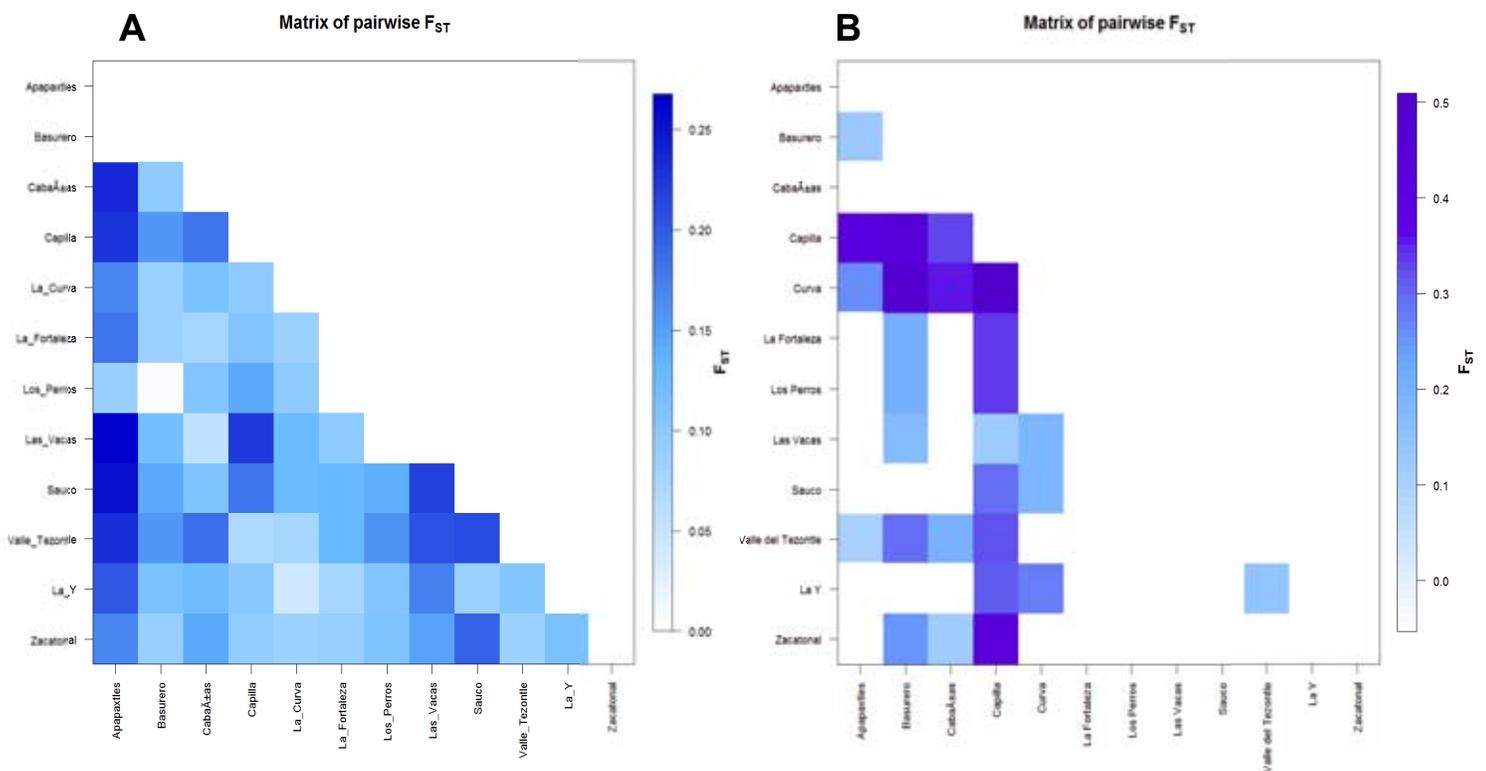


Figura 6. Distancias genéticas entre 12 localidades de *A. religiosa* ubicadas en el volcán Ajusco calculadas con dos tipos de marcadores moleculares. (A) Matriz pareada de F_{ST} utilizando 394 AFLPs (N=213). (B) Matriz de pareada de F_{ST} , análogo a R_{ST} , utilizando 3 microsatélites de cloroplasto (N=229). En blanco se eliminaron aquellas comparaciones que no fueron significativas después de utilizar la corrección de Bonferroni (A) y del valor q para múltiples comparaciones (B).

Los análisis realizados en Geneland v4.0.3 (Guillot *et al.*, 2005) y BAPS v6.0 (Corander *et al.*, 2008) mostraron resultados concordantes. Utilizando AFLPs se detectaron cuatro grupos

genéticos cuya distribución espacial fue a manera de parches. Si bien entre los dos análisis algunos individuos fueron asignados a distintos grupos genéticos, el patrón general resultó ser concordante (Fig. 7, B y D); por ejemplo, ambos análisis presentaron tres localidades cuyos individuos pertenecen a un solo grupo genético (“Las Vacas”, “La Curva” y “Cabañas”). Así mismo las localidades “Los Perros” y “Capilla” se dividen espacialmente en dos grupos genéticos, en los dos análisis (Fig. 7, B y D).

La población de *A. religiosa* perteneciente al volcán Ajusco esta compuesta solo por dos linajes de cloroplasto (Fig 7, C). “Capilla” fue la localidad con mayor número de individuos pertenecientes al linaje 2 (Fig 7, C; color rojo) el resto de los individuos pertenecientes a este linaje se encontraron esparcidos en una baja frecuencia dentro de cuatro localidades más.

El AMOVA realizado tomando en cuenta la estructura genética obtenida con Geneland v4.0.3 (Guillot *et al.*, 2005) indicó que la mayor parte de la variación genética se distribuyó dentro de las localidades (78.37%); mientras que el resto lo hace entre grupos genéticos (14.11%) y entre localidades dentro del mismo grupo genético (7.52%). El índice de fijación entre localidades dentro del mismo grupo genético (F_{SC}) fue de 0.08759, entre localidades sin importar el grupo genético (F_{ST}) fue de 0.2136 y entre grupos genéticos (F_{CT}) fue de 0.14107.

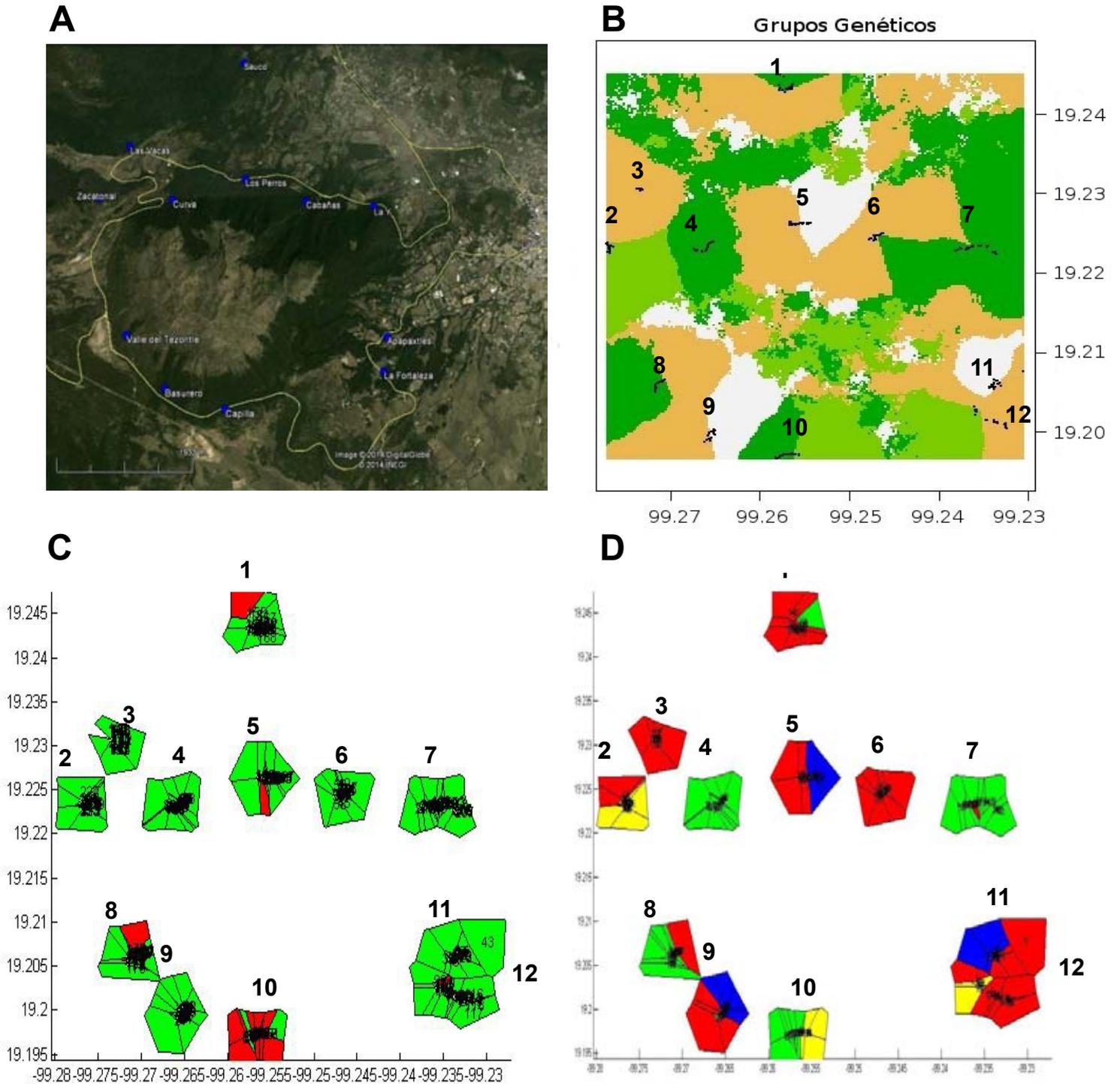


Figura 7. Estructura genética de la población de *A. religiosa* perteneciente al volcán Ajusco utilizando dos tipos de marcadores moleculares. (A) Mapa con la ubicación de las doce localidades muestreadas. (B) Distribución espacial de los cuatro grupos genéticos detectados por el programa Geneland v4.0.3 (Guillot *et al.*, 2005) utilizando 394 AFLPs (N=213): los puntos negros indican la ubicación espacial de cada individuo, los colores la

distribución espacial de cada grupo genético y por lo tanto, la pertenencia de cada individuo a determinado grupo genético. (C) Diagrama de Voronoi representando la ubicación espacial de los dos linajes de cloroplasto detectados por el programa BAPS v6.0 (Corander *et al.*, 2008) utilizando tres microsátélites de cloroplasto (N=229): cada color representa un linaje distinto. (D) Diagrama de Voronoi representando la ubicación espacial de los cuatro grupos genéticos detectados por el programa BAPS v6.0 (Corander *et al.*, 2008) utilizando 394 AFLPs (N=213): cada color representa un grupo genético.

Nombre de cada localidad: 1 = Sauco, 2 = Zacatonal, 3 = Las Vacas, 4 = La Curva, 5 = Los Perros, 6 = Cabañas, 7 = La Y, 8 = Valle del Tezontle, 9 = Basurero, 10 = Capilla, 11 = Apapaxtles, 12 = La Fortaleza.

El bajo número de individuos con mezcla y la relativa homogeneidad de los grupos genéticos indican que el flujo genético entre localidades es limitado (Fig. 8, A). La mayor parte de la migración se dio entre individuos pertenecientes al mismo grupo. Los grupos 1 y 2 recibieron migrantes de todos los grupos genéticos, el grupo 4 no recibió migrantes del grupo 3 pero si de los otros dos grupos, mientras que el grupo 3 solo donó y no recibe migrantes de ningún otro grupo (Fig. 8, B).

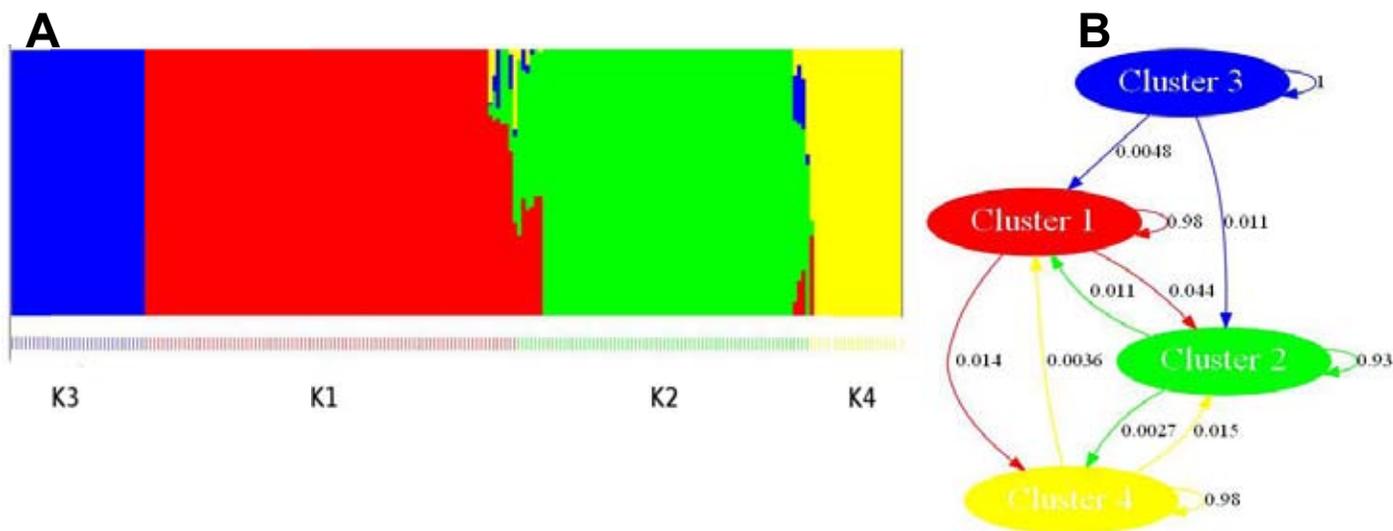


Figura 8. (A) Probabilidad de asignación de cada individuo de *Abies religiosa* (N=213, colectados en doce localidades distribuidas en el volcán Ajusco) a K grupos genéticos. Este análisis se realizó con 394 AFLPs bajo un modelo de mezcla. Cada línea vertical representa un individuo y el color indica la probabilidad de asignación a determinado grupo genético así como el porcentaje individual de mezcla. (B) Inferencias *a posteriori* sobre la intensidad y direccionalidad del flujo genético entre los cuatro grupos genéticos previamente detectados (B). Ver Fig. 7 (D) para la ubicación espacial de los grupos genéticos.

6.4 Modelo causal

Para realizar la matriz de distancia ambiental se utilizaron las dos primeras dimensiones obtenidas en el MCA. La primera dimensión explicó el 60.9% de la variación ambiental y se encontró compuesta principalmente por las variables “Tipo de Suelo” y “Tipo de vegetación”(30.6% y 31.6%, respectivamente; la segunda explicó el 29.44% y se encontró compuesta principalmente por la variable “Tipo de vegetación” (60.1%). En conjunto ambas dimensiones explicaron el 89.84% de la variación ambiental. Ninguna de las dos hipótesis puestas a prueba (es decir, IBD y IBA) explicó la diferenciación genética mediada por la dispersión de las semillas+polen (AFLPs). Por otro lado, la distancia ecológica entre localidades explicó la diferenciación genética mediada por la dispersión del polen (cpADN), aún cuando se controló por el efecto de la distancia geográfica (Tabla 7). La diferenciación genética producto de la dispersión de las semillas+polen generó un patrón distinto a aquel producto de la dispersión del polen. La distancia geográfica y la distancia ecológica mostraron una correlación positiva y significativa (Tabla 7), esta correlación no tiene implicaciones en el análisis, ya que al utilizar las pruebas de Mantel parcialse controla el efecto de una variable sobre otra (Wang & Summers, 2009).

Tabla 7. Resultados de las pruebas de Mantel y de Mantel parcial realizadas en el paquete vegan (Oksanen *et al.*, 2013) para R (R Core Team, 2013) con el fin de encontrar la hipótesis que mejor explica la diferenciación genética entre 12 localidades de *Abies religiosa* del volcán Ajusco, bajo el marco de un modelo causal. Las dos hipótesis puestas a prueba fueron: aislamiento por distancia (IBD) y aislamiento por adaptación (IBA). En negritas y cursivas se resaltan las asociaciones estadísticamente significativas ($p < 0.05$).

	R2	p
G.D	0.023	0.423
G.A	-0.005	0.483
A.D	0.507	0.001
g.G	0.094	0.338
G.D~A	0.030	0.451
G.A~D	-0.020	0.501
g.D	-0.126	0.774
g.A	0.499	0.015
g.D~A	-0.507	0.966
g.A~D	0.657	0.022

Abreviaciones: G, diferenciación genética calculada con AFLPs; g, diferenciación genética calculada con cpADN; D, distancia geográfica entre localidades; A, distancia ecológica entre localidades. El “~” indica la variable “z” utilizada en la prueba de Mantel parcial.

6.5 Detección de loci bajo selección natural

En su conjunto los cinco modelos utilizados bajo las tres aproximaciones distintas detectaron 28 loci con señales de selección natural. En Bayescan v2.1 (Foll & Gaggiotti, 2008) se probaron dos modelos, el primer modelo utilizó un intervalo para F_{IS} de 0.0 a 1.0 y detectó dos *outliers*, el segundo modelo utilizó un intervalo de F_{IS} de 0.192 a 0.216 (ver Múgica-Gallart, 2013; Aguirre-Planter *et al.*, 2000) y detectó tres *outliers*; dos de estos fueron los mismos que en el primer modelo (Anexo 5a), por lo que se decidió utilizar los resultados del segundo modelo. Los modelos jerárquicos de islas, implementados en el programa Arlequin v3.5.1.2 (Excoffier *et al.*, 2005), detectaron 9 (MJI-G), 22 (MJI-S) y 18 (MJI-V) *outliers* respectivamente (Anexo 5b). La aproximación basada en SAM, implementada en el programa matsam v2.0 (Joost *et al.*, 2007) fue realizada en dos etapas, en la primera se detectaron 12 loci putativamente selectivos, algunos sin valor de significancia para la prueba G. Luego de eliminar las muestras con falta de información, se calculó el valor de este estadístico y se confirmó el carácter putativamente selectivo de siete de los 12 loci identificados en la primer etapa (Anexo 5c).

