



**UNIVERSIDAD NACIONAL AUTONOMA
DE MEXICO**

FACULTAD DE CIENCIAS

**DIVERSIDAD GENETICA EN POBLACIONES DE
CHILE (*Capsicum annum* L.) CON DIFERENTES
GRADOS DE DOMESTICACION EN EL ESTADO DE
OAXACA, MEXICO.**

T E S I S

QUE PARA OBTENER EL TITULO DE:

BIOLOGA

P R E S E N T A:

ANA LAURA PEREZ MARTINEZ



**DIRECTOR DE TESIS:
DR. LEV ORLANDO JARDON BARBOLLA**

2018

CIUDAD UNIVERSITARIA, CD. MX.



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

HOJA DE DATOS DEL JURADO

1. Datos del alumno
Pérez
Martínez
Ana Laura
55 64 38 77 48
Universidad Nacional Autónoma de México
Facultad de Ciencias
Biología
309208707
2. Datos del tutor
Dr
Lev Orlando
Jardón
Barbolla
3. Datos del sinodal 1
Dr
Daniel Ignacio
Piñero
Dalmau
4. Datos del sinodal 2
Dra
Mariana
Benítez
Keinrad
5. Datos del sinodal 3
Dra
Erika
Aguirre
Planter
6. Datos del sinodal 4
Dra
Xitlali
Aguirre
Dugua
7. Datos del trabajo escrito
Diversidad genética en poblaciones de chile (*Capsicum annuum* L.) con diferentes grados de domesticación en el estado de Oaxaca, México
165 p
2018

Pidieron entonces a su abuela su comida.

*“Moled solamente nuestro alimento;
no deseamos más que un caldo con pimienta,
oh abuela nuestra”, dijeron.*

Popol Vuh

A mis padres y hermano

AGRADECIMIENTOS INSTITUCIONALES

A la Universidad Nacional Autónoma de México (UNAM), a la Facultad de Ciencias de la UNAM y al Centro de Investigaciones Interdisciplinarias en Ciencias y Humanidades (CEIICH) por ser mi segunda casa y brindarme la oportunidad de estudiar una carrera universitaria.

Al Dr. Luis Eguiarte y al Laboratorio de Evolución Molecular y Experimental del Instituto de Ecología de la UNAM por brindarme los medios necesarios para la realización práctica de este trabajo.

Al proyecto PAPIIT IA-202515: Diversidad agroecológica en evolución: del organismo al manejo y la cultura, y al proyecto CONACyT-Problemas Nacionales, 247730. La variación genética de las plantas cultivadas en México: estrategias para enfrentar el cambio climático por los apoyos económicos otorgados para la realización del presente trabajo.

A los miembros del jurado, el Dr. Daniel Piñero, la Dra. Mariana Benítez, la Dra. Erika Aguirre y la Dra. Xitlali Aguirre.

Al Dr. Lev Orlando Jardón Barbolla por dirigir la presente tesis.

AGRADECIMIENTOS A TITULO PERSONAL

A mis padres por apoyarme siempre en mis decisiones, por sus consejos, por su paciencia y sobre todo por el apoyo cuando elegí mi carrera universitaria. Gracias mamá, gracias papá, por creer en mí, ser mis guías de vida, por sus sacrificios, su amor y por mostrarme que trabajando y esforzándose diario es como se sobresale en la vida. Hoy les digo: ¡Por fin lo logré!

A mi hermano que me aguantó las depresiones, los momentos de frustración, felicidad y locura. Siempre me animas a seguir adelante.

Al Dr. Lev Jardón Barbolla por su apoyo, ánimos, consejos y sobre todo por su paciencia. Gracias por brindarme la oportunidad de realizar esta tesis, por la confianza, disponibilidad y los conocimientos brindados durante todo este tiempo. Gracias por ser mi guía y un ejemplo a seguir.