Después de eliminar aquellos loci que no cumplieron con los criterios descritos previamente (ver Materiales y Método) se conservaron 14 loci, que constituyen el 3.55% del total de AFLPs genotipados (Tabla 8; Fig. 9). Estos incluyeron el 100% de los *outliers* detectados con Bayescan v2.1 (Foll & Gaggiotti, 2008), el 55.55% de los *outliers* detectados con MJI-G, el 50% de los *outliers* detectados con MJI-S, el 63.15% de los *outliers* detectados con MJI-V y el 100% de los loci putativamente selectivos detectados con SAM (Tabla 8; Fig. 9).

Las pruebas de Mantel no mostraron una correlación entre la diferenciación genética calculada con base en las frecuencias alélicas de los loci putativamente selectivos y la diferenciación ambiental o la distancia geográfica ($p > 0.05$, en ambos casos).

Tabla 8. Loci putativamente selectivos en la población de *A. religiosa* del volcán Ajusco. Cinco métodos fueron utilizados: Bayescan, modelo jerárquico de islas basado en la estructura genética detectada con Geneland (MJI-G), modelo jerárquico de islas basado en la estructura genética arbitraria con base en el tipo de suelo (MJI-S), modelo jerárquico de islas basado en la estructura genética arbitraria con base en el tipo de vegetación (MJI-V) y el método de análisis espacial (SAM). Estos loci fueron conservados ya que cumplieron con los criterios establecidos en el método. Para estos análisis se utilizaron 213 individuos genotipados para 394 AFLPs.

		Loci con señales de selección natural											
Bayescan						126					321	383	
MJI-G			79	82				191	197	199			
MJI-S	19		27	79	82	126		191	197	199	321	354	383
MJI-V	19	20	27	79	82	126		191	197	199	321	354	383
SAM	19	20	27				145	164				354	383

Loci con señales de selección natural

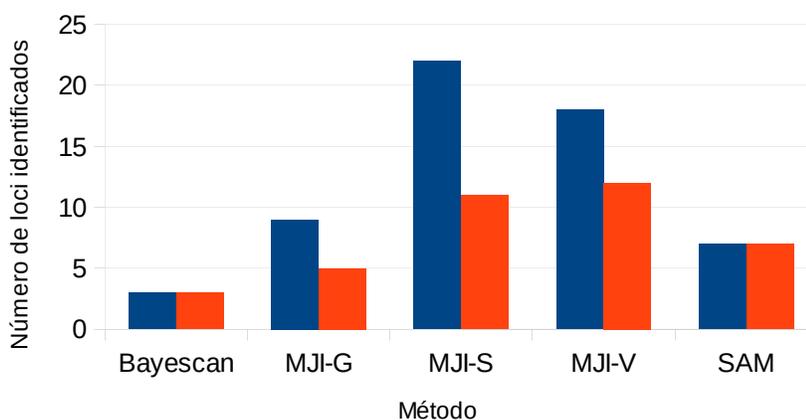


Figura 9. Número de loci putativamente selectivos en la población de *A. religiosa* del volcán Ajusco detectados con diferentes métodos. Se utilizaron cinco métodos: Bayescan, modelo jerárquico de islas basado en la estructura genética detectada con Geneland (MJI-G), modelo jerárquico de islas basado en la estructura genética arbitraria con base en el tipo de suelo (MJI-S), modelo jerárquico de islas basado en la estructura genética arbitraria con base en el tipo de vegetación (MJI-V) y el método de análisis espacial (SAM). En azul se muestra el total de loci detectados por cada método, en naranja los loci retenidos de acuerdo a los criterios descritos en el método. Para estos análisis se utilizaron 213 individuos genotipados para 394 AFLPs.

6.6 Procesos selectivos y diferenciación genética

Con el fin de contrastar la importancia de los procesos selectivos y neutrales en la diferenciación genética se construyó un modelo con el cual se pusieron a prueba tres hipótesis utilizando *Random Forest* (Breiman, 2001). La coordenada 1 del PCoA explicó el 27.64% de la diferenciación genética, mientras que la coordenada 2 explicó el 20.85%. Las hipótesis selectivas explicaron siempre el mayor porcentaje de la varianza en cuanto a la diferenciación genética intrapoblacional, mientras que las hipótesis nulas y neutrales explicaron en ambos casos un porcentaje de la varianza similar (Tabla 9). Para detallar la hipótesis selectiva con base en el modelo de mayor parsimonia posible se calculó la importancia de cada variable en cada una de las hipótesis selectivas (Fig. 10). Los modelos elegidos fueron el “Selectivo 9.1”, que utilizó como variable de respuesta la coordenada 1 y estuvo compuesto por las frecuencias alélicas de los loci 82, 79, 126, 321, 354 y 383 (Fig. 10); y el “Selectivo 9.2”, que utilizó como variable de respuesta la coordenada 2 y estuvo compuesto por las frecuencias alélicas de los loci 19, 164, 27, 197, 199, 20 (Fig. 10). Para mayor información sobre el desempeño de cada modelo probado véase el Anexo 6a. Ningún loci coincidió entre los dos modelos de mayor parsimonia. Ninguna de las 550 regresiones aleatorias explicó un porcentaje de la varianza mayor al explicado por la hipótesis selectiva, tres de éstas explican un porcentaje mayor al del modelo nulo, y solo una un porcentaje mayor al del modelo neutral (35.36%, 35.96% y 43.62%) (Anexo 6b). En conjunto estos resultados indican que los procesos selectivos tienen una mayor importancia en la diferenciación genética entre localidades que los procesos neutrales, como podrían ser el flujo y/o la deriva genética.

Tabla 9. Porcentaje de la varianza explicada por cada modelo realizado a partir de una regresión múltiple *Random Forest* (Breiman, 2001). Coeficiente de Pearson entre los eigenvectores predichos y observados y el valor de p como medida de significancia.

Modelo	Coordenada 1			Coordenada 2		
	Varianza Explicada %	R	p	Varianza Explicada %	R	p
Nulo	33.85	0.9799	2.45E-008	35.09	0.9793	2.85E-008
Neutral	36.30	0.9777	4.11E-008	33.32	0.9827	1.17E-008
Selectivo	59.45	0.9864	3.50E-009	60.82	0.9803	2.22E-008

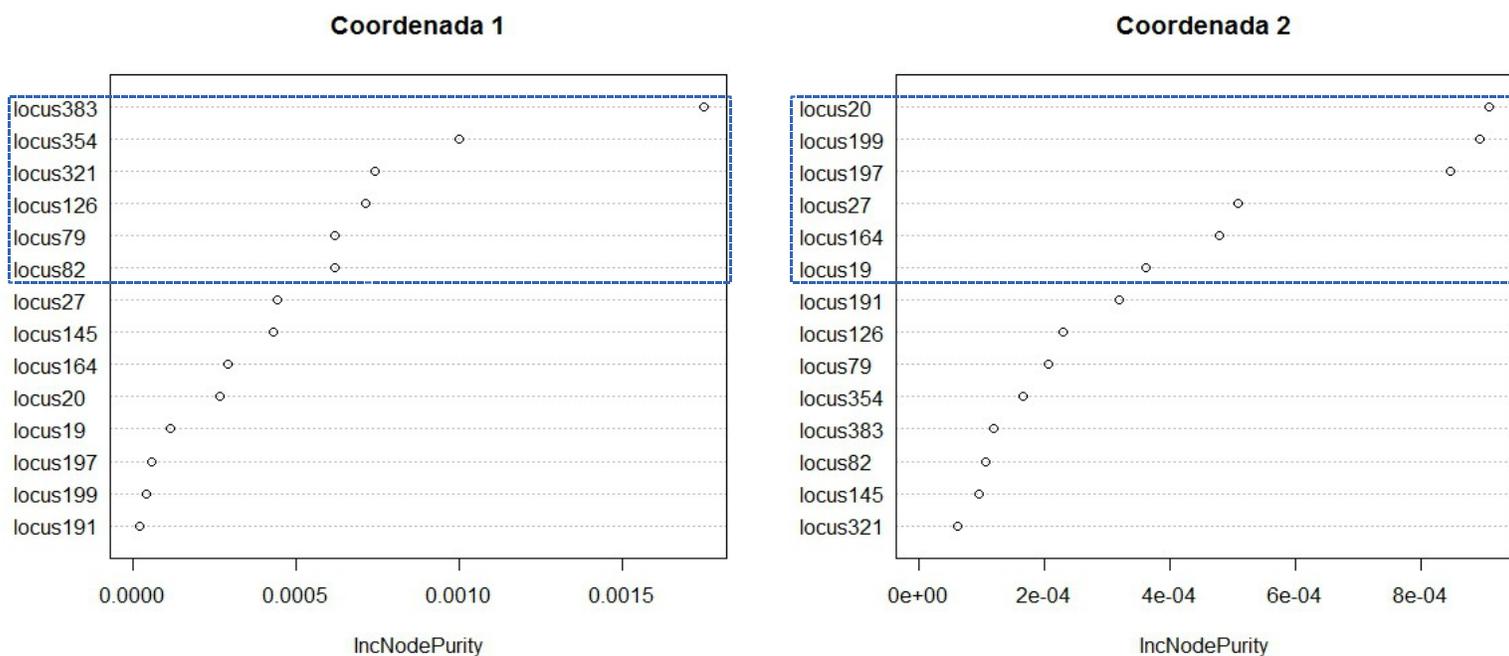


Figura 10. Importancia de cada locus putativamente selectivo en la predicción de la diferenciación genética. Esta medida fue obtenida realizando una regresión múltiple *Random Forest* (Breiman, 2001) entre las frecuencias alélicas de los loci putativamente selectivos y la primera (izquierda) y la segunda coordenada (derecha) resultado del PCoA de todos los loci. La importancia se evaluó a partir del incremento en la pureza del nodo (eje x) debido a la inclusión de cada variable predictiva. El recuadro punteado enmarca los loci contenidos en el modelo de mayor parsimonia. Nótese la diferencia en la escala del eje x.

7. Discusión

El curso de la evolución es afectado cualitativa y cuantitativamente por la subdivisión espacial de las poblaciones. Algunos patrones de evolución molecular solo pueden ser observados en poblaciones con estructura genética. Incluso en presencia de un régimen selectivo uniforme, el efecto de la deriva genética y la respuesta a la selección natural cambian sustancialmente (Whitlock, 2004). Utilizando una escala geográfica fina no sólo fue posible detectar, con una buena resolución, una fuerte estructura genética; sino que también se pudieron obtener conclusiones sobre cómo ésta determina el efecto de ciertas fuerzas evolutivas como la deriva genética y la selección natural. A continuación se discuten los resultados a detalle.

Diversidad genética nuclear

La población de *Abies religiosa* del volcán Ajusco presentó una diversidad genética alta utilizando AFLPs (%Pol = 74.21), promedio entre todas las localidades, con excepción de “Las Vacas” (%Pol = 32.08). El tamaño de muestra se correlacionó significativamente con el %Pol ($r = 0.87$, $p < 0.001$), lo que en parte explica la baja diversidad genética de “Las Vacas” donde apenas se colectaron 10 individuos. Esta diversidad fue similar o mayor a la estimada en otras especies de coníferas con distribuciones restringidas y utilizando los mismos marcadores, por ejemplo, *Picea chihuahuana*, especie endémica de México (%Pol = 74.92, en promedio para cinco poblaciones; Wehenkel & Sáenz-Romero 2012), *Picea asperata* (%Pol = 67.6; Xue *et al.*, 2005), y *Abies ziyuanensi* (%Pol = 35.73; Tang *et al.*, 2008), estas dos últimas endémicas de China. Comparada con la de algunas especies de amplia distribución, como *Pinus pinaster* (%Pol = 48.0; Mariette *et al.*, 2001) y *Pinus oocarpa* (%Pol = 37.86; Díaz *et al.*, 2001), la diversidad observada resultó considerablemente mayor. Cabe resaltar que estos estudios reportan la diversidad intraespecífica a partir de un enfoque interpoblacional, mientras que los valores aquí reportados son para una sola población. Sin embargo, al ser los primeros para *Abies religiosa* utilizando AFLPs, es la única comparación posible. La alta diversidad genética nuclear observada concuerda con lo reportado utilizando microsatélites (Mugica-Gallart, 2013), y podría ser explicada por ciertos rasgos de la especie,

como su carácter monoico, su largo tiempo generacional y la polinización y dispersión de sus semillas por viento (Loveless & Hamrick, 1984).

La actual distribución restringida de *Abies religiosa*, muy evidente en el área estudiada, esta relacionada con niveles bajos de diversidad genética (Loveless & Hamrick, 1984), y de hecho ha sido empleada, en conjunto con otros rasgos de la historia de vida, para explicar la baja diversidad observada en algunas especies de *Abies* mexicanos utilizando isoenzimas (Aguirre-Planter *et al.*, 2000). Lo anterior sugiere que la gran diversidad genética observada aquí y en estudios previos (Mugica-Gallart, 2013) es resultado tanto de características propias del marcador molecular empleado, como de la compleja interacción entre distintos rasgos de la historia de vida de la especie y de factores ambientales e históricos, de la especie y de la población, que podrían haber variado en el tiempo y el espacio. Por ejemplo, la distribución restringida actual podría ser una característica reciente, ya que es probable que en el pasado esta especie haya tenido una distribución más amplia y continua (Aguirre-Planter *et al.* 2000, Jaramillo-Correa *et al.*, 2008), lo que explicaría la cantidad de diversidad observada. Esto podría significar que un tamaño poblacional histórico alto ha permitido conservar una cantidad considerable de polimorfismos; y que los cuellos de botella por los que ha pasado esta especie no han sido lo suficientemente intensos o han sido muy recientes, por lo que su efecto aún no se refleja en la diversidad genética de la(s) población(es) (Jaramillo-Correa *et al.*, 2008).

Diversidad genética del cloroplasto

En concordancia con lo encontrado con los marcadores nucleares (AFLPs), la población estudiada presentó niveles de diversidad altos en el genoma del cloroplasto. La diversidad clorotípica (0.956) estuvo dentro del rango previamente reportado para esta especie utilizando los mismos marcadores (0.795 a 0.991; Jaramillo-Correa *et al.*, 2008). De los 65 clorotipos encontrados, el 53% se observaron solo en un individuo; el hecho de que la mayoría de los clorotipos se encuentren en una baja frecuencia concuerda con lo observado previamente para los *Abies* mexicanos (Jaramillo-Correa *et al.*, 2008). La gran diversidad genética del cpADN podría deberse a un gran tamaño poblacional histórico (Aguirre-Planter

et al. 2000, Jaramillo-Correa *et al.*, 2008); a una alta tasa de mutación, característica de este tipo de marcador molecular; y a un considerable nivel de flujo genético, situación presente en la población estudiada, donde solo el 40.9% de las F_{ST} pareadas fueron significativas; el cual podría deberse a la polinización por viento, que permite mantener niveles altos de variación genética e incluso introducir nuevos alelos a la población mediante eventos de dispersión a grandes distancias (Loveless & Hamrick, 1984).

Estructura genética mediada por la dispersión de las semillas

La población de *Abies religiosa* del volcán Ajusco presenta una estructura genética compleja. Los valores de F_{ST} pareados fueron significativos en el 95.45% de las comparaciones, lo que indica un alto grado de diferenciación genética entre ellas. Esto concuerda con los resultados del AMOVA, que indican que la diferenciación genética entre localidades fue mayor que entre grupos genéticos ($F_{ST} > F_{SC}$); aunque el mayor porcentaje de la diversidad genética se encontró dentro de las localidades (86.74% y 78.37%, sin tomar y tomando en cuenta la estructura genética, respectivamente). Por otro lado, los análisis bayesianos de agrupamiento indican la existencia de cuatro grupos genéticos, distribuidos a manera de parches y sin seguir ninguna tendencia lineal, lo que se refleja en la ausencia de un patrón de IBD.

Tomados en conjunto, estos resultados pueden ser explicados por la contingencia de la dinámica de colonización/extinción propia de un comportamiento metapoblacional. En un escenario de metapoblaciones, las localidades del Ajusco pudieron haber sido originadas a partir de un pequeño número de migrantes provenientes de pocas poblaciones, no necesariamente cercanas entre ellas (González-Martínez *et al.*, 2010). Esto explicaría la gran diferenciación entre localidades y la falta de relación entre los grupos genéticos y su ubicación espacial (González-Martínez *et al.*, 2010). Asimismo, la dinámica metapoblacional no permitiría alcanzar el equilibrio necesario para observar un patrón de IBD (Slatkin, 2000). Por otro lado, el efecto de fundador, causado por un limitado número de propágulos sexuales, aumentaría directamente la diferenciación genética entre localidades si existe un equilibrio entre el establecimiento de éstas en hábitats nuevos y la extinción de otras

(Senneville *et al.*, 2000). Resultados similares fueron observados en *Taxus canadensis*. En ésta especie la gran diferenciación genética, la heterogeneidad de las frecuencias alélicas y del nivel de heterocigosis entre poblaciones, así como el patrón errático de estructura genética tampoco siguieron ninguna tendencia lineal o estuvieron relacionados con algún tipo de variable geográfica (Senneville *et al.*, 2001). De igual manera, en *Taxus baccata* se recurrió a este concepto para explicar la gran diferenciación genética local, más no regional, y la ausencia de patrones de aislamiento por distancia (Hilfiker *et al.*, 2004; Dubreuil *et al.*, 2010; González-Martínez *et al.*, 2010).

Se puede inferir que en el Ajusco han existido las condiciones necesarias para provocar este tipo de dinámica poblacional. En esta zona se tiene evidencia de al menos cinco glaciaciones ocurridas durante los últimos 28,000 años (White, 1990). La alternancia de los períodos glaciales pudo haber provocado fluctuaciones en el tamaño poblacional y fragmentaciones en la distribución de la especie alrededor del Ajusco. Además de éstas, hay que tomar en cuenta la intensa actividad volcánica de la zona, como los episodios comprendidos entre las glaciaciones Santo Tomás y Albergue, que duraron aproximadamente 10,000 años, y la erupción del volcán Xitle hace aproximadamente 2,000 años (White, 1990). Con seguridad estos eventos provocaron la extinción de grandes partes de la población, y tomando en cuenta la naturaleza estocástica de dichos fenómenos, sería de esperarse que las extinciones ocurrieran en distintas zonas del volcán. La recolonización de estas áreas pudo haber sido a partir de diferentes localidades o poblaciones, no necesariamente las más cercanas, o incluso a partir de poblaciones vecinas, lo que habría originado la falta de relación entre los grupos genéticos y su ubicación espacial, y la gran diferenciación local observada para *Abies religiosa*. Si bien los eventos de dispersión a larga distancia son más bien raros, se sabe que estos tienen una gran importancia en el éxito de colonización de hábitats adecuados nuevos, y por lo tanto son especialmente relevantes para una dinámica metapoblacional (Whitlock, 2004).

Además de los patrones de diferenciación genética antes mencionados, observaciones hechas durante el muestreo soportan la hipótesis metapoblacional. A pesar de que el nivel de regeneración no fue medido cuantitativamente, en la mayoría de las localidades se encontró

una predominancia de árboles adultos que probablemente conforman una única cohorte. Esto sugiere que los individuos que conforman una misma localidad probablemente se establecieron simultáneamente. Si bien no existe una distribución de edades característica de bosques de coníferas maduros (Parker & Peet, 1984), se sugiere que una población estable es aquella en la que la mayoría de los individuos son juveniles (Juárez-Guerrero, 2010), ya que esto permite su continuidad. Al menos de manera cualitativa, la población estudiada no presenta esta distribución de edades, lo que sugiere que aún no alcanza su madurez (es decir, probablemente las localidades se establecieron en un tiempo geológico reciente), por lo que presiones ambientales o antropogénicas podrían amenazar su existencia. Se sabe que la regeneración de esta especie depende en primer lugar del área basal de la misma y en segundo de la edad del estrato superior (May, 2001); sin embargo, para desarrollar planes de manejo adecuados para esta población hace falta conocer la cantidad de reclutamiento necesario para sobrevivir bajo una dinámica metapoblacional (Senneville *et al.*, 2000).

Otra característica ecológica que podría contribuir al comportamiento metapoblacional en *Abies religiosa* es la dormancia de las semillas. Teóricamente las especies con una gran dormancia tendrían una mayor capacidad de sobrevivir mediante una dinámica metapoblacional, ya que la tasa de extinción y la tasa de colonización necesaria para mantener una metapoblación estable disminuyen conforme aumenta la dormancia (Husband & Barret, 1996). La dormancia fisiológica es una característica ampliamente distribuida en las coníferas, y previene la germinación durante el otoño-invierno para evitar una alta mortalidad debido a las bajas temperaturas (Kolotelo, 1997). Por ejemplo, en *Abies amabilis*, *A. grandis* y *A. lasiocarpa*, la mayor capacidad de germinación (70%) se da entre las 13 y 16 semanas a una humedad del 35% (Edwards, 1981). Además, se ha visto que *Abies balsamea* tiene una gran tolerancia a la sombra, por lo que puede sobrevivir bajo el dosel en poblaciones dominadas por otras especies de coníferas. Posteriormente perturbaciones naturales permiten el rápido crecimiento de esta especie aprovechando la baja competencia interespecífica (Blum *et al.*, 1981). Sin embargo, las características climáticas particulares de los bosques de *Abies* mexicanos, como la presencia de una estación seca bien definida y nevadas casi ausentes (Rzedowski, 2006), podrían implicar una dinámica de las semillas

distinta. En este sentido, es necesario estudiar las características de las semillas de los *Abies* mexicanos, incluyendo su porcentaje de germinación y dormancia, así como su capacidad para generar bancos de semillas y permanecer en el sotobosque luego de varias generaciones. Características que de ser similares a las de *Abies balsamea* podrían contribuir a la supervivencia a través de una dinámica metapoblacional y al mismo tiempo traducirse en niveles altos de diversidad genética.

La elevada diversidad genética nuclear de *Abies religiosa* podría servir como un reservorio genético, a partir del cual podrían originarse adaptaciones locales, posiblemente necesarias para la supervivencia a través de esta dinámica metapoblacional (véase más adelante). Por otro lado, contrario a lo esperado en un escenario metapoblacional, la diversidad genética es relativamente homogénea entre las localidades, con excepción de “Las Vacas” (aunque esto podría deberse al menor tamaño de muestra de esta localidad), lo que sugiere que todas tienen un tiempo similar de haberse establecido (Senneville *et al.*, 2000). Esto podría explicarse por los procesos históricos antes mencionados ocurridos en el Ajusco; si el último evento de esta naturaleza afectó de igual manera a las localidades muestreadas, entonces la recolonización de estas zonas tiene aproximadamente el mismo tiempo de ocurrencia, con la probable excepción de “Las Vacas”. A pesar de las diferencias en el tamaño de muestra, esta localidad también mostró el valor más bajo para el índice $D - W$ (65.75), lo que podría sugerir que se estableció recientemente, por lo que no ha tenido el tiempo suficiente para acumular una gran cantidad de mutaciones (Schönswetter & Tribsch 2005). Sin embargo, considerando un escenario metapoblacional y el largo tiempo generacional de los *Abies*, esta explicación parece poco probable. Es posible que en realidad esta localidad sea la que tenga mayor tiempo de haberse establecido; en dicho caso, la deriva genética podría haber provocado una pérdida importante de diversidad genética.

Si bien no existe una manera totalmente correcta de establecer los límites geográficos y las localidades dentro de una metapoblación, se deben considerar criterios acerca de la diferenciación ambiental de los hábitats ocupados, la estructura genética espacial y la capacidad de dispersión de la especie (Husband & Barrett, 1996). Se sabe que existe una fuerte estructura interpoblacional en los *Abies* mexicanos (Aguirre-Planter *et al.* 2000,

Jaramillo-Correa *et al.*, 2008, Mugica-Gallart, 2013). Asimismo, un modelado de la dispersión primaria de las semillas de *Abies alba* reveló que ésta se da mayormente de manera local, entre 13.2 a 19.2 metros (Sagnard *et al.*, 2006). Por lo tanto, la escala geográfica del presente estudio y la baja capacidad potencial de dispersión de *Abies religiosa*, justifican utilizar el concepto de metapoblación como una forma de obtener información relevante y proponer escenarios realistas sobre la demografía histórica de la población estudiada (Husband & Barrett, 1996).

Por otro lado, la localidad “Capilla” tuvo el mayor valor de divergencia para el índice $D - W$ (154.5), fue la única con valores de diferenciación significativos con respecto a todas las localidades utilizando cpADN y se encontró conformada, casi en su totalidad, por un solo linaje de cloroplasto distinto al del resto de las localidades. En conjunto estos resultados sugieren que esta localidad se originó a partir de una población genéticamente divergente con respecto a la(s) que fundaron las demás localidades. Otra posible explicación es que la deriva genética fijó este linaje en esta localidad, mientras que en las otras se encuentra en baja frecuencia o incluso ha sido eliminado. Es reconocido que distintos procesos históricos, como los mencionados anteriormente, han provocado una gran asimetría entre la pendiente norte y sur del volcán Ajusco, lo que ha tenido una fuerte influencia en la distribución de la vegetación (García-Aguirre *et al.* 2007). Sin embargo, aún queda por resolverse si esta asimetría ha tenido un impacto en los procesos de colonización a partir de diferentes poblaciones, y por lo tanto en la distribución espacial de la diversidad genética, como lo podrían sugerir los resultados de la localidad “Capilla”, ubicada en la pendiente sur del volcán y a solo un kilómetro de su vecino más cercano (“Basurero”) del cual está altamente diferenciado ($F_{ST}(\text{cpADN}) = 0.43272$ y $F_{ST}(\text{AFLP}) = 0.15611$).

Para poner a prueba explícitamente las hipótesis mencionadas anteriormente, como la del comportamiento metapoblacional, las diferencias en el tiempo de establecimiento de cada localidad y el posible origen de “Capilla” a partir de una poza genética distinta a la de las demás localidades, se podrían usar diversas aproximaciones, incluyendo las gráficas tipo Skyline (Pybus *et al.*, 2000), que permiten estimar patrones históricos sobre los tamaños poblacionales a partir de una genealogía construida con base en secuencias de nucleótidos y

que prescinden de restricciones *a priori* sobre posibles modelos demográficos (Ho & Shapiro, 2011). Estas serían especialmente útiles para investigar más acerca de la demografía histórica de esta población. Otra aproximación útil sería realizar simulaciones sobre la demografía histórica, tomando en cuenta tanto las características biológicas de la especie como la heterogeneidad ambiental, e integrarlas en simulaciones coalescentes (Knowles & Alvarado-Serrano, 2010). De esta forma sería posible examinar directamente la influencia del paisaje actual, la de la distribución histórica de la especie (o en este caso de la población) y la del proceso de colonización en sí mismo, sobre la diversidad y estructura genética contemporánea de *Abies religiosa* (ver por ejemplo Knowles & Alvarado-Serrano, 2010). Asimismo, podría utilizarse el Cómputo Bayesiano Aproximado (Beaumont *et al.*, 2002), que es un marco estadístico que permite poner a prueba hipótesis demográficas específicas a partir de la simulación de genealogías construidas con diversos tipos de marcadores moleculares (Bertorelle *et al.*, 2010). Esta aproximación permite comparar probabilísticamente varios modelos y evaluar la credibilidad de los resultados de manera cuantitativa (Bertorelle *et al.*, 2010).

Dispersión diferencial del polen y de las semillas

El índice de diferenciación genética entre localidades (F_{ST}) calculado con AFLPs fue significativo en 95.45% de las comparaciones, mientras que al utilizar cpADN solo fue significativo en 40.9% de ellas. Además, los resultados del modelo de mezcla realizado en BAPS para AFLPs, donde pocos individuos presentan un nivel de mezcla significativo y el flujo genético se da principalmente entre individuos del mismo grupo genético, sugieren que las semillas no tienen una gran capacidad de dispersión, como ya se ha señalado en *Abies alba* (Sagnard *et al.*, 2006), cuyas semillas también son dispersadas por viento, y para la gran mayoría de las coníferas. Tomados en conjunto, estos resultados sugieren que el flujo genético mediado por el polen es mayor que el mediado por las semillas. Dicho resultado es de esperarse, ya que la dispersión del polen a través del viento es más eficaz (Loveless & Hamrick, 1984). Por otra parte, las pruebas de Mantel realizadas bajo el marco de un modelo causal además indicaron, como era de esperarse, que la estructura genética mediada por la dispersión del polen es diferente de aquella mediada por la dispersión de las

semillas ($p > 0.05$)

A pesar de que el flujo genético mediado por el polen es alto, es posible detectar un patrón de estructura genética, aunque débil, dado por este tipo de dispersión. Debido a la escala espacial del estudio y a la gran capacidad de dispersión del polen, podría esperarse que el flujo genético tuviera la suficiente intensidad para homogeneizar las distintas localidades. En esta situación se esperaría que no existiera diferenciación genética entre ellas. Sin embargo, los valores de diferenciación genética calculados con cpADN, cuando significativos, fueron relativamente altos. Esto puede explicarse por la dispersión local de las semillas y por la dinámica metapoblacional mencionada anteriormente, estas reducen significativamente el número de árboles que contribuyen propágulos durante el proceso de recolonización y aunque haya una buena dispersión del polen, esta resulta insuficiente para compensar el efecto a corto plazo de la baja dispersión de las semillas (Jordano, 2010). Otro factor a tomar en cuenta es que dentro de la familia Pinaceae los granos de polen de los *Abies* tienen la mayor velocidad de descenso debido a efectos gravitacionales, característica que podría limitar la dispersión a larga distancia y el flujo genético entre localidades (Cinget, 2014). Esto resulta particularmente relevante si se considera la topografía del volcán, ya que limitaría en gran medida la dispersión del polen pendiente arriba.

Determinantes de la estructura genética poblacional

Las pruebas de Mantel y de Mantel parcial realizadas bajo el marco de un modelo causal no permitieron identificar el patrón adecuado para describir la estructura genética obtenida con los AFLPs. Tanto la hipótesis de aislamiento por distancia (IBD) (Wright, 1943) como la de aislamiento por adaptación (IBA) (Nossil *et al.*, 2009) fueron rechazadas ($p > 0.05$). La ausencia de IBD concuerda con lo esperado en el escenario de metapoblación, ya que los procesos de recolonización podrían darse a partir de migrantes provenientes de diferentes localidades, no necesariamente cercanas entre sí, resultado de eventos raros de dispersión a grandes distancias; o podrían darse a partir de individuos en estado de latencia presentes en el banco de semillas del subsuelo. A pesar de que los *Abies* mexicanos muestran una fuerte estructura geográfica interpoblacional, solo en un estudio se ha detectado un patrón de IBD

(Aguirre-Planter *et al.*, 2000), justamente para *Abies religiosa*. Esto podría deberse a que las poblaciones estudiadas no han existido por suficiente tiempo y/o que han experimentado diversos eventos demográficos, lo que les han impedido alcanzar un equilibrio deriva genética-migración (Aguirre-Planter *et al.*, 2000; Mugica-Gallart, 2013). Por otro lado, es reconocido que las determinantes de la estructura genética pueden variar a diferentes escalas geográficas, como se ha observado en *Abies alba* (Mosca *et al.*, 2014). Los resultados del presente estudio indican que a una escala geográfica fina el patrón de IBD no se encuentra en esta población de *Abies religiosa*. Sin embargo, esto no implica que al igual que en *Abies alba*, no existan patrones diferentes en otras poblaciones de la especie.

El hecho de no haber encontrado un patrón de IBA podría sugerir que el volcán Ajusco no presenta la heterogeneidad ambiental suficiente para inducir procesos de adaptación local. Sin embargo, esta explicación es poco probable ya que se reconoce que la historia natural de esta zona, en conjunto con las diferencias entre las características ambientales de las pendientes norte y sur tienen una gran influencia en la distribución de la vegetación (García-Aguirre *et al.* 2007). Es probable que la resolución de los datos ambientales haya disminuido nuestro poder estadístico para evaluar la hipótesis de IBA de manera correcta, ya que solo se usaron cuatro variables ambientales cualitativas. Esto resalta la necesidad de cuantificar y caracterizar la variación ambiental a niveles microespaciales, ya que procesos ecológicos y/o evolutivos importantes podrían estar ocurriendo a esta escala y solo teniendo la suficiente resolución en los datos ambientales sería posible identificarlos. Un enfoque que podría utilizarse, y se adapta a la escala espacial del estudio, sería la de caracterizar de manera cuantitativa los nutrientes del suelo; de esta forma sería posible identificar las variables ambientales con mayor varianza entre las localidades muestreadas. Este enfoque se utilizó anteriormente en *Helianthus petiolaris*, con una escala espacial similar a la del presente estudio, y permitió identificar elementos del suelo que están promoviendo la diferenciación genética a escala fina de esta especie (Andrew *et al.*, 2012).

A pesar de no haber encontrado evidencia de IBA y tomando en cuenta la dinámica metapoblacional, la posibilidad de eventos de adaptación local, resultado de la heterogeneidad ambiental y la compleja historia natural del volcán Ajusco, no queda

descartada. En este escenario los primeros colonizadores podrían haber desarrollado adaptaciones locales, que brindarían una ventaja sobre los migrantes; de esta manera el establecimiento de estos se vería disminuido, conservando así el efecto de fundador y aumentando la diferenciación genética (Orsini *et al.*, 2013). Sin embargo, este proceso (en el que el efecto de fundador se conserva y la diferenciación genética aumenta) también podría darse de manera neutral. En dicho escenario el efecto de fundador se conservaría gracias a la superioridad numérica de los primeros colonizadores, resultado del proceso de colonización, y al impacto de la deriva genética en las localidades (Orsini *et al.*, 2013). La baja capacidad de dispersión de las semillas y la presencia de loci con señales de selección sugieren que en la población estudiada podría haber ocurrido una combinación de ambos escenarios. De esta manera la adaptación a las condiciones locales, combinada con una ventaja numérica de los primeros colonizadores, contribuiría a un fuerte efecto de prioridad de los residentes sobre los migrantes (Orsini *et al.*, 2013).

Además de la variación ambiental, es muy probable que las características físicas del paisaje tengan una gran influencia en la distribución de la variación genética (Manel *et al.*, 2003). Esto puede ser particularmente importante en los ecosistemas montañosos debido a su complicada topografía, que aumenta el número de barreras físicas (Mosca *et al.*, 2014). La distribución de los manchones de bosque de *Abies religiosa* está muy relacionada con la presencia de cañadas, lo que podría determinar la intensidad y direccionalidad del flujo genético. Para medir la influencia de estas características físicas en el patrón de estructura genética de *Abies religiosa* sería necesario poner a prueba la hipótesis de aislamiento por resistencia (IBR) (McRae, 2006), en donde se prediga la relación entre la conectividad del paisaje y el flujo genético.

La estructura genética mediada por la dispersión del polen estuvo relacionada con el ambiente ($p < 0.05$) incluso luego de haber controlado el efecto de la distancia geográfica ($p < 0.05$). Cabe señalar que este patrón es contra intuitivo debido a la naturaleza neutral de los microsatélites. Una posible explicación es que el flujo genético diferencial entre hábitats ecológicamente divergentes (o similares), también podría influir en la variación genética neutral (Orsini *et al.*, 2013). Sin embargo, si este fuera el caso, se esperaría encontrar una

correlación entre los loci putativamente selectivos y la distancia ambiental, situación que no ocurrió ($p > 0.05$), aunque podría deberse a la poca resolución de los datos ambientales disponibles. Otra explicación es que, nuevamente, la baja resolución de los datos ambientales influyó en la correlación, provocando un falso positivo. Una forma de discernir entre causalidad y casualidad sería realizar experimentos de jardín común. Estos consistirían en examinar la distribución espacial de los haplotipos y seleccionar aquellos que se distribuyan únicamente en el sur, en el norte o lo hagan en ambas zonas. Posteriormente se podría probar si los individuos sembrados cerca de su ubicación “original” tienen mayor adecuación que cuando crecen en el lado opuesto; mientras que aquellos con haplotipos de amplia distribución no deberían de mostrar diferencias en la adecuación entre los dos sitios. Por otro lado, la hipótesis de IBR (McRae, 2006) también resultaría importante para hacer conclusiones acerca de la estructura genética mediada por la dispersión del polen en relación a sus determinantes.