A los miembros del jurado el Dr. Daniel Piñero, la Dra. Mariana Benítez, la Dra. Erika Aguirre y la Dra. Xitlali Aguirre, así como al Dr. Luis Eguiarte por su tiempo, revisiones y comentarios que permitieron este trabajo fuera mucho mejor.

Al Dr. Luis Eguiarte y la Dra. Valeria Souza por permitirme trabajar en su laboratorio.

Al Laboratorio de Evolución Molecular y Experimental y a sus integrantes por su apoyo, experiencia y comentarios. A los técnicos la Dra. Erika Aguirre y la Dra. Laura Espinosa por su apoyo durante la realización práctica de esta tesis. A la Sra. Silvia por el apoyo con el material del laboratorio.

A la Dra. Gabriela Castellanos por su ayuda en el laboratorio y en los análisis genéticos. Al Dr. Jaime Gasca y al Dr. Enrique Scheinvar por los materiales y conocimientos que me ayudaron a saber por dónde empezar a programar. A la Biól. Helena Hernández por su ayuda en las lecturas de fragmentos. Gracias infinitas a la Biól. Anayansi Sierralta y a Omar Aguilar por su trabajo, apoyo y ánimos durante la extracción de ADN.

A todas las personas que colectaron las muestras de chiles. Gracias por su esfuerzo.

A toda la gente que trabaja en el campo mexicano especialmente en Oaxaca. A las familias, campesinos, agricultores y productores que sin su trabajo diario esta tesis no existiría. Gracias por compartir su sabiduría, conocimientos y plantas con nosotros.

A mi mamá y abuelita Ana que gracias a sus historias de vida contadas mientras crecía, me hicieron apreciar, respetar y admirar la vida en el campo mexicano. A mi tía Lupe por brindarles calma a mis papás cuando tenía que salir de campo.

A Yuridia que desde la prepa me apoyó y brindó su amistad. Gracias por los consejos, regaños y pláticas amenas. A Gaby por ser amiga y apoyo desde el jardín de niños.

A Alí por brindarme su amistad desde el comienzo de la carrera. Gracias por los consejos, apoyo y tardes de café. A Ara y Rosa por acompañarme desde el comienzo de esta aventura llamada tesis, por los momentos de risas y también de llanto, por las clases y el servicio social compartidos, por sus palabras de aliento y por aguantarme en los días difíciles. Un agradecimiento especial a Ara por sus comentarios a la tesis y su apoyo en todo momento. Gracias amigas.

A los compañeros del seminario de tesis. A Anayansi por su gran compañerismo y por brindarme su apoyo siempre. A Cristina por sus ánimos, comentarios y fotos que complementaron esta tesis. A Julia y Andrea por compartir conmigo mi primera salida al campo para ver chiles. A Omar y Tania gracias por su apoyo y ánimos. De cada uno de ustedes aprendí mucho.

A todos los innumerables compañeros, maestros y amigos de la carrera, equipos, voluntariados, seminarios, servicio social y taller que me hicieron crecer y madurar de una manera impresionante como profesionista. A Anita y Kevin por compartir intereses, ideas y preocupaciones con respecto a la agroecología. A Dulce, Gaby, Cris y Francisco por los momentos vividos al inicio de la carrera.

A todas las personas que conocí dentro y fuera de la UNAM, que cada vez que les platicaba acerca de mi trabajo de tesis me animaban a seguir adelante. En especial, a los amigos y garnacheros del SEFI por brindarme la oportunidad de despejarme un poco de la realidad durante todo este proceso.

Gracias a la música, a todas aquellas canciones, intérpretes y compositores que me acompañaron y dieron momentos de tranquilidad mientras escribía la tesis.

A los que se tomarán el tiempo de leer esta tesis.

A todos ustedes gracias.