Genómica ecológica y procesos selectivos

En total se encontraron 14 loci putativamente selectivos, que corresponden al 3.55% del total de AFLPs genotipados, este porcentaje se encuentra en concordancia con respecto al encontrado en otros estudios que han utilizado AFLPs (2.7%, McEwen *et al.*, 2013; 6% Richardson *et al.*, 2009) y SNPs (2.7%, Eckert *et al.* 2010, 4.4%, Prunier *et al.*, 2011 y 1.7% Mosca *et al.*, 2014). Dicho resultado es de esperarse debido a lo conservador del enfoque utilizado y demuestra las ventajas de emplear múltiples aproximaciones para identificar loci putativamente selectivos (De Mita *et al.*, 2013). Por otro lado, la diferencia en el número de loci detectados y conservados con cada método refleja la complejidad de encontrar con exactitud los loci responsables de la adaptación ambiental (Mosca *et al.*, 2014).

A pesar de haber utilizado un método basado en correlaciones entre variables ambientales y frecuencias alélicas (SAM), la resolución de los datos ambientales, y su carácter cualitativo, no permiten hacer una relación directa entre los loci detectados y la variable ambiental de interés. A pesar de estas limitantes, se pueden descifrar ciertas asociaciones. De los siete loci detectados con SAM, cinco estuvieron relacionados con la primera dimensión del MCA,

compuesta principalmente por el tipo de suelo y el tipo de vegetación (31.6% y 30.6 %, respectivamente); y los restantes con la segunda dimensión, compuesta principalmente por el tipo de vegetación (60.1%). Con respecto al tipo de suelo, hubo una serie de variables de gran importancia que no estaban disponibles para este estudio, como la acidez, la humedad relativa o el porcentaje de materia orgánica que podrían estar generando procesos microadaptativos en esta metapoblación de *Abies religiosa*. Este tipo de variables estuvieron fuertemente correlacionados con la diferenciación genética dentro de una población de *Helianthus petiolaris* (Andrew *et al.*, 2012).

Por otro lado, en nuestro estudio el tipo de vegetación estuvo fuertemente correlacionado con las dos dimensiones del MCA. A pesar de que esta variable es muy general, se esperaría que esté correlacionada con un gran número de variables ambientales, especialmente climáticas. Estudios previos han detectado que en general, la temperatura, el nivel de precipitación (Mosca *et al.*, 2012 y 2014) y el de aridez (Eckert *et al.*, 2010), representan importantes presiones adaptativas en diversas especies de coníferas. La escala espacial utilizada en la presente tesis hace difícil esperar una gran variación en estos factores dentro del área de estudio, por lo que es posible que a esta escala sean otras variables climáticas las que tengan mayor importancia. Sin embargo, la distribución de *Abies religiosa*, principalmente en cañadas podría significar la presencia de microclimas, cuya variación ambiental podría resultar en procesos selectivos. Tomando en cuenta tanto la ecología de los bosques de *Abies* mexicanos como la de los ecosistemas alpinos, algunas de estas variables podrían ser, por ejemplo, la exposición al viento y la cantidad de radiación solar; ambas variables seguramente dependen de la ubicación espacial alrededor del volcán y se relacionan directamente con la cantidad de evaporación (y por lo tanto de humedad). Estos resultados sugieren que incluso a esta escala microespacial procesos adaptativos importantes, posiblemente relacionados con variables climáticas y/o del suelo, podrían estar ocurriendo.

En el presente estudio se encontró una gran diferenciación genética entre localidades, debida probablemente al limitado flujo genético. Sin embargo, es necesario hacer notar que el flujo genético en sentido estricto depende tanto de la movilidad como de la capacidad de establecimiento de la especie; esta última determinada por los procesos adaptativos que

proceden a la dispersión. Construyendo un modelo basado en regresiones de tipo *Random Forest* (Breiman, 2001) se buscó evaluar la importancia de cada componente del flujo genético; esto se hizo contrastando el papel de la variación genética neutral contra la variación genética adaptativa en la diferenciación genética.

Los resultados del modelo indican que, las hipótesis selectivas de mayor parsimonia explican un mayor porcentaje de la varianza que las hipótesis neutrales. Además, ninguno de los 550 modelos aleatorios explican un porcentaje mayor o similar que los modelos selectivos. Esto sugiere que la diferenciación genética observada es mantenida por la interacción entre la baja capacidad de dispersión de las semilla y la ocurrencia de ciertos procesos adaptativos, que no se limitan a unos cuantos loci selectivos. Es decir, el limitado flujo genético detectado esta determinado, en cierta medida, por la selección en contra del migrante, y no depende exclusivamente de la limitada de dispersión de la especie . Esta baja capacidad de establecimiento podría ser resultado de la divergencia adaptativa entre localidades, y resalta la influencia del régimen selectivo en el flujo genético dentro la población estudiada. Sin embargo, esta interpretación debe tomarse con cautela por las siguientes razones; i) las regresiones de tipo Random Forest (Breiman, 2001) han sido utilizadas, comúnmente, para detectar las variables de mayor importancia; por ejemplo, para elegir genes candidatos a partir de datos de microarreglos (Díaz-Uriarte & Álvarez de Andres, 2006) y no para contrastar modelos utilizando el porcentaje de la varianza explicada; ii) algunos de los loci putativamente selectivos fueron escogidos por presentar valores inusualmente altos para el coeficiente de diferenciación F_{ST} , por lo tanto es de esperarse que dichos loci tengan una mayor importancia en la diferenciación entre localidades. Sin embargo, hay que notar que los modelos de mayor parsimonia para la coordenada uno y dos incluyeron loci detectados con el método SAM, que no esta basado en estas medidas de diferenciación.

Los dos modelos de máxima parsimonia fueron los “Selectivo 9.1” y “Selectivo 9.2”. Para la coordenada 1 este estuvo compuesto por los loci 82, 79, 126, 321, 354 y 383, mientras que para la coordenada 2 este incluyó los loci 19, 164, 27, 197, 199, 20. Debido a la naturaleza predominantemente no codificante de los AFLPs es probable que estos loci en realidad se encuentren ligados a genes o regiones reguladoras, que serían los verdaderos blancos de las

fuerzas selectivas. Para poder saber cuales son estos genes o regiones reguladoras se podrían realizar dos procedimientos. El primero sería aislar el fragmento de interés y secuenciarlo. Posteriormente se podría alinear esta secuencia para buscar su identidad putativa mediante una base de datos de referencia (BLAST). El segundo sería realizar un mapeo genético del fragmento de interés para identificar la región a la cual se encuentra ligada. De hecho, la técnica de AFLPs es considerada como una alternativa poderosa para realizar mapeos genómicos de alta densidad (Marques *et al.*, 1998), y ha sido utilizada con este fin en diversas especies de coníferas, por ejemplo, *Pinus taeda* (Remington *et al.*, 1998), *Picea abies* (Paglia *et al.*, 1998) y *Pinus pinaster* (Chagne *et al.*, 2002).

Como ya se mencionó antes, la estructura genética observada puede ser resultado de un comportamiento tipo metapoblación. En este escenario, aunque no se descarta la existencia de procesos de adaptación local, se da más importancia a procesos neutrales como la deriva genética, y a la contingencia de la dinámica de colonización/extinción. Por otro lado, el modelo basado en *Random Forest* (Breiman, 2001) sugiere que la diferenciación genética entre localidades también esta influenciada por procesos selectivos. Una posible explicación para unificar ambas hipótesis es que la gran variación genética observada posibilita el surgimiento de adaptaciones locales incluso bajo una dinámica metapoblacional. Bajo este escenario los primeros colonizadores en establecerse serían los que tuvieran una mayor adecuación para cada sitio determinado y/o tendrían una mayor adecuación con respecto a migrantes futuros (Orsini *et al.*, 2013). Si los alelos adaptativos se originan a partir de la variación genética preexistente, entonces el proceso de adaptación sería relativamente rápido; ya que dichos alelos ventajosos ya estarían presentes en la población, posiblemente con una frecuencia inicial alta, lo que aumentaría de manera importante su probabilidad de fijación (Barret & Schluter, 2007). De esta forma, aunque la dinámica metapoblacional podría ocasionar la extinción de las localidades en un tiempo relativamente corto, esto no sería un impedimento para el surgimiento de adaptaciones locales. Una forma de comprobar si los alelos adaptativos surgieron a partir de la variación genética pre existente, y no son producto de nuevas mutaciones, sería demostrar que el alelo fijado en el nuevo ambiente también se encuentra en la población ancestral en forma de variación genética preexistente (Barret & Schluter, 2007). Otra aproximación sería examinar la distribución espacial de los alelos

candidatos, si estos son resultado de mutaciones nuevas, entonces se esperaría que mostrarán señales de “barridos selectivos fuertes”, y por lo tanto tuvieran una distribución geográfica estrecha (Hancock *et al.*, 2011).

A pesar de las dificultades para interpretar el modelo basado en regresiones tipo *Random Forest* (Breiman, 2001) y del poco conocimiento que se tiene con respecto a la importancia de los procesos neutrales contra los procesos selectivos en la diferenciación genética, una explicación puramente demográfica y neutralista sería insuficiente dada la gran variación y complicada estructura genética observada en la población de *Abies religiosa* del Ajusco. Por ejemplo, en *Pinus monticola* existe evidencia de que la presencia de dos ecotipos podría ser resultado de la expansión postglacial a partir de dos refugios. Sin embargo, utilizando loci de AFLPs putativamente selectivos y un modelo basado en regresiones tipo *Random Forest* (aunque dicho modelo es distinto al utilizado en el presente estudio), se encontró evidencia de selección divergente aún en presencia de flujo genético. Lo anterior indica que la estructura genética de dicha especie no puede ser explicada exclusivamente con base en procesos demográficos y, en cambio, es resultado de la compleja interacción entre estos y la presencia de adaptaciones locales (Richardson *et al.*, 2009).

8. Conclusión

En el presente estudio se describió la estructura genética de la población de *Abies religiosa* del volcán Ajusco, para posteriormente identificar algunas de sus determinantes. Esta población se caracteriza por tener una diversidad genética elevada, tanto nuclear como cloroplástica; y estar subdividida espacialmente en cuatro grupos genéticos y dos linajes de cloroplasto. Los resultados obtenidos indican que los factores paisajísticos, en este caso la distancia geográfica y la distancia ambiental, no son los responsables de la estructura genética observada. En cambio, esta se debe a la compleja interacción entre procesos demográficos (neutrales) y selectivos. La complicada historia natural del volcán Ajusco probablemente ha provocado un comportamiento tipo metapoblación en las localidades de *Abies religiosa*. El carácter estocástico de la dinámica de extinción/colonización, en conjunto con la baja capacidad de dispersión de la especie, en la que los eventos de dispersión a larga distancia son más bien raros, y la presencia de procesos selectivos que influyen en el régimen de flujo genético, a la vez que permiten el surgimiento de adaptaciones locales, potencialmente reforzaron el efecto de fundador, lo que a su vez ha provocado una gran diferenciación genética entre las localidades.

Tanto el índice de diferenciación entre localidades (F_{ST} pareada), como los análisis Bayesianos de agrupamiento sugieren que, conforme a lo esperado, la dispersión del polen es más efectiva que la de las semillas. Sin embargo, a pesar de la poca distancia entre las localidades muestreadas, fue posible encontrar diferenciación genética significativa entre algunas localidades utilizando cpADN. Lo anterior sugiere que a pesar de que el flujo genético mediado por polen es más eficiente que el de las semillas, su eficiencia se ve limitada debido a la poca dispersión de estas últimas, así como al comportamiento metapoblacional.

Se detectaron 14 loci candidatos a ser adaptativos y relacionados, principalmente, con el tipo de suelo y el tipo de vegetación. La dinámica metapoblacional de colonización/extinción podría sugerir una baja probabilidad para el surgimiento de adaptaciones locales intrapoblacionales, debido al corto tiempo entre el establecimiento de una localidad y su

extinción. A pesar de ésto, es posible que la gran variación genética observada sirva como un reservorio genético a partir del cual se originaran dichas adaptaciones en un período de tiempo relativamente corto; de esta manera la dinámica metapoblacional no representaría un impedimento para el surgimiento de adaptaciones locales intrapoblacionales. A pesar de las limitaciones del presente estudio, la evidencia aquí colectada sugiere la necesidad de acumular datos más finos sobre la variación ambiental, para así poder entender de mejor manera los procesos evolutivos que ocurren a esta escala espacial.

La alta diferenciación genética observada anteriormente en los *Abies* mexicanos ha sido explicada por una fuerte influencia de la deriva genética y una cantidad limitada de flujo genético. En el presente estudio se utilizó un modelo para analizar la importancia de cada componente del flujo genético en sentido estricto, es decir, de la capacidad de dispersión y de la de establecimiento. Los resultados sugieren que la divergencia genética entre localidades no se debe exclusivamente a la baja capacidad de dispersión de la especie; y que el efecto homogeneizador del flujo genético se ve contrarrestado, en cierta medida, por procesos selectivos que limitan el establecimiento de los migrantes.

Anexo 1. Recolección de muestras

Anexo 1a. Geo referencia de cada uno de los 240 individuos de *A. religiosa* colectados en doce localidades distribuidas en el volcán Ajusco, durante diciembre del 2013.

Individuo	Geo referencia		Individuo	Geo referencia		Individuo	Geo referencia				
Sa1	N	19	14.69	Y1	N	19	13.41	ZA1	N	19	13.36
	W	99	15.44		W	99	14.17		W	99	16.61
Sa2	N	19	14.70	Y2	N	19	13.41	ZA2	N	19	13.37
	W	99	15.46		W	99	14.17		W	99	16.62
Sa3	N	19	14.67	Y3	N	19	13.40	ZA3	N	19	13.37
	W	99	15.45		W	99	14.18		W	99	16.61
Sa4	N	19	14.62	Y4	N	19	13.40	ZA4	N	19	13.39
	W	99	15.47		W	99	14.19		W	99	16.60
Sa5	N	19	14.59	Y5	N	19	13.39	ZA5	N	19	13.39
	W	99	15.47		W	99	14.21		W	99	16.61
Sa6	N	19	14.60	Y6	N	19	13.39	ZA6	N	19	13.40
	W	99	15.46		W	99	14.21		W	99	16.61
Sa7	N	19	14.58	Y7	N	19	13.38	ZA7	N	19	13.39
	W	99	15.44		W	99	14.22		W	99	16.61
Sa8	N	19	14.57	Y8	N	19	13.41	ZA8	N	19	13.40
	W	99	15.43		W	99	14.13		W	99	16.61
Sa9	N	19	14.59	Y9	N	19	13.41	ZA9	N	19	13.41
	W	99	15.42		W	99	14.12		W	99	16.61
Sa10	N	19	14.59	Y10	N	19	13.40	ZA10	N	19	13.41
	W	99	15.41		W	99	14.08		W	99	16.61
Sa11	N	19	14.60	Y11	N	19	13.41	ZA11	N	19	13.42
	W	99	15.41		W	99	14.07		W	99	16.62
Sa12	N	19	14.60	Y12	N	19	13.36	ZA12	N	19	13.42
	W	99	15.41		W	99	14.04		W	99	16.62
Sa13	N	19	14.59	Y13	N	19	13.36	ZA13	N	19	13.43
	W	99	15.40		W	99	14.04		W	99	16.62
Sa14	N	19	14.59	Y14	N	19	13.36	ZA14	N	19	13.43
	W	99	15.39		W	99	14.03		W	99	16.62
Sa15	N	19	14.61	Y15	N	19	13.36	ZA15	N	19	13.44
	W	99	15.39		W	99	14.02		W	99	16.62
Sa16	N	19	14.61	Y16	N	19	13.35	ZA16	N	19	13.43
	W	99	15.39		W	99	14.01		W	99	16.62
Sa17	N	19	14.61	Y17	N	19	13.39	ZA17	N	19	13.41
	W	99	15.38		W	99	14.25		W	99	16.60
Sa18	N	19	14.63	Y18	N	19	13.39	ZA18	N	19	13.41
	W	99	15.38		W	99	14.26		W	99	16.59
Sa19	N	19	14.61	Y19	N	19	13.38	ZA19	N	19	13.41
	W	99	15.38		W	99	14.28		W	99	16.59
Sa20	N	19	14.63	Y20	N	19	13.38	ZA20	N	19	13.40
	W	99	15.38		W	99	14.29		W	99	16.58