RESUMEN

El chile *Capsicum annuum* es una de las especies de mayor importancia biológica, económica y cultural para México. En el estado de Oaxaca existen alrededor de 20 variedades locales de chile presentes en un gradiente de domesticación que incluye poblaciones silvestres, arvenses y domesticadas. Los chiles arvenses se pueden encontrar como plantas toleradas creciendo en sistemas de manejo como traspatios, milpas y huertos familiares bajo un proceso de domesticación inicial. En el presente trabajo se analizó la diversidad y estructura genética de 28 poblaciones y 16 variedades de chile *C. annuum* en tres regiones de las ocho reportadas para el estado de Oaxaca. Se extrajo ADN total de 341 individuos y se amplificaron 10 loci de microsatélites (SSRs) marcados con fluorescencia. La región de la Costa presentó niveles de diversidad genética ligeramente mayores ($H_E= 0.28$) a las regiones de Valles Centrales ($H_E= 0.21$) y la Cañada ($H_E= 0.24$). Se encontró una moderada diferenciación genética ($F_{ST}= 0.11$) con la mayor cantidad de variación dentro de las poblaciones, y altos niveles de endogamia ($F_{IS}= 0.32$ por región geográfica y $F_{IS}= 0.62$ por grado de domesticación). La moderada diferenciación fue atribuida al factor de la polinización cruzada (*C. annuum* es una especie autógena facultativa) al observarse niveles de consanguinidad más bajos para silvestres $F_{IS}= 0.25$ y arvenses $F_{IS}= 0.17$, las cuales son dispersadas principalmente por aves. Los patrones de diversidad y estructura genética de las poblaciones se encuentran influenciados por la altitud, el sistema de apareamiento y los sistemas de manejo tradicional (traspatios, huertos familiares y milpas) en la que crecen las poblaciones de chiles. Los datos de estructura genética obtenidos sugieren que la variedad de chile Tusta, probablemente se ha diferenciado de otras variedades cultivadas de *C. annuum* debido a su posible flujo génico con los chiles arvenses y las especies *Capsicum frutescens* y *Capsicum chinense*, mientras que el chile Costeño parece presentar altos niveles de diversidad genética asociados a la polinización cruzada que experimenta con los chiles silvestres. Los chiles arvenses presentaron niveles de diversidad genética similares a los silvestres ($H_E= 0.37$ para arvenses y $H_E= 0.35$ para silvestres). La permanencia de estos individuos en sistemas de manejo tradicional como los traspatios permitirá mantener una importante fuente de recursos genéticos para el chile *C. annuum* ante el posible declive de las poblaciones silvestres.

ABSTRACT

Capsicum annuum is one of the species of most biological, economic and cultural importance for Mexico. In the state of Oaxaca there are about 20 local varieties of pepper present in a gradient of domestication that includes wild, weedy and domesticated populations. Weed pepper can be found as tolerated plants growing in management systems such as backyards and family gardens under an initial domestication process. In the present work, the diversity and genetic structure of 28 populations and 16 varieties of *C. annuum* were analyzed in three of the eight regions reported for the state of Oaxaca. The total DNA was extracted from 341 individuals and 10 microsatellite loci (SSRs) marked with fluorescence were amplified. Coast region presented slightly higher levels of genetic diversity ($H_E = 0.28$) than Central Valleys ($H_E = 0.21$) and Cañada ($H_E = 0.24$) regions. Moderate genetic differentiation ($F_{ST} = 0.11$) was found with the greatest amount of variation within the populations, and high levels of inbreeding ($F_{IS} = 0.32$ per geographical region and $F_{IS} = 0.62$ per degree of domestication). This differentiation was attributed to the cross-pollination factor (*C. annuum* is a facultative self-pollinated species) when lower inbreeding levels were observed for wild ($F_{IS} = 0.25$) and weedy ($F_{IS} = 0.17$) populations, which are dispersed mainly by birds. The diversity patterns and genetic structure of the populations are influenced by the altitude, the mating system and the traditional management system (backyards and family gardens) in which the populations of pepper grow. The genetic structure data obtained suggest that the Tusta variety has probably differentiated from other cultivars of *C. annuum* due to its possible gene flow with the weed peppers and the *Capsicum frutescens* and *Capsicum chinense* species, while the Costeño pepper seems to present high levels of genetic diversity associated with the cross-pollination it experiences with wild peppers. The weed peppers showed levels of genetic diversity similar to the wild ones ($H_E = 0.37$ for weeds and $H_E = 0.35$ for wild). The permanence of these individuals in traditional management systems such as backyards, will allow to maintain an important source of genetic resources for *C. annuum* in the face of the possible decline of wild populations.

INDICE

I. INTRODUCCION.....	14
II. MARCO TEORICO.....	16
II.1. Diversidad genética.....	16
II.1.1. Domesticación y datos genéticos en plantas cultivadas.....	18
II.1.2. Conservación genética.....	22
II.1.3. Marcadores moleculares, microsatélites (SSRs).....	23
II.2. Mesoamérica: centro de origen y diversidad de plantas cultivadas.....	25
II.2.1. Estudios de diversidad genética de plantas cultivadas en México.....	27
II.3. El género <i>Capsicum</i>	28
II.3.1. La especie <i>Capsicum annuum</i>	29
II.3.2. Domesticación y diversificación de <i>C. annuum</i> en México.....	32
II.3.3. <i>C. annuum</i> en Oaxaca.....	35
III. JUSTIFICACION.....	37
IV. HIPOTESIS.....	38
V. OBJETIVOS.....	39
V.1. Objetivo General.....	39
V.2. Objetivos Particulares.....	39
VI. METODOLOGIA.....	40
VI.1. Sitio de estudio.....	40
VI.2. Obtención de las muestras.....	40
VI.3. Extracción de ADN.....	45
VI.4. Amplificación de microsatélites (SSRs).....	45
VI.5. Análisis de diversidad genética.....	47
VI.6. Análisis de varianza molecular.....	50
VI.7. Distancia genéticas y aislamiento por distancia.....	51
VI.7.1. Distancias genéticas.....	51
VI.7.2. Aislamiento por distancia.....	52

VI.7.3. Barreras geográficas.....	53
VI.8. Estructura poblacional.....	54
VI.8.1. Análisis multivariado.....	54
VI.8.2. Análisis bayesiano.....	55
VI.8.3. Análisis gráfico poblacional.....	56
VI.8.4. Cuellos de botella y tamaño efectivo poblacional.....	58
VII. RESULTADOS.....	60
VII.1. Amplificación y análisis de microsatélites (SSRs).....	60
VII.2. Análisis poblacional.....	62
VII.2.1. Diversidad genética.....	62
VII.2.2. Análisis de varianza molecular.....	65
VII.2.3. Distancias genéticas y aislamiento por distancia.....	67
VII.2.3.1. Distancias genéticas.....	67
VII.2.3.2. Aislamiento por distancia.....	70
VII.2.3.3. Barreras geográficas.....	72
VII.2.4. Estructura poblacional.....	75
VII.2.4.1. Análisis multivariado (DAPC).....	75
VII.2.4.2. Análisis bayesiano (STRUCTURE).....	78
VII.2.4.3. Análisis gráfico poblacional.....	82
VII.2.4.4. Cuellos de botella y tamaño efectivo poblacional (N_e).....	85
VII.3. Análisis por grado de domesticación (variedades).....	87
VII.3.1. Diversidad genética.....	87
VII.3.2. Análisis de varianza molecular.....	90
VII.3.3. Distancias genéticas y aislamiento por distancia.....	91
VII.3.3.1. Distancias genéticas.....	91
VII.3.3.2. Aislamiento por distancia.....	93
VII.3.4. Estructura poblacional.....	93
VII.3.4.1. Análisis multivariado (DAPC).....	93
VII.3.4.2. Análisis bayesiano (STRUCTURE).....	96
VII.3.4.3. Análisis gráfico poblacional.....	101
VII.3.4.4. Cuellos de botella y tamaño efectivo poblacional (N_e).....	102