Individuo	Geo referencia			Individuo	Geo referencia			Individuo	Geo referencia		
Ca1	N	19	13.51	LP1	N	19	13.92	VT1	N	19	12.31
	W	99	14.77		W	99	15.41		W	99	16.30
Ca2	N	19	13.51	LP2	N	19	13.59	VT2	N	19	12.31
	W	99	14.78		W	99	15.40		W	99	16.30
Ca3	N	19	13.50	LP3	N	19	13.59	VT3	N	19	12.33
	W	99	14.79		W	99	15.39		W	99	16.30
Ca4	N	19	13.49	LP4	N	19	13.59	VT4	N	19	12.33
	W	99	14.80		W	99	15.40		W	99	16.30
Ca5	N	19	13.49	LP5	N	19	13.57	VT5	N	19	12.35
	W	99	14.80		W	99	15.39		W	99	16.30
Ca6	N	19	13.49	LP6	N	19	13.57	VT6	N	19	12.35
	W	99	14.82		W	99	15.40		W	99	16.30
Ca7	N	19	13.49	LP7	N	19	13.57	VT7	N	19	12.34
	W	99	14.82		W	99	15.38		W	99	16.29
Ca8	N	19	13.48	LP8	N	19	13.57	VT8	N	19	12.36
	W	99	14.83		W	99	15.37		W	99	16.30
Ca9	N	19	13.49	LP9	N	19	13.57	VT9	N	19	12.37
	W	99	14.83		W	99	15.36		W	99	16.30
Ca10	N	19	13.48	LP10	N	19	13.57	VT10	N	19	12.37
	W	99	14.83		W	99	15.35		W	99	16.29
Ca11	N	19	12.46	LP11	N	19	13.58	VT11	N	19	12.38
	W	99	13.83		W	99	15.34		W	99	16.28
Ca12	N	19	13.47	LP12	N	19	13.58	VT12	N	19	12.38
	W	99	14.83		W	99	15.33		W	99	16.29
Ca13	N	19	13.47	LP13	N	19	13.58	VT13	N	19	12.38
	W	99	14.83		W	99	15.32		W	99	16.28
Ca14	N	19	13.45	LP15	N	19	13.58	VT14	N	19	12.38
	W	99	14.84		W	99	15.32		W	99	16.27
Ca15	N	19	13.46	LP16	N	19	13.59	VT15	N	19	12.39
	W	99	14.85		W	99	15.30		W	99	16.26
Ca16	N	19	13.47	LP17	N	19	13.58	VT16	N	19	12.39
	W	99	14.85		W	99	15.29		W	99	16.25
Ca17	N	19	13.47	LP18	N	19	13.58	VT17	N	19	12.39
	W	99	14.86		W	99	15.27		W	99	16.25
Ca18	N	19	13.46	LP19	N	19	13.58	VT18	N	19	12.40
	W	99	14.87		W	99	15.27		W	99	16.25
Ca19	N	19	13.47	LP20	N	19	13.57	VT19	N	19	12.40
	W	99	14.86		W	99	15.28		W	99	16.24
Ca20	N	19	13.46	LP21	N	19	13.58	VT20	N	19	12.40
	W	99	14.82		W	99	15.26		W	99	16.24

Individuo	Geo referencia			Individuo	Geo referencia			Individuo	Geo referencia		
LV1	N	19	13.84	Cur1	N	19	13.39	Cap1	N	19	11.82
	W	99	16.36		W	99	16.04		W	99	15.49
LV2	N	19	13.83	Cur2	N	19	13.38	Cap2	N	19	11.81
	W	99	16.38		W	99	16.03		W	99	15.49
LV3	N	19	13.84	Cur3	N	19	13.37	Cap3	N	19	11.82
	W	99	16.40		W	99	16.03		W	99	15.47
LV4	N	19	13.83	Cur4	N	19	13.38	Cap4	N	19	11.82
	W	99	16.40		W	99	16.02		W	99	15.46
LV5	N	19	13.83	Cur5	N	19	13.36	Cap5	N	19	11.83
	W	99	16.40		W	99	16.00		W	99	15.46
LV6	N	19	13.83	Cur6	N	19	13.38	Cap6	N	19	11.83
	W	99	16.40		W	99	16.00		W	99	15.44
LV7	N	19	13.84	Cur7	N	19	13.37	Cap7	N	19	11.84
	W	99	16.39		W	99	15.98		W	99	15.44
LV8	N	19	13.84	Cur8	N	19	13.36	Cap8	N	19	11.84
	W	99	16.40		W	99	15.98		W	99	15.43
LV9	N	19	13.84	Cur9	N	19	13.37	Cap9	N	19	11.84
	W	99	16.39		W	99	15.97		W	99	15.41
LV10	N	19	13.84	Cur10	N	19	13.38	Cap10	N	19	11.84
	W	99	16.40		W	99	15.96		W	99	15.42
LV11	N	19	13.84	Cur11	N	19	13.39	Cap11	N	19	11.84
	W	99	16.40		W	99	15.95		W	99	15.41
LV12	N	19	13.84	Cur12	N	19	13.40	Cap12	N	19	11.84
	W	99	16.40		W	99	15.95		W	99	15.40
LV13	N	19	13.85	Cur13	N	19	13.42	Cap13	N	19	11.84
	W	99	16.42		W	99	15.95		W	99	15.39
LV14	N	19	13.85	Cur14	N	19	13.42	Cap14	N	19	11.85
	W	99	16.42		W	99	15.94		W	99	15.39
LV15	N	19	13.84	Cur15	N	19	13.43	Cap15	N	19	11.85
	W	99	16.42		W	99	15.94		W	99	15.38
LV16	N	19	13.84	Cur16	N	19	13.43	Cap16	N	19	11.84
	W	99	16.42		W	99	15.94		W	99	15.37
LV17	N	19	13.84	Cur17	N	19	13.44	Cap17	N	19	11.84
	W	99	16.41		W	99	15.93		W	99	15.36
LV18	N	19	13.84	Cur18	N	19	13.43	Cap18	N	19	11.83
	W	99	16.41		W	99	15.92		W	99	15.34
LV19	N	19	13.84	Cur19	N	19	13.44	Cap19	N	19	11.84
	W	99	16.41		W	99	15.92		W	99	15.33
LV20	N	19	13.84	Cur20	N	19	13.44	Cap20	N	19	11.84
	W	99	16.41		W	99	15.91		W	99	15.34

Individuo	Geo referencia			Individuo	Geo referencia			Individuo	Geo referencia		
LF1	N	19	12.20	APA1	N	19	12.36	BA1	N	19	11.93
	W	99	14.16		W	99	14.05		W	99	15.97
LF2	N	19	12.18	APA2	N	19	12.35	BA2	N	19	11.95
	W	99	14.15		W	99	14.05		W	99	15.96
LF3	N	19	12.17	APA3	N	19	12.34	BA3	N	19	11.95
	W	99	14.16		W	99	14.04		W	99	15.95
LF4	N	19	12.17	APA4	N	19	12.34	BA4	N	19	11.94
	W	99	14.16		W	99	14.06		W	99	15.93
LF5	N	19	12.16	APA5	N	19	12.35	BA5	N	19	11.95
	W	99	14.15		W	99	14.06		W	99	15.94
LF6	N	19	12.14	APA6	N	19	12.33	BA6	N	19	11.96
	W	99	14.16		W	99	14.03		W	99	15.94
LF7	N	19	12.13	APA7	N	19	12.35	BA7	N	19	11.96
	W	99	14.15		W	99	14.02		W	99	15.93
LF8	N	19	12.10	APA8	N	19	12.35	BA8	N	19	11.97
	W	99	14.09		W	99	14.04		W	99	15.92
LF9	N	19	12.10	APA9	N	19	12.36	BA9	N	19	11.97
	W	99	14.08		W	99	14.05		W	99	15.91
LF10	N	19	12.10	APA10	N	19	12.36	BA10	N	19	11.99
	W	99	14.03		W	99	14.04		W	99	15.90
LF11	N	19	12.08	APA11	N	19	12.37	BA11	N	19	12.00
	W	99	14.03		W	99	14.03		W	99	15.91
LF12	N	19	12.08	APA12	N	19	12.38	BA12	N	19	12.01
	W	99	14.02		W	99	14.03		W	99	15.91
LF13	N	19	12.07	APA13	N	19	12.40	BA13	N	19	12.02
	W	99	14.01		W	99	14.03		W	99	15.90
LF14	N	19	12.08	APA14	N	19	12.39	BA14	N	19	12.02
	W	99	14.00		W	99	14.01		W	99	15.91
LF15	N	19	12.08	APA15	N	19	12.39	BA15	N	19	12.02
	W	99	13.99		W	99	14.00		W	99	15.91
LF16	N	19	12.08	APA16	N	19	12.37	BA16	N	19	12.01
	W	99	13.96		W	99	14.01		W	99	15.92
LF17	N	19	12.07	APA17	N	19	12.37	BA17	N	19	12.00
	W	99	13.96		W	99	14.00		W	99	15.94
LF18	N	19	12.06	APA18	N	19	12.38	BA18	N	19	11.99
	W	99	13.95		W	99	13.99		W	99	15.94
LF19	N	19	12.05	APA19	N	19	12.37	BA19	N	19	11.99
	W	99	13.95		W	99	13.98		W	99	15.95
LF20	N	19	12.03	APA20	N	19	12.36	BA20	N	19	11.98
	W	99	13.94		W	99	13.99		W	99	15.95

Anexo 2. Métodos de laboratorio para la obtención de AFLPs

Anexo 2a. Secuencias de los cebadores utilizados, se señalan los oligonucleótidos selectivos de cada cebador. Estas combinaciones de cebadores y la secuencia de oligonucleótidos selectivos fueron seleccionados a partir de estudios previos que los reportan como altamente polimórficos y reproducibles en especies de árboles (Hasbún *et al.*, 2011). En negritas se señalan las combinaciones seleccionadas.

Oligonucleótidos selectivos	Clave	Combinaciones probadas
<i>Eco</i> + ACA	1	1a
<i>Eco</i> + ACC	2	1b
<i>Eco</i> + ACG	3	1c
<i>Mse</i> + AGCT	a	2d
<i>Mse</i> + CCCG	b	2c
<i>Mse</i> + CCGA	c	2e
<i>Mse</i> + AGGC	d	3f
<i>Mse</i> + CCTA	e	3a
<i>Mse</i> + AGCA	f	3c
<i>Mse</i> + CCAA	g	3g
<i>Mse</i> + CCCA	h	3h
<i>Mse</i> + CCTC	i	3i
<i>Mse</i> + CCTG	j	3j

Nota: Secuencia del cebador *Eco* (GACTGCGTACCAATTC), secuencia del cebador *Mse* (GATGAGTCCTGAGTAA).

Anexo 2b. Detalles del procedimiento realizado para obtener la matriz binaria final.

Los electroforegramas obtenidos utilizando los servicios de secuenciación de la Universidad de Illinois en Urbana-Champaign fueron analizados con el método de análisis sugerido en el manual de RawGeno (Arrigo *et al.*, 2009) en el programa Peak Scanner (Applied Biosystems). El electroforegrama se exportó en forma de un tabla que contenía el color de la marca fluorescente, el número del pico, el nombre de la muestra, y el tamaño, altura y área (en pares de bases) del pico. Posteriormente se utilizó el programa RawGeno (Arrigo *et al.* 2009) para R (R core team, 2013) para definir los fragmentos confiables y repetibles para ser incluidos en la matriz final. El primer paso constó en importar el perfil genético de las muestras, esto se hizo de manera independiente para cada combinación de cebadores utilizados, utilizando la tabla obtenida en el análisis efectuado en Peak Scanner (Applied Biosystems). En seguida, RawGeno muestra una gráfica que indica la cantidad de fragmentos amplificados por individuo (Figura 2b). Para cada análisis se excluyó el 5% de los individuos con menor número de fragmentos amplificados indicativos de una reacción de PCR que no fue exitosa; esto ayuda a disminuir la cantidad de falsos negativos.

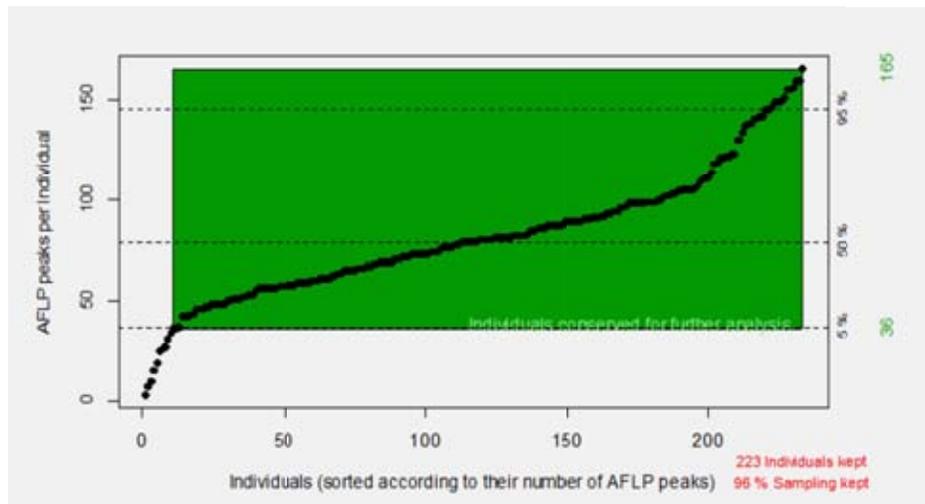


Figura 2b. Número de fragmentos amplificados por individuo utilizando la combinación 1a. El área verde comprende los individuos que fueron retenidos para continuar el análisis de esta combinación de cebadores, el área gris (esquina inferior izquierda) agrupa los individuos excluidos del análisis para esta combinación de cebadores en particular.

Posteriormente RawGeno analizó los perfiles de los individuos conservados y eliminó aquellos fragmentos que no cumplieron con las características enunciadas en la Tabla 2b; para ello utilizó el siguiente algoritmo:

A. Tamaño y altura de los fragmentos a partir del análisis hecho en Peak Scanner.

B. Al principio de la definición de las unidades de análisis (bin), todos los picos son considerados dentro de un sólo bin.

C. División del bin.
 1- Se calcula el tamaño de intervalo entre picos.
 2- Se divide el bin en el intervalo de mayor tamaño.

D. ¿Es el ancho del sub bin mayor al ancho máximo de bin?

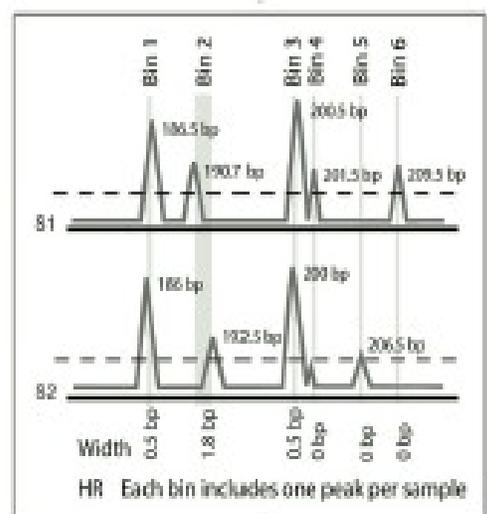
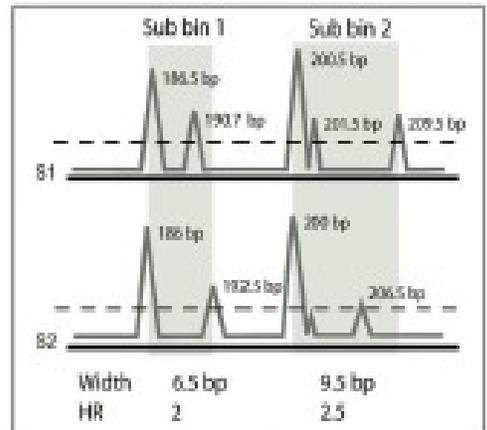
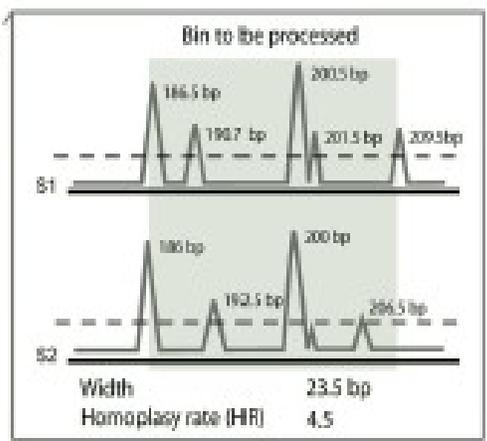
E. ¿Es el ancho del sub bin menor al ancho mínimo de bin?

F. ¿El sub bin incluye varios picos que pertenecen a la misma muestra?

G. Pasa al siguiente bin.
 Así procede hasta que cada sub bin queda conforme a los parámetros seleccionados.

H. Produce la matriz binaria.
 Registra la presencia o ausencia de cada pico en cada muestra y bin.

I. Revisión de calidad.
 Elimina los bins fuera del rango de búsqueda.
 Elimina los picos que no superan el umbral de intensidad.
 Elimina los picos no reproducibles.
 Elimina los picos en frecuencia menor a la deseada.



	Bin 1	Bin 2	Bin 3	Bin 4	Bin 5	Bin 6
S1	1	1	1	1	0	1
S2	1	1	1	0	1	0

Tabla 2b. Parámetros utilizados para el análisis de los perfiles genéticos, previamente analizados en el programa Peak Scanner y la definición de los fragmentos confiables. Estos parámetros fueron utilizados para las ocho combinaciones de cebadores.