VIII. DISCUSION.....	104
VIII.1. Diversidad genética.....	104
VIII.2. Estructura genética.....	115
VIII.3. Manejo humano moldeador de la diversidad genética.....	121
VIII.4. <i>C. annuum</i> y las especies <i>C. frutescens</i> y <i>C. chinense</i> en la Costa de Oaxaca.....	125
IX. CONCLUSIONES.....	130
X. REFERENCIAS.....	132
ANEXOS.....	150
ANEXO 1. Fotos de las variedades de chiles <i>C. annuum</i> analizadas.....	150
ANEXO 2. Extracción de ADN, protocolo CTAB-miniprep (Doyle y Doyle, 1987). Modificado.....	151
ANEXO 3. SCRIPTS para análisis genéticos en R.....	152
ANEXO 4. Frecuencias de alelos nulos por locus en las poblaciones.....	155
ANEXO 5. Desequilibrio de ligamiento.....	156
ANEXO 6. Resultados de la prueba ANOVA.....	157
ANEXO 7. Valores F_{CT} obtenidos de las 15 corridas en el análisis SAMOVA.....	158
ANEXO 8. Valores obtenidos del DAPC.....	159
ANEXO 9. Valores obtenidos de STRUCTURE.....	163

I. INTRODUCCION

Se considera que los chiles (*Capsicum spp.*) son de las primeras plantas domesticadas en Mesoamérica (Nuez, *et al.*, 2003). Por lo menos una de sus especies, *Capsicum annuum*, fue domesticada en esta zona, en la época prehispánica, y fue un importante producto alimenticio y de tributo en la época de la conquista europea, además de formar parte de una larga tradición cultural en México (Long-Solís, 2012). Su cultivo tiene una gran importancia económica para México, al ser el segundo productor y el tercer exportador de chile verde en el mundo (SAGARPA, 2016). En 2015, se sembraron 153 mil hectáreas con una producción de 2.8 millones de toneladas, siendo Chihuahua y Sinaloa los principales productores de chile verde en el país con 800, 000 y 600, 000 toneladas producidas respectivamente (SIAP, 2015).

Existen 60 variedades de chile cultivadas localmente en México y la gran mayoría de estas variedades no se encuentran en los grandes centros de producción, sino en sistemas de traspatios, jardines o huertos familiares (Jardón-Barbolla, 2017). Estos sistemas de manejo son comunes en el estado de Oaxaca, donde las familias tienden a cultivar sus propios chiles para consumo familiar, o los llegan a vender, pero con bajos volúmenes en el mercado (Castellón-Martínez, *et al.*, 2012). En 2015, Oaxaca aportó alrededor de 5 mil toneladas a la producción de chile verde incluyendo a las variedades Jalapeño, Serrano, Soledad y de Agua; y 900 toneladas a la producción de chile seco de las variedades Ancho, Costeño, Pasilla y Tabaquero (SIAP, 2015). En términos porcentuales, la contribución de Oaxaca al volumen nacional de producción es del 0.3%, pero en términos del número de variedades locales su contribución es de más del 30% del total nacional.

En México, podemos encontrar la mayor diversidad morfológica y genética de chile distribuida a lo largo y ancho de todo el país, tanto de formas cultivadas como silvestres (Aguilar-Rincón, *et al.*, 2010). Esta diversidad y estructura genética es notable gracias a sus sistemas de dispersión, distribución geográfica, diversos sistemas de manejo humano y cuellos de botella (Loaiza-Figueroa, *et al.*, 1989; Hernández-Verdugo, *et al.*, 2001; Oyama, *et al.*, 2006; Aguilar-Meléndez, *et al.*, 2009; González-Jara, *et al.*, 2011). En el estado de Oaxaca resulta interesante conocer cómo se encuentra distribuida esta diversidad y los procesos evolutivos que actúan sobre ésta, al existir alrededor de 20 variedades locales de chile *C. annuum* en formas silvestres, arvenses y domesticadas sometidas a diversos sistemas de manejo y a un proceso de domesticación activo hoy en día (Caballero, *et al.*, 2004; Jardón-Barbolla, 2017).