Parámetro	Valor
Ancho máximo del bin	1.5 pb
Ancho mínimo del bin	1.0 pb
Intensidad mínima del pico	100 rfu
Frecuencia mínima del bin	11 individuos
Reproducibilidad	90%
Mantener bins sin reproducibilidad analizada	TRUE
Tamaño mínimo de los fragmentos analizados	100 pb
Tamaño máximo de los fragmentos analizados	500 pb

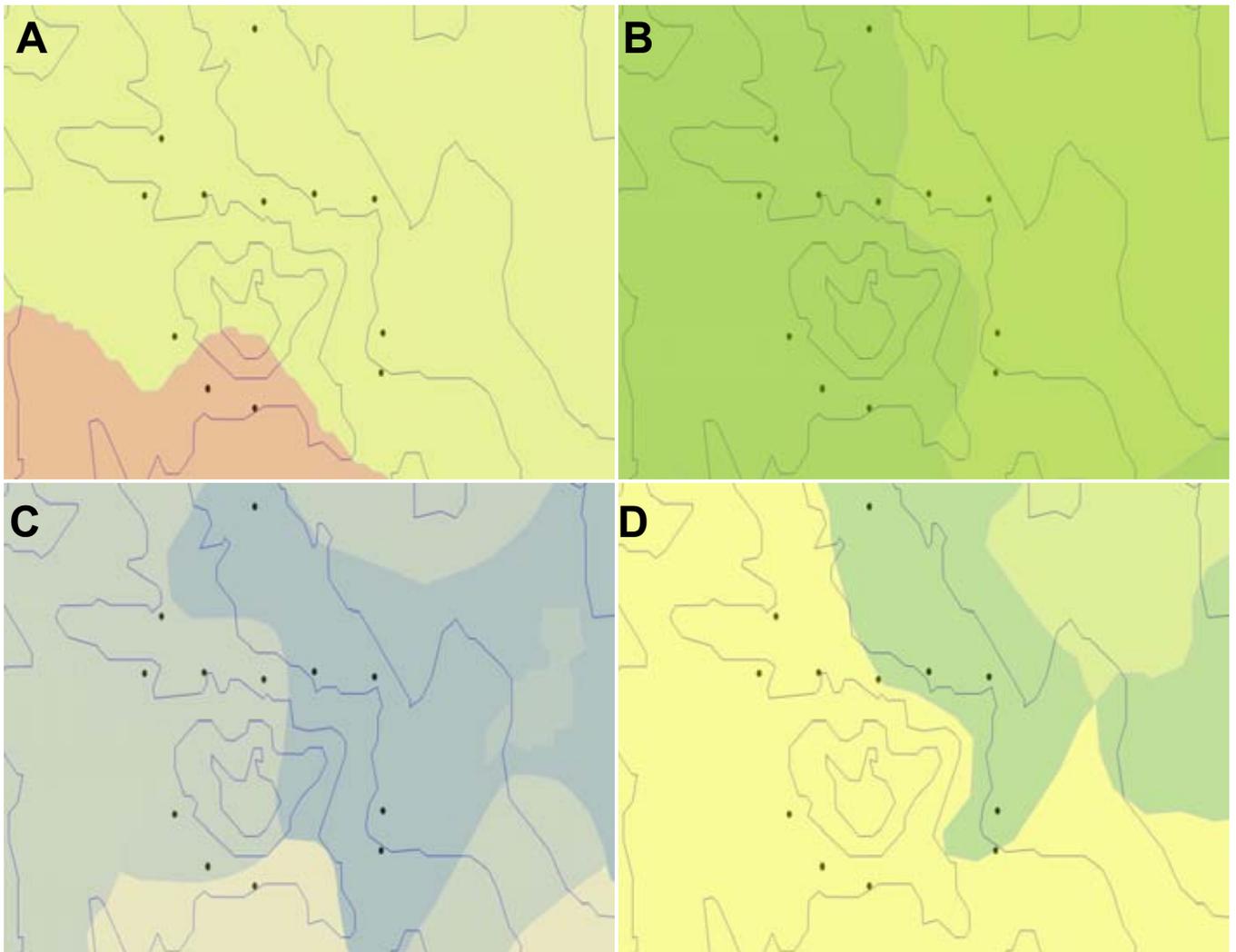
Abreviaciones: pb = pares de bases, rfu = unidades de fluorescencia relativa

El valor del ancho máximo y mínimo del *bin* se utilizó de acuerdo a lo sugerido en el manual de RawGeno (Arrigo *et al.*, 2009). La frecuencia mínima del *bin* correspondió al 5% del tamaño total de la muestra. Mantener los *bins* sin reproducibilidad analizada indica que aquellos bins que hayan presentado un estado de ausencia en todas las muestras con réplicas (es decir 0,0 en ambas muestras) serán conservados en la base de datos. Si bien esto genera cierta incertidumbre sobre la confiabilidad de dichos fragmentos, excluirllos de la base de datos podría provocar una subestimación de la diversidad genética. Se sabe que existe una correlación negativa entre el tamaño del fragmento y su probabilidad de presentar eventos de homoplasia por tamaño (es decir, fragmentos no homólogos del mismo tamaño que pertenecen al mismo individuo y que son considerados como un solo locus). Análisis teóricos señalan a los fragmentos de tamaño pequeño (<125 pb) como más propensos a la homoplasia por tamaño, ya que son más abundantes durante la obtención de AFLPs. En este estudio se utilizó un valor mínimo de 100 pb ya que esto permite incluir un mayor número de

fragmentos (lo que evita subestimar la diversidad genética). Además utilizando RawGeno se calculó la correlación entre la frecuencia de los fragmentos y el tamaño, solo en 3 de las 8 combinaciones de cebadores utilizadas se obtuvo una correlación significativa, aunque esto también está influenciado por el número de fragmentos obtenidos con cada combinación (véase resultados).

Anexo 3. Variables ambientales presentes en la zona de estudio

Anexo 3a. Distribución espacial de las variables ambientales utilizadas para construir la matriz de distancia ambiental entre localidades utilizada en el modelo causal y para los modelos jerárquicos de islas realizados en Arlequin v3.5.1.2 (Excoffier *et al.*, 2005) para detectar loci con señales de selección natural. (A) Regiones hidrológicas: en amarillo se muestra la cuenca del Pánuco y en rosa la del Balsas. (B) Tipo de rocas: en verde claro se ilustran rocas básicas y en verde oscuro las intermedias. (C) Tipo de vegetación y uso de suelo: en azul se indica el bosque de oyamel, en gris la vegetación inducida y en beige el bosque de pino. (D) Tipo de suelo: en verde se muestra el Inceptisol y en amarillo el Entisol. Los puntos negros representan las doce localidades muestreadas, las líneas negras son las curvas de nivel.



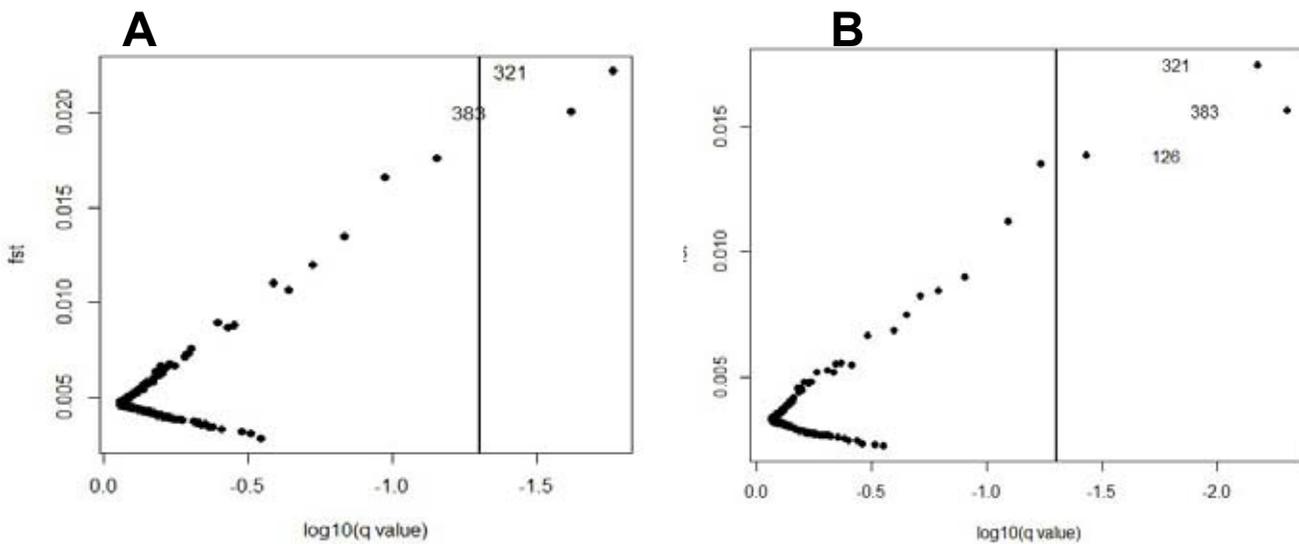
Anexo 4. Análisis de reproducibilidad para construir la matriz binaria final

Anexo 4a. Resultados del análisis de reproducibilidad para definir la matriz binaria final, una vez que se eliminaron aquellos individuos con reacciones de PCR negativas. Se utilizaron 213 individuos de *A. religiosa*, genotipados para ocho combinaciones de cebadores. # de fragmentos iniciales se refiere al número total de fragmentos reportados en el electroforegrama a partir del análisis realizado en Peak Scanner. Removidos por baja intensidad se refiere al número de fragmentos que no sobrepasaron el umbral de intensidad (100 rfu). Removidos por baja reproducibilidad se refiere al número de fragmentos que presentaron una reproducibilidad por debajo del 87.5%. Removidos por baja frecuencia se refiere al número de fragmentos con una frecuencia menor al 5% dentro de la población. El número de fragmentos finales se refiere a aquellos fragmentos que cumplieron con los requisitos establecidos y fueron utilizados para construir la matriz binaria final. Significancia de la correlación entre el tamaño del fragmento y su frecuencia, en negritas y cursiva se resaltan las correlaciones estadísticamente significativas.

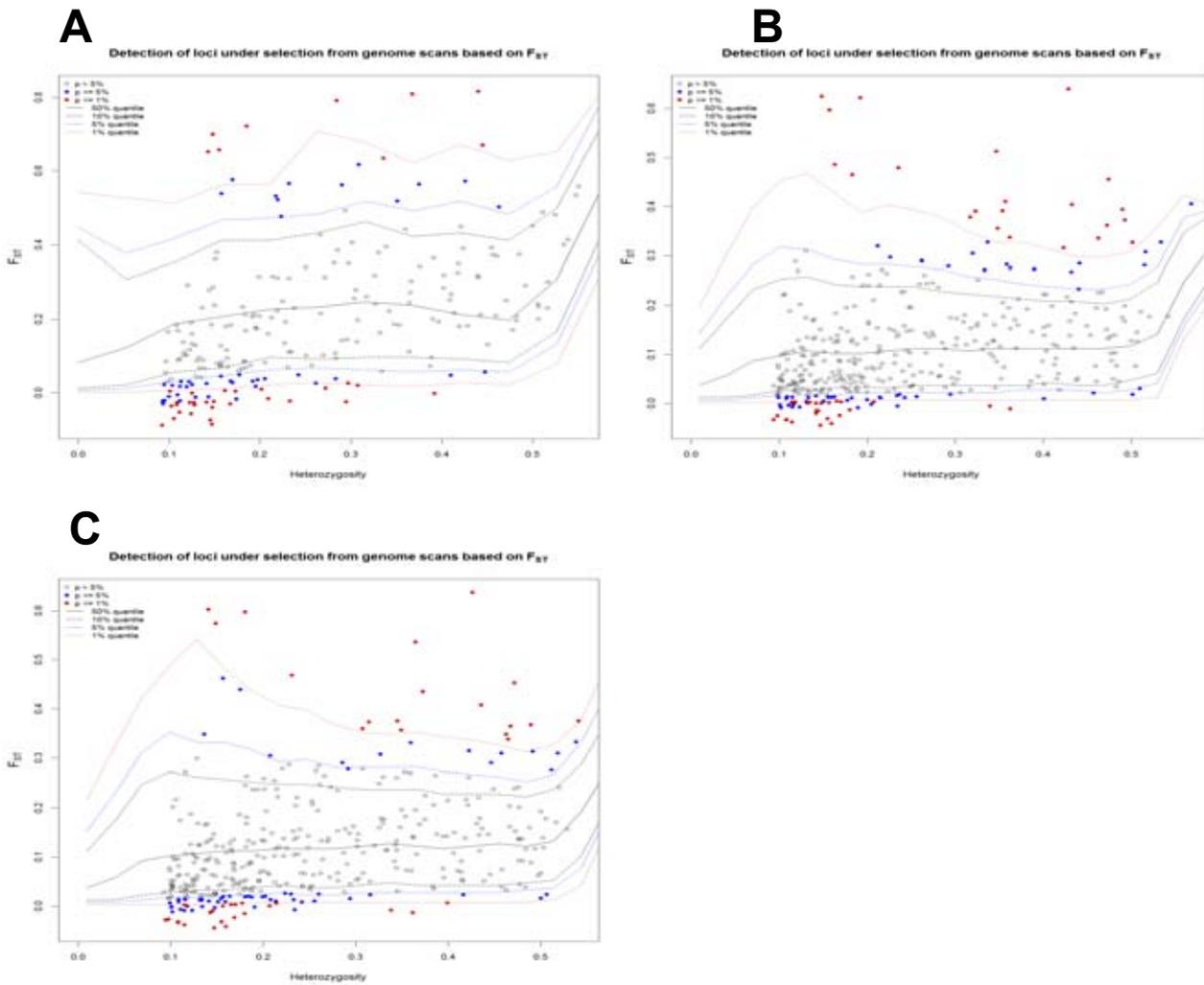
	1a	2c	3a	3c	3f	3g	3i	3j	Total	Porcentaje
Num. Bin iniciales	283	373	324	470	320	515	443	474	3202	100.00
Removidos por baja intensidad	129	192	243	191	180	297	275	191	1698	53.03
Removidos por baja reproducibilidad	81	67	39	78	66	85	61	85	562	17.55
Removidos por baja frecuencia	45	67	32	107	48	62	64	123	548	17.11
Num. Bin finales	28	47	10	94	26	71	43	75	394	12.30
Relación entre el tamaño y la frecuencia. Valor de <i>p</i>	0.01	0.895	0.221	0.404	0.034	0.002	0.364	0.257	-	-

Anexo 5. Detección de loci bajo selección natural.

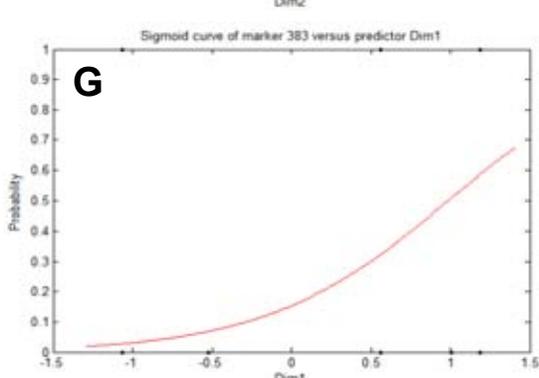
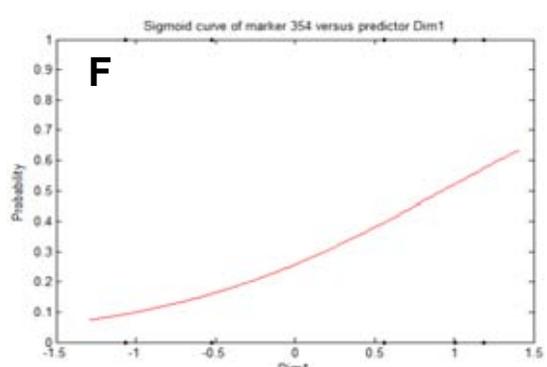
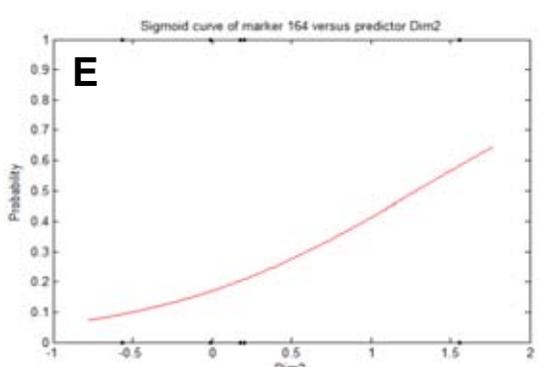
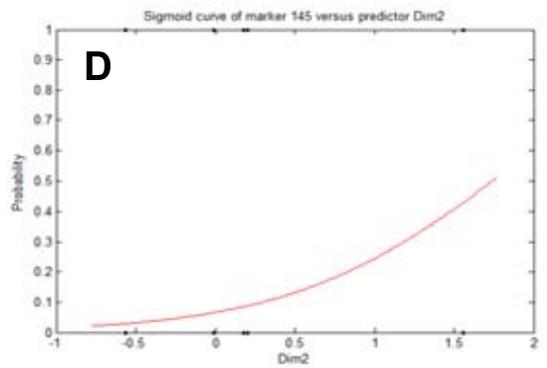
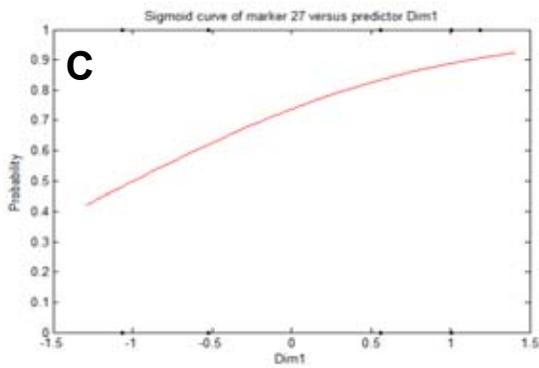
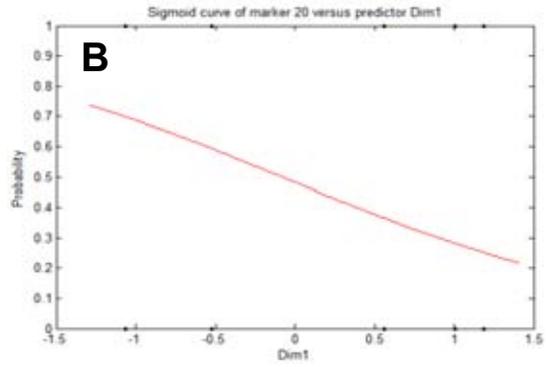
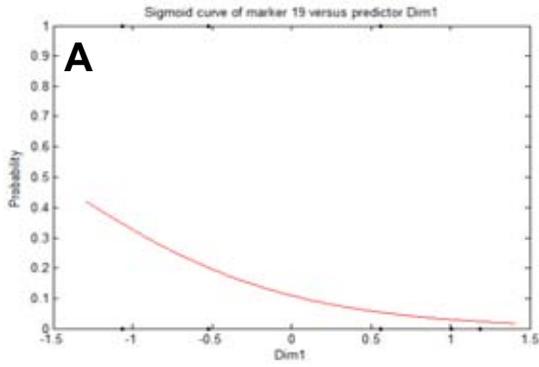
Anexo 5a. Resultados de los dos modelos puestos a prueba para detectar loci bajo selección natural en el programa Bayescan v2.1 (Foll & Gaggiotti, 2008). (A) Modelo 1 supone valores para el coeficiente de endogamia F_{is} entre 0.0 y 1.0; con él se detectaron dos loci. (B) Modelo 2 supone que F_{is} toma valores de un intervalo de 0.192 a 0.216, estos valores están reportados en la literatura y fueron calculados empíricamente para *A. religiosa*, utilizando microsatélites nucleares (Múgica-Gallart, 2013) e isoenzimas (Aguirre-Planter *et al.*, 2000) respectivamente; con este modelo se detectaron tres loci.



Anexo 5b. Detección de loci bajo selección natural en el programa Arlequin v3.5.1.2 (Excoffier *et al.*, 2005). (A) MJJ-G, basado en la estructura genética poblacional obtenida con el análisis realizado en Geneland v4.0.3 (Guillot *et al.*, 2005), nueve loci fueron detectados. (B) MJJ-S, basado en la estructura genética poblacional arbitraria utilizando la variable ambiental “Tipo de Suelo”, 22 loci fueron detectados (C) MJJ-V basado en la estructura genética poblacional arbitraria utilizando la variable ambiental “Vegetación”, 18 loci fueron detectados. Para los tres modelos se muestran los valores de F_{ST} de los 394 AFLPs genotipados, en gris se indican aquellos loci con una probabilidad nula de ser selectivos ($p > 0.05$), en azul aquellos con una probabilidad alta de ser selectivos ($p \leq 0.05$) y en rojo aquellos loci con una probabilidad muy alta de ser selectivos ($p < 0.01$). Solo se tomaron en cuenta los loci con señales de selección direccional; es decir, aquellos con valores inusualmente altos para F_{ST} , ubicados por encima de la distribución neutral.

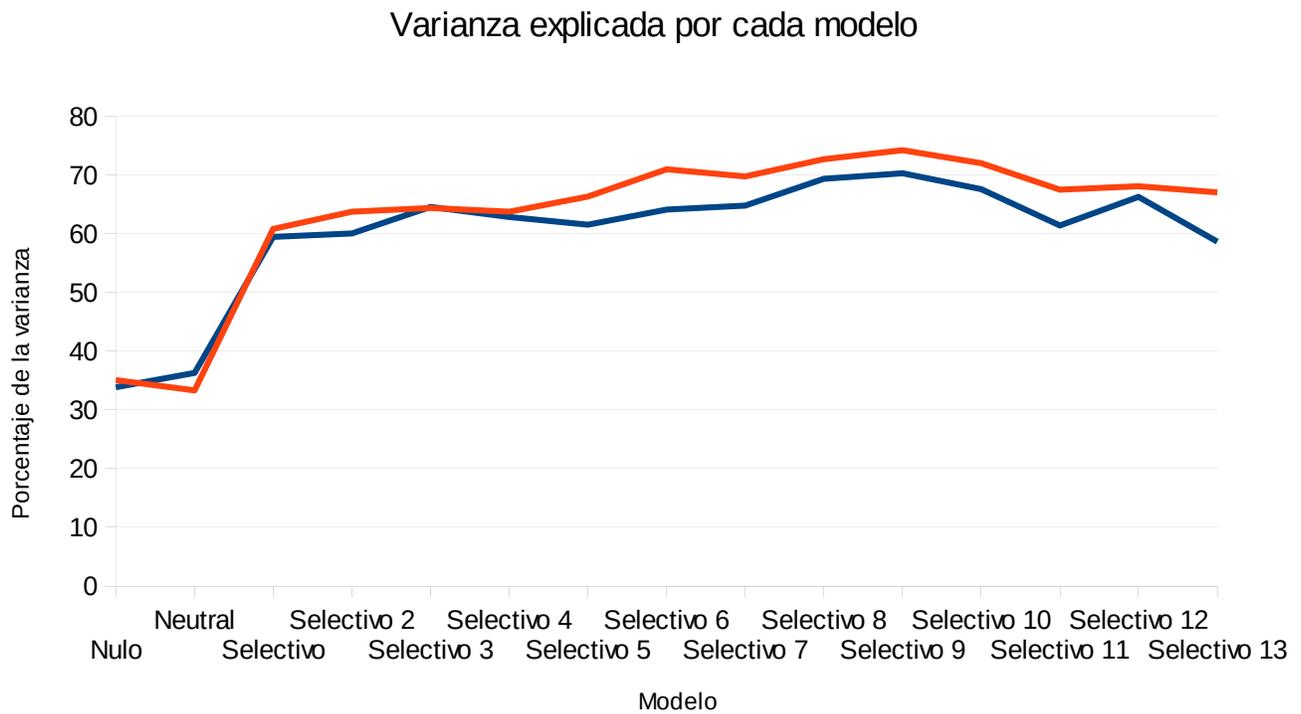


Anexo 5c. Curva sigmoide de cada uno de los siete loci detectados como selectivos por el programa matsam v2.0 (Joost *et al.*, 2007) contra la variable ambiental predictiva, dimensión 1 (dim1) o dimensión 2 (dim2). (A) Locus 19 contra dim1. (B) Locus 20 contra dim1. (C) Locus 27 contra dim1. (D) Locus 145 contra dim2. (E) Locus 164 contra dim2. (F) Locus 354 contra dim1. (G) Locus 383 contra dim1.

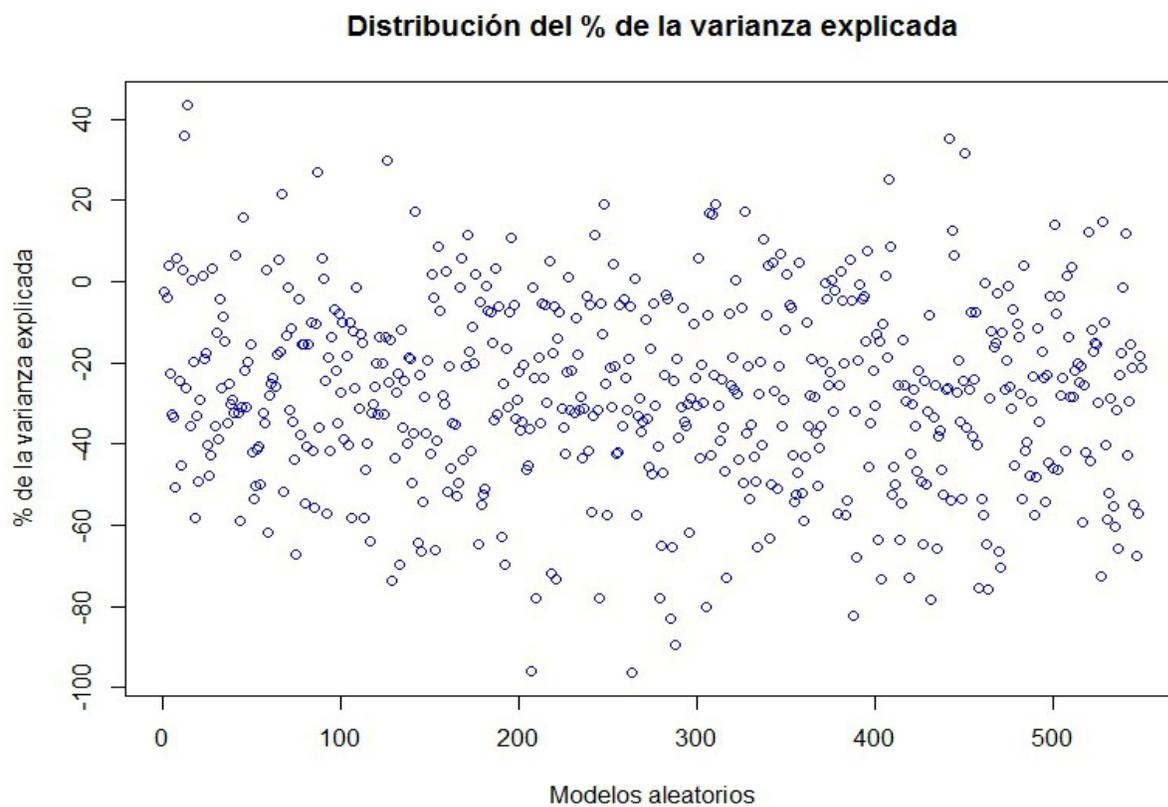


Anexo 6. Procesos selectivos y diferenciación genética-migración

Anexo 6a. Porcentaje de la varianza explicada por cada modelo construido utilizando el algoritmo *Random Forest* (Breiman, 2001). La línea azul representa los modelos que utilizaron la coordenada 1 (resultado del PcoA) como variable de respuesta. Esta explica el 27.64% de la diferenciación genética total. La línea naranja representa los modelos que utilizaron la coordenada 2 (resultado del PCoA) como variable de respuesta y que explicó el 20.85% de la diferenciación genética total.



Anexo 6b. Distribución del porcentaje de la varianza explicada por cada uno de los 550 modelos aleatorios construidos utilizando el algoritmo *Random Forest* (Breiman, 2001).



9. Referencias

- Aguirre-Planter E, Furnier GR & Eguiarte LE. (2000). Low levels of genetic variation within and high levels of genetic differentiation among populations of species of *Abies* from southern Mexico and Guatemala. *American Journal of Botany*, 87(3), 362-371.
- Aguirre-Planter E, Jaramillo-Correa JP, Gómez-Acevedo S, Khasa DP, Bousquet J & Eguiarte LE. (2012). Phylogeny, diversification rates and species boundaries of Mesoamerican firs (*Abies*, Pinaceae) in a genus-wide context. *Molecular Phylogenetics and Evolution*, 62(1), 263-274.
- Alberto F, Aitken SN, Alía R, González-Martínez SC, Hänninen H, Kremer A, Lefevre F, Lenormand T, Yeaman S, Whetten R (2013). Potential forevolutionary responses to climate change – evidence from tree populations. *Global Change Biology*, 19, 1645–1661.
- Anderson CD, Epperson BK, Fortin M, Holderegger R, James PM, Scribner, KT, Spear S (2010). Considering spatial and temporal scale in landscape-genetic studies of gene flow. *Molecular ecology* 19, 3565–75.
- Anderson JB, Browler LP (1996). Freeze protection of overwintering monarch butterflies in México. Critical role of the forest as a blanket and an umbrella. *Ecological Entomology* 21, 107 – 116.
- Andrew RL, Ostevik KL, Ebert DP & Rieseberg LH (2012). Adaptation with gene flow across the landscape in a dune sunflower. *Molecular ecology* 21, 2078–91.
- Arrigo N, Tuszynski JW, Ehrich D, Gerdes T, Alvarez N (2009). Evaluating the impact of scoring parameters on the structure of intra-specific genetic variation using RawGeno, an R package for automating AFLP scoring. *BMC Bioinformatics* 10:33.
- Arrigo N, Tuszynski JW, Ehrich D, Gerdes T & Alvarez N (2009). Evaluating the impact of scoring parameters on the structure of intra-specific genetic variation using RawGeno, an R package for automating AFLP scoring. *BMC Bioinformatics* 10, 33.
- Barret RDH & Schluter D (2007). Adaptation from standing genetic variation. *Trends in ecology and evolution*, 38 – 44.
- Beaumont MA, Nichols RA (1996). Evaluating loci for use in the genetic analysis of population structure. *Proceedings: Biological Sciences* 263, 1619–1626.

Beaumont M A, Zhang W & Balding D J (2002). Approximate Bayesian computation in population genetics. *Genetics* 162, 2025 – 2035.

Bertorelle G, Benazzo A & Mona S (2010). ABC as a flexible framework to estimate demography over space and time: some cons, many pros. *Molecular Ecology* 19, 2609 – 2625.

Blum BM, Klaider H M & Randal AG (1981). Northeastern spruce-fir. Choices in silviculture for American forests. *Society of American Foresters*, Washington, DC.

Bonin A, Ehrich D & Manel S (2007) Statistical analysis of amplified fragment length polymorphism data: a toolbox for molecular ecologists and evolutionists. *Molecular Ecology* 16, 3737 – 3758.

Bonin A, Taberlet P, Miaud C & Pompanon F (2006). Explorative Genome Scan to Detect Candidate Loci for Adaptation Along a Gradient of Altitude in the Common Frog (*Rana temporaria*). *Molecular Biology and Evolution*, 23, 773 – 778.

Breiman L (2001). Random Forests. *Machine Learning* 45, 5 – 32.

Caballero A, Quesada H & Rolan-Alvarez E (2008). Impact of Amplified Fragment Length Polymorphism Size Homoplasmy on the Estimation of Population Genetic Diversity and the Detection of Selective Loci. *Genetics* 179:539 – 554.

Cavalli-Sforza L L, Edwards A (1967). Phylogenetic analysis: models and estimation procedures. *Evolution* 21, 550 – 570.

Chagne D, Lalanne C, Madur D, Kumar S, Frigerio JM, Krier C, Decroocq S, Savoure A, Bou-Dagher-Kharrat M, Bertocchi E, Brach J & Plomion C (2002). A high genetic map of maritime pine based on AFLPs. *Ann. For. Sci*, 59, 627 – 636.

Cinget B (2014). Étude phylogéographique pancanadienne du sapin baumier (*Abies balsamea* [L.] Mill.) et de ses relations avec le sapin subalpin (*Abies lasiocarpa* [Hook] Nutt.) dans l'ouest du Canada. Tesis de Doctorado. Universidad de Laval, Canadá.

Corander J, Siren J & Arjas E (2008). Bayesian spatial modeling of genetic population structure. *Computational Statistics*, 23, 111-129.

Cuenca A, Escalante A E., and Piñero D (2003). Long-distance colonization, isolation by distance, and historical demography in a relictual Mexican pinyon pine (*Pinus nelsonii* Shaw) as re- vealed by paternally inherited

genetic markers (cpSSRs). *Molecular Ecology*, 12, 2087 – 2097.

Cushman S A & Landguth E L (2010). Scale dependent inference in landscape genetics. *Landscape Ecology* 25, 967 – 979.

Cutler R D, Edwards T C, Beard K H, Cutler A, Hess K T, Gibson J & Lawler J J (2007). Random Forests for classification in Ecology. *Ecology* 88, 2783 – 2792.

Dabney A, Storey J D and with assistance from Gregory R. Warnes (). qvalue: Q-value estimation for false discovery rate control. R package version 1.36.0.

Dagher-Kharrat M B, Mariette S, Lefèvre F, Fady B, Frenier-de March G, Plomion C & Sauré A (2006). Geographical diversity and genetic relationships among *Cedrus* species estimated by AFLP. *Tree Genetics & Genomes* 3, 275 – 285.

De Mita S, Thuillet A C, Gay L, Ahmadi N, Manel S, Ronfort J & Vigoroux Y (2013). Detecting selection along environmental gradients: analysis of eight methods and their effectiveness for outbreeding and selfing populations. *Molecular Ecology*, 22, 1383-1399.

Delgado H, Molinero R, Cervantes P, Nieto-Obregón J, Lozano-Santa Cruz R, Macías-González H L, Mendoza-Rosales & Silva-Romo G (1998). Geology of Xitle volcano in Southern Mexico City – a 2000 year old monogenetic volcano in an urban area. *Revista Mexicana de Ciencias Geológicas*, 15, 115-131.

Delgado P, Salas-Lizana R, Vázquez-Lobo A, Alvarez-Buylla E R, Vendramin G G, and Piñero D (2007). Introgressive hybridization in *Pinus montezumae* Lamb. and *P. pseudostrobus* Lindl. (Pinaceae): morphological and molecular (cpSSR) evidence. *International Journal of Plant Sciences*, 168, 861 – 875.

Diaz V, Muniz L M & Ferrer E (2001). Random amplified polymorphic DNA and amplified fragment length polymorphism assessment of genetic variation in Nicaraguan populations of *Pinus oocarpa*. *Molecular Ecology* 10, 2593 – 2603.

Dubreuil M, Riba M, González-Martínez S C, Vendramin G G, Sebastiani F & Mayol M (2010). Genetic effects of chronic habitat fragmentation revisited: Strong genetic structure in a temperate tree, *Taxus Baccata* (Taxaceae), with great dispersal capability. *American Journal of Botany* 97, 303 – 310.

Echt C S, DeVerno L L, Anziedi M & Vendramin G G (1998). Chloroplast microsatellites reveal population genetic diversity in red pine, *Pinus resinosa* Ait. *Molecular Ecology*, 7, 307 – 316.

Eckenwalder J E (2009). Conifers of the World. The Complete Reference. Ed. Timber Press, Inc. Portland, Oregon.

Eckert A J, Bower A D, González-Martínez S C, Wegrzyn J L, Coop G & Neale D B (2010). Back to nature: ecological genomics of loblolly pine (*Pinus taeda*, Pinaceae). *Molecular Ecology*, 19, 3789-3805.

Edwards D G W (1981). Improving seed germination in *Abies*. *Proceedings of the International Plant Propagation Society* 31, 69 – 78.

Excoffier L, Laval G & Schneider S (2005). Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* 1:47- 50.

Excoffier L & Lischer H (2011). *Manual de usuario Arlequin v. 3.5*.

Farjon A (1990). Pinaceae: drawings and descriptions of the genera *Abies*, *Cedrus*, *Pseudolarix*, *Keteleeria*, *Nothotsuga*, *Tsuga*, *Cathaya*, *Pseudotsuga*, *Larix* and *Picea*. Koeltz Scientific Books, Alemanis, 330 pp.

Foulkes A S (2009). *Applied Statistical Genetics with R*. Springer Science+Business Media. Amherst, MA. USA.

Fournier-Level A, Korte A, Cooper M D, Nordborg M, Schmitt J & Wilczek A M (2011). A map of local adaptation in *Arabidopsis Thaliana*. *Science*, 334, 86 – 89.

Frichot E, Schoville S D, Bouchard G & Francois O (2013). Testing for Associations between Loci and Environmental Gradients Using Latent Factor Mixed Models. *Molecular Biology Evolution*, 30, 1687-1699.

Galarza J A, Carreras-Carbonell J, Macpherson E, Pascual M, Roques S, Turner G F & Rico C (2009). The influence of oceanographic fronts and early-life-history traits on connectivity among littoral fish species. *Proceedings of the National Academy of Sciences*, 106, 1473-1478.

García-Aguirre M C, Ortiz M A, Zamorano J J & Reyes Y (2007). Vegetation and landform relationships at Ajusco volcano Mexico, using a geographic information system (GIS). *Forest Ecology and Management*, 239, 1-12.

Garcia L V (2004). Escaping the Bonferroni iron claw in ecological studies. *Oikos* 105, 657–663.

Garza J C & Williamson E G (2001). Detection of reduction in population size using data from microsatellite loci. *Mol Ecol* 10, 305–318.