En el presente trabajo se analiza la diversidad y estructura genética de poblaciones de chile (*C. annuum*) en el estado de Oaxaca y se describen los patrones de diversidad genética presentes según el grado de domesticación (poblaciones silvestres, poblaciones arvenses y poblaciones de variedades cultivadas localmente).

II. MARCO TEORICO

II.1. Diversidad genética

La diversidad genética o variabilidad genética puede definirse como las variaciones heredables que ocurren en el material genético de cada organismo, entre los individuos de una población y entre las poblaciones dentro de una especie, siendo uno de los componentes fundamentales de la biodiversidad (Piñero, *et al.*, 2008a). Para su estudio, la genética de poblaciones trata de explicar cómo cambia la variación genética dentro y entre poblaciones de una especie, así como comprender el proceso de la adaptación y los fundamentos de la biología evolutiva moderna, a través de aproximaciones empíricas y experimentales, y modelos teóricos (Eguiarte, 2009; Hedrick, 2011).

Con el descubrimiento de la estructura de la molécula del ADN (ácido desoxirribonucleico), en 1953, se sentaron las bases que permitieron la identificación de variantes en el material hereditario (Engels, *et al.*, 2006). Por otro lado, el desarrollo de la reacción en cadena de la polimerasa (PCR, por sus siglas en inglés) condujo a una explosión de diferentes herramientas que tomaron ventaja de los métodos de amplificación de segmentos de ADN, incluyendo no solo los métodos actuales de secuenciación, sino también de varios tipos de marcadores polimórficos basados en PCR, dando paso a un enfoque en la producción de huellas genéticas, que suplantaron las técnicas basadas en hibridación (Emshwiller, 2006; Romero, *et al.*, 2009). Hoy en día, la huella genética se ha aplicado en una gran gama de investigación biológica que incluyen la biología forense, la medicina y la agricultura (Romero, *et al.*, 2009).

En la agricultura, los estudios pioneros de Darwin (1868) y De Candolle (1883) sobre el origen y la evolución de las plantas cultivadas se basaban principalmente en

evidencia morfológica y arqueológica de las plantas (Burger, *et al.*, 2008). En 1926, el genetista ruso Nicolai Vavilov (Vavilov, 1926), incorpora evidencia genética por medio de evidencia morfológica, para definir los centros de origen y domesticación de las plantas cultivadas, en relación a la presencia de sus formas silvestres y la variabilidad de la planta (Harlan, 1975; Jardón-Barbolla, 2015). Posteriormente, la propuesta del genetista Jack Harlan, de la existencia de un conjunto de rasgos morfológicos que se seleccionan en las etapas iniciales de la domesticación, permitiría estimar la divergencia morfológica, y por tanto genética de la planta domesticada en relación a su pariente silvestre más cercano (Harlan, 1973, 1975; Burger, *et al.*, 2008).

El mayor cambio en el estudio de la variación genética en la evolución y la domesticación ocurrió a mediados del siglo XX tras el desarrollo y la introducción de las herramientas moleculares (Burguer, *et al.*, 2008; Charlesworth y Charlesworth, 2017). Estos avances aumentaron el conocimiento de los rasgos morfológicos de la planta domesticada definidos por un gran número de genes relacionados al proceso de domesticación, identificados y aislados por medio de genes candidatos, loci de rasgos cuantitativos (QTLs), mapeo y clonación, y técnicas de secuenciación y genotipado de muchas plantas cultivadas (Meyer y Purugganan, 2013; Olsen y Wendel, 2013).