Gaudeul M, Till-Boutraud I, Barjon F & Manel S (2004). Genetic diversity and differentiation in *Eryngium alpinum*

L. (Apiaceae): comparison of AFLP and microsatellite markers. *Heredity* 92, 508-518.

Gort G, Van Hintum T, Van Eeuwijk F (2009). Homoplasy corrected estimation of genetic similarity from AFLP bands, and the effect of the number of bands on the precision of estimation. *Theoretical Applied Genetics*, 119, 397-416.

González-Martínez S C, Gerber S, Cervera M T, Martínez-Zapater J M, Gil L, Alía R (2002). Seed gene flow and fine-scale structure in a Mediterranean pine (*Pinus pinaster* Ait.) using nuclear microsatellite markers. *Theoretical and Applied Genetics*, 104, 1290-1297.

González-Martínez S C, Dubreuil M, Riba M, Vendramin G G, Sebastiani F & Mayol M (2010). Spatial genetic structure of *Taxus baccata* L. in the western Mediterranean Basin: Past and present limits to gene movement over a broad geographic scale. *Molecular Phylogenetics and Evolution* 55, 805 – 815.

Graham A. (1993). Historical factors and biodiversity in Mexico. In T . P . Ramammoorthy, R. Bye, A. Lot, and J. Fa [eds.], *Biological diversity of Mexico: origins and distribution*, 10

Guillot G, Mortier F & Estoup A (2005). Geneland: A computer package for landscape genetics. *Molecular Ecology Notes*, 5,708–711.

Guillot G & Santos F (2010). Using AFLP markers and the Geneland program for the inference of population genetic structure. *Molecular Ecology* 10, 1082 – 1084.

Hamrick J L (2004). Response of forest trees to global environmental changes. *Forest Ecology and Management* 197, 323-335.

Hancock A M, Brachi B, Faure N, Horton M W, Jarymowycz L B, Sperone F G, Toomajian C, Roux F & Bergelson J (2011) Adaptation to Climate Across the *Arabidopsis thaliana* Genome. *Science*, 334, 83 – 86.

Hasbún R, Iturra C, Moraga P, Wachtendorff P, Quiroga P & Valenzuela S (2012). An efficient and reproducible protocol for production of AFLP markers in tree genomes using fluorescent capillary detection. *Tree Genetics & Genomes* 8(4), 925-931.

Hedrick P W (2011). *Genetics of Populations*. Jones & Barttlet Learning.

Heuertz M, Vekemans X, Hausman F, Paladas M & Hardy O J (2003). Estimating seed vs pollen dispersal from spatial genetic structure in the common ash. *Molecular Ecology* 12, 2483 – 2495.

- Hilfiker K, Gugerli F, Schütz J P, Rotach P & Holderegger R (2004). Low RAPD variation and female-biased sex ratio indicate genetic drift in small populations of the dioecious conifer *Taxus baccata* in Switzerland. *Conservation Genetics* 5, 357 – 365.
- Ho S Y & Shapiro B (2011). Skyline-plot methods for estimating demographic history from nucleotide sequences. *Molecular Ecology* 11, 423 – 434.
- Holsinger K E & Weir B S (2009). Genetics in geographically structured populations: defining, estimating and interpreting FST. *Genetics*, 639-650.
- Hosmer D W, Lemeshow S (2000). *Applied Logistic Regression*. John Wiley, New York.
- Husband B C & Barrett S C (1996). A Metapopulation Perspective in Plant Biology. *Journal of Ecology* 84, 461 – 469.
- Hughes R, Inouye BD, Johnson MT, Underwood N & Vellend, M (2008). Ecological consequences of genetic diversity. *Ecology letters* 11, 609–23.
- Husson F, Josse J, Le S & Mazet J (2013). FactoMineR: Multivariate Exploratory Data Analysis and Data Mining with R. R package version 1.25. <http://CRAN.R-project.org/package=FactoMineR>
- Jaramillo-Correa JP, Aguirre-Planter E, Khasa DP, Eguiarte LE, Piñero D, Furnier GR, & Bousquet J (2008). Ancestry and divergence of subtropical montane forest isolates: molecular biogeography of the genus *Abies* (Pinaceae) in southern México and Guatemala. *Molecular ecology*, 17(10), 2476-2490.
- Jones F C, *et al.* (2012). The genomic basis of adaptive evolution in threespine sticklebacks. *Nature*, 484, 55 – 61.
- Joost S, Bonin A, Bruford M, Despres L, Conord C, Erhardt G & Taberlet P (2007). A spatial analysis method (SAM) to detect candidate loci for selection: towards a landscape genomics approach to adaptation. *Molecular Ecology* 16, 3955 – 3969.
- Jordano P (2010). Pollen, seeds and genes: the movement ecology of plants. *Heredity*, 105, 329 – 330.
- Juárez-Guerrero L (2010). Estructura del Bosque de *Abies religiosa* (H.B.K.) Schl. et Cham. en la cuenca de Zempoala, Morelos. Tesis de Biología. Facultad de Ciencias. UNAM.
- Knowles L L & Alvarado-Serrano D F (2010). Exploring the population genetic consequences of the colonization

process with spatio-temporally explicit models: insights from coupled ecological, demographic and genetic models in montane grasshoppers. *Molecular Ecology* 19, 3727 – 3745.

Kolotelo D (1997). Cone and Seed Improvement Program BCMoF Tree Seed Centre. Forest Nursery Association of British Columbia Proceedings.

Latch E K, Dharmarajan G, Glaubitz J C & Rhodes O E (2006). Relative performance of Bayesian clustering software for inferring population substructure and individual assignment at low levels of population differentiation. *Conservation Genetics* 7, 295 – 302.

Lewontin R C, Krakauer J (1973). Distribution of gene frequency as a test of the theory of selective neutrality of polymorphisms. *Genetics* 74, 175–195.

Liaw A & Wiener M (2002). Classification and Regression by randomForest. *R News* 2, 18 — 22.

Liu T S (1971). A monograph of the genus *Abies*. Department of Forestry, College of Agriculture. National Taiwan University, China, 608.

Lotterhos K E & Withlock M C (2014) Evaluation of demographic history and neutral parameterization on the performance of FST outliers tests. *Molecular Ecology* 23, 2178 – 2192.

Loveless M D & Hamrick J L (1984). Ecological determinants of genetic structure in plant populations. *Annual Review of Ecology and Systematics* 15, 65 – 95.

Manel S, Schwartz MK, Luikart G & Taberlet P (2003). Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology and Evolution*, 18, 189–197.

Manel S, Conord C & Després L (2009). Genome scan to assess the respective role of host-plant and environmental constraints on the adaptation of a widespread insect. *BMC Evolutionary Biology* 9, 288.

Mariette S, Chagne D, Lezier C, Pastuszka P, Raffin A, Plomion C & Kremer A (2001). Genetic diversity within and among *Pinus pinaster* populations: comparison between AFLP and microsatellite markers. *Heredity*, 86, 469-479.

Marques CM, Araujo JA, Ferreira JG, Whetten R, O'Malley DM, Liu BH & Sederoff R (1998). AFLP genetic maps of *Eucalyptus globulus* and *E. tereticornis*. *Theoretical Applied Genetics*, 96, 727 – 737.

Martínez M (1948). Los *Abies* mexicanos. *Anales del Instituto de Biología de México* 19, 11-104.

May E N G (2001). Dinámica de la regeneración de *Abies religiosa* (H.B.K.) Schl. Et Cham., y *Pinus Hartwegii* Lindl., en la estación forestal experimental Zoquiapan, estado de México. Tesis de licenciatura ENA Chapingo México pp 8-25.

McRae BH (2006). Isolation by resistance. *Evolution* 60, 1551–1561.

McRae BH. & Beier P (2007). Circuit theory predicts gene flow in plant and animal populations. *Proceedings of the National Academy of Sciences of the United States of America* 104, 19885–90.

McRae BH & Shah VB (2009). Circuitscape User Guide. ONLINE. The University of California, Santa Barbara. Available at: <http://www.circuitscape.org>.

Meudt H M & Clarke A C (2006). Almost Forgotten or Latest Practice? AFLP applications, analyses and advances. *Trends in Plant Science*, 12, 106 – 117.

Mimura M, & Aitken S (2007). Adaptive gradients and isolation- by-distance with postglacial migration in *Picea sitchensis*. *Hereditas*, 99: 224–23

Mosca E, Eckert A J, Di Pierro E A, Rocchini D, La Porta N, Belletti P & Neale D (2012). The geographical and environmental determinants of genetic diversity for four alpine conifers of the European Alps. *Molecular Ecology*, 21, 5530-5545.

Mosca E, González-Martínez S C, Neale D B (2014). Environmental versus geographical determinants of genetic structure in two subalpine conifers. *New Phytologist*, 201, 180-192.

Múgica-Gallart J. (2013). Patrones de diversidad y detección de grupos genéticos en dos especies de abies con distribuciones contrastantes en el Eje volcánico Transversal. Tesis de Biología. Facultad de Ciencias. UNAM.

Munshi-South J (2012). Urban landscape genetics: canopy cover predicts gene flow between white-footed mouse (*Peromyscus leucopus*) populations in New York City. *Molecular ecology* 21, 1360–78.

Murphy M A, Evans J S, Cushman S A & Storfer A (2008). Representing genetic variation as continuous surface: an approach for identifying spatial dependency in landscape genetics. *Ecography* 31, 685 – 697.

Narum S R (2006). Beyond Bonferroni: Less conservative analyses for conservation genetics. *Conservation Genetics* 7, 783 – 787.

Nei M, & Li W H (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences* 76, 5269 – 5273.

Nosil P, Funk DJ, Ortiz-Barrientos D (2009). Divergent selection and heterogeneous genomic divergence. *Molecular Ecology*, 18:375–402.

Oksanen J, Guillaume F, Kindt R, Legendre P, Minchin P, O'Hara R, Simpson G, Solymos P, Stevens M & Wagner H (2013). *vegan: Community Ecology Package*. R package version 2.0-10. <http://CRAN.R-project.org/package=vegan>

Paglia GP, Olivieri AM & Morgante M (1998). Towards second-generation STS (sequence-tagged sites) linkage maps in conifers: a genetic map of Norway spruce (*Picea abies* K.) *Mol. Gen. Genet.*, 258, 466 – 478.

Pannell JR & Fields PD (2014). Evolution in subdivided plant populations: concepts, recent advances and future directions. *New Phytologist*, 201: 417–432.

Parker A J & Peet R K (1984). Size and age structure of conifer forests. *Ecology*, 65, 1685 – 1689.

Penrod K, Cabanero C, Beier P, Luke C, Spencer W, Rubin E, Sauvajot R, Riley S & Kamrat D (2006). South coast missing linkages project: a linkage design for the Santa Monica-Sierra Madre Connection. South Coast Wildlands, Idyllwild, USA

Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey S & Rafalski A (1996). The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Mol Breed* 2:225–238.

Pybus O G, Rambaut A & Harvey P H (2000). An integrated framework for the inference of viral population history from reconstructed genealogies. *Genetics* 155, 1429 – 1437.

R Core Team (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.

Remington DL, Whetten RW, Liu BH & O'Malley DM (1999). Construction of an AFLP genetic map with nearly complete genome coverage in *Pinus taeda*. *Theor. Appl. Genet.*, 98, 1279 – 1292.

Richardson B, Rehfeldt G & Kim M (2009). Congruent climate-related genealogical responses from molecular markers and quantitative traits for western white pine (*Pinus monticola*). *International Journal of Plant Sciences* 170, 1120 – 1131.

- Ross-Ibarra J, Tenaillon M & Gaut B S (2009). Historical divergence and gene flow in the genus *Zea*. *Genetics* 181, 1399 – 1413.
- Rushforth, K D (1989). Two new species of *Abies* (Pinaceae) from western México. *Notes from the Royal Botanic Garden Edinburgh*, 46.
- Rzedowski, J (2006). *Vegetación de México*. 1ra. Edición digital, Comisión Nacional para el Conocimiento y Uso de la Biodiversidad. México.
- Sagnar F, Pichot C, Dreyfus P, Jordano P & Fady B (2006). Modelling seed dispersal to predict seedling recruitment: Recolonization dynamics in a plantation forest. *Ecological Modelling* 203, 464 – 474.
- Safner T, Miller M P, McRae B H, Fortin M J & Manel S (2011). Comparison of Bayesian Clustering and Edge Detections Methods for Inferring Boundaries in Landscape Genetics. *International Journal of Molecular Sciences*, 12, 865-889.
- Schlaepfer C P (1968). *Resumen de la geología de la hoja México, Distrito Federal y estados de México y Morelos: México, DF, Universidad Nacional Autónoma de México, Instituto de Geología, Carta Geológica de México, Serie 1 : 100,000.*
- Schönswetter P & Tribsch A (2005). Vicariance and Dispersal in the Alpine Perennial *Bupleurum stellatum* L. (Apiaceae). *Taxon* 54, 725 – 732.
- Schlüter P M & Harris S (2006). Analysis of multilocus fingerprinting data sets containing missing data. *Molecular Ecology Notes* 6, 569 – 572.
- Selkoe K A & Toonen R J (2006). Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters*, 9, 615-629.
- Senneville S, Beaulieu J, Daoust G, Deslauriers M & Bousquet J (2001) Evidence for low genetic diversity and metapopulation structure in Canada yew (*Taxus canadensis*): considerations for conservation. *Canadian Journal of Forest Research* 31, 110 – 116.
- Shaffer J P (1995). Multiple hypothesis-testing. *Annual Review of Psychology* 46, 561–584.
- Slatkin M, Voelm L (1991). FST in a hierarchical island model. *Genetics* 127, 627.- 629.

- Slatkin M (1993). Isolation by distance in equilibrium and non-equilibrium populations. *Evolution*, 47:264 – 279.
- Slatkin M (1995). A measure of population subdivision based on microsatellite allele frequencies. *Genetics* 139, 457 – 462.
- Slatkin M (2000). A Coalescent view of Population Structure. En: *Evolutionary Genetics*. Editado por Singh R S & Krimbas C B. Cambridge University Press. Pp, 418 – 429.
- Spear S F, Balkenhol N, Fortin M J, McRae B H & Scribner K (2010). Use of resistance surfaces for landscape genetic studies: considerations for parameterization and analysis. *Molecular Ecology*, 19, 3576 – 3591.
- Spurgin L G, Illera J C, Jorgensen T H, Dawson D A & Richardson D S (2014). Genetic and phenotypic divergence in an island bird: isolation by distance, by colonization or by adaptation. *Molecular Ecology* 23, 1028 – 1039.
- Storey J D & Tibshirani R (2003) Statistical significance for genomewide studies. *Proceedings of the National Academy of Sciences* 100, 9440 – 9445.
- Tajima F (1983). Evolutionary relationship of DNA sequences in finite populations. *Genetics* 105, 437 – 460.
- Tang S, Dai W, Li M, Zhang Y, Geng Y, Wang L & Zhong Y (2008). Genetic diversity of relictual and endangered plant *Abies ziyuanensis* (Pinaceae) revealed by AFLP and SSR markers. *Genetica* 133, 21- 30.
- Tsumura Y, Uchiyama K, Moriguchi Y, Ueno S & Ihara-Ujino T (2012). Genome scanning for detecting adaptive genes along environmental gradients in the Japanese conifer, *Cryptomeria japonica*. *Heredity*, 109, 349 – 360.
- Vendramin GG, Lelli L, Rossi P, Morgante M (1996). A set of primers for the amplification of 20 chloroplast microsatellites in Pinaceae. *Molecular Ecology* 5, 111–114.
- Vekemans X, Beauwens T, Lemaire M & Roldan-Ruiz I (2002). Data from amplified fragment length polymorphism (AFLP) markers show indication of size homoplasy and of a relationship between degree of homoplasy and fragment size. *Molecular Ecology*, 11, 139-151.
- Villemereuil P, Frichot E, Bazin E, Francois O & Gaggiotti O (2014). Genome scan methods against more complex models: when and how much should we trust them?. *Molecular Ecology* 23, 2006 – 2019.
- Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M & Zabeau M (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res.* 23, 4407–4414.

Wall C A, Koniges G J, Miller S R (2014). Divergence with gene flow in a population of thermophilic bacteria: a potential role for spatially varying selection. *Molecular Ecology* 23, 3371 – 3383.

Wang I, & Summers K (2010). Genetic structure is correlated with phenotypic divergence rather than geographic isolation in the highly polymorphic strawberry poison-dart frog. *Molecular Ecology* 19, 447 – 458.

Wang IJ (2013). Examining the Full Effects of Landscape Heterogeneity on Spatial Genetic Variation: a Multiple Matrix Regression Approach for Quantifying Geographic and Ecological Isolation. *Evolution* n/a–n/a.

Watterson G (1975). On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology* 7, 256 – 276.

Wehenkel C & Sáenz-Romero C (2012). Estimating genetic erosion using the example of *Picea chihuahuana* Martínez. *Tree Genetics & Genomes* 8, 1085 – 1094.

Whitlock M C (2004). Selection and Drift in Metapopulations. En: *Ecology, Genetics and Evolution in Metapopulations*. Editado por Hanski I & Gaggiotti O. Elsevier Inc. Pp153 - 173

White S E & Valastro S (1984). Pleistocene Glaciation of Volcano Ajusco, Central Mexico, and Comparison with the Standard Mexican Glacial Sequence. *Quaternary Research*, 21, 21-35.

White S E (1990). El ajusco: Geomorfología volcánica y Acontecimientos Glaciales Durante el Pleistoceno Superior y Comparación con Las Series Glaciales Mexicanas y Las de Las Montañas Rocallosas. Instituto Nacional de Antropología e Historia.

Wright S (1943). Isolation by distance. *Genetics*, 28, 139–156.

Wright S (1951). The genetical structure of populations. *Annals of Eugenics*, 15, 323-354.

Xue X, Wang Y, Korpelainen H & Li C (2005). Assessment of AFLP-based genetic variation in the populations of *Picea asperata*. *Silvae Genetica* 54, 24 – 30.

<http://www.conafor.gob.mx:8080/documentos/docs/13/873Abies%20religiosa.pdf>

http://www.conabio.gob.mx/conocimiento/regionalizacion/doctos/rtp_108.pdf