Desde 1980, los marcadores moleculares nos han permitido conocer cómo están distribuidas las frecuencias alélicas dentro y entre poblaciones de plantas cultivadas (Hartl y Clark, 2007; Burger, *et al.*, 2008). Esta distribución, también conocida como estructura genética, puede ser afectada, tanto por las fuerzas evolutivas de selección, deriva génica, flujo génico y mutación, como por la historia de vida, sistema de apareamiento, distribución geográfica y mecanismos de dispersión de la especie (Brown, 1978; Hernández-Verdugo, *et al.*, 2001; Hedrick, 2011).

En los últimos años, el interés en el estudio de los genes relacionados a la domesticación, las diferencias en las secuencias de ADN o proteínas y los análisis de los genomas entre diversas especies ha aumentado el conocimiento de las huellas genéticas del proceso de domesticación, la historia demográfica y el origen de un gran número de plantas (Azofeita-Delgado, 2006; Hartl y Clark, 2007; Meyer y Purugganan, 2013; Olsen y Wendel, 2013; Kalia, *et al.*, 2011).

II.1.1. Domesticación y datos genéticos en plantas cultivadas

La domesticación es el resultado de un proceso de selección que tiende a incrementar la adaptación de plantas o animales que son cultivados, criados y utilizados por el ser humano (Gepts, 2004; Emshwiller, 2006; Zeder, *et al.*, 2006). En las plantas, la domesticación se basa en el cambio genético de las poblaciones domesticadas, resultando en una mejor adaptación al medio ambiente creado para su cultivo, pero al mismo tiempo mal adaptadas a su hábitat original en el medio silvestre, haciéndolas totalmente dependientes del ser humano al perder su mecanismo de dispersión, ser trasladadas a nuevos ambientes o ser propagadas vegetativamente (Doebley, *et al.*, 2006; Ladizinskiv, 2012).

La domesticación es un proceso evolutivo continuo, que actúa sobre plantas silvestres, semi-silvestres (similares a las silvestres con un proceso de domesticación inicial) y plantas domesticadas. Estas últimas ya presentan diferencias morfológicas y genéticas que pueden ayudar a distinguirlas de su pariente silvestre (Casas, *et al.*, 2007; Pickersgill, 2007). A estas diferencias morfológicas y fisiológicas de las plantas

domesticadas se le conoce como síndrome de domesticación (Gepts, 2004; Emswiller, 2006; Pickersgill, 2007).

Los rasgos asociados al síndrome de domesticación en plantas anuales van desde cambios en el tamaño del fruto o semilla, crecimiento determinado, dominancia apical y una gran variabilidad de formas; hasta cambios fisiológicos en la latencia de la semilla, la pérdida de protecciones químicas o mecánicas, sincronización floral, hábito de la planta y la pérdida de los mecanismos de dispersión de semillas (Gepts y Papa, 2002; Doebley, *et al*, 2006; Pickersgill, 2007). Los genes que controlan estos rasgos son afectados por la domesticación (Emswiller, 2006).

En la domesticación de las plantas las fuerzas evolutivas de mutación, selección, deriva génica y flujo génico se ven involucradas en la reestructuración de la diversidad genética de la planta domesticada en relación a su pariente silvestre, dejando huellas genómicas del proceso de domesticación en las plantas cultivadas actuales (Heslop-Harrison y Schwarzacher, 2011; Ladizinskiy, 2012; Olsen y Wendel, 2013). El proceso de mutación implica cambios en nucleótidos (por inversión, supresión, sustitución o adición), genes o cromosomas (Hedrick, 2011). La mutación es la única medida por la cual nuevos alelos se forman incrementando la diversidad genética siendo un elemento esencial en la evolución y por consiguiente en la domesticación de cultivos (Ladizinskiy, 2012).

Existen dos tipos de selección asociadas a cómo opera la domesticación cuando las plantas silvestres comienzan a domesticarse: i) la selección consciente o intencional, aplicada deliberadamente sobre rasgos de interés para el ser humano; y ii) la selección inconsciente o automática aplicada involuntariamente cuando las plantas son trasladadas de su hábitat silvestre original a un nuevo ambiente creado por el ser humano,

conduciendo a cambios drásticos en las presiones de selección, al generar adaptaciones a nuevas condiciones ambientales o al reducir la adecuación del individuo (Zohary, 2004).

La selección afecta diferencialmente la diversidad dependiendo de la naturaleza de los rasgos seleccionados y de su base genética, conduciendo frecuentemente a un barrido selectivo, producto de la pérdida de diversidad genética en sitios ligados al gen seleccionado, dejando huellas de selección o de adaptación reducida en regiones genómicas específicas con baja o ninguna tasa local de recombinación genética (Emshwiller, 2006; Olsen y Wendel, 2013; Charlesworth y Charlesworth, 2017).

La deriva génica es el cambio al azar de las frecuencias alélicas de generación a generación a causa de un pequeño tamaño poblacional (Allendorf y Luikart, 2009; Hedrick, 2011). Este proceso puede influir en la pérdida de diversidad genética en las poblaciones de plantas domesticadas y en el incremento de la diferenciación genética entre poblaciones (Ellstrand y Elam, 1993; Olsen y Wendel, 2013; Ellegren y Galtier, 2016). Un fuerte cuello de botella resultaría en un rápido decaimiento de la heterocigosis y una eventual fijación de alelos, disminuyendo la variación genética dentro de las poblaciones de plantas cultivadas (Ellstrand y Elam, 1993; Emshwiller, 2006; van de Wouw, *et al.*, 2009; Ellegren y Galtier, 2016).

En cambio, el flujo génico puede mantener o incrementar la diversidad genética de la planta domesticada, en el caso de la exogamia, o disminuir su diversidad por medio de la endogamia, dependiendo de si los alelos migrantes son neutrales, detrimentales o benéficos para la población receptora (Ellstrand y Elam, 1993; Ellstrand, *et al.*, 1999; Olsen y Wendel, 2013). En las plantas domesticadas el flujo génico puede ser nulo si son autocompatibles, se encuentran aisladas geográficamente, o no sobrelapan sus tiempos de floración propiciando la diferenciación genética, mientras que en las especies

incompatibles, con fenologías de floración similares y crecimiento espacial cercano, los niveles de flujo génico pueden variar (Ellstrand, *et al.*, 1999; Ellegren y Galtier, 2016).

Los datos genéticos han comenzado a utilizarse para investigar los efectos y eventos de la domesticación en muchas plantas cultivadas (Emshwiller, 2006; Gross y Olsen, 2010). Por ejemplo, la duración del evento de domesticación potencialmente puede deducirse de la gravedad del cuello de botella en todo el genoma. Los datos genéticos son complementarios a datos morfológicos y arqueológicos que proveen información sobre las poblaciones silvestres (Emshwiller, 2006).

Los marcadores de ADN han hecho posible producir mapas de ligamiento de muchas plantas cultivadas y determinar mediante la búsqueda de asociaciones con esos marcadores, cuántos y qué regiones del genoma afectan a un rasgo cuantitativo dado (o loci de rasgos cuantitativos, QTL) (Pickersgill, 2007). Además, nos permiten contestar preguntas evolutivas muy puntuales relacionadas con el monitoreo del flujo génico, introgresión y análisis de paternidad para hacer inferencias de parámetros demográficos y determinar patrones evolutivos de los procesos históricos del origen de las especies, formas o razas, a través de las frecuencias alélicas de la planta domesticada en relación con las poblaciones silvestres. Se cree que la diversidad alélica de la planta cultivada es un subconjunto de la población silvestre de la cual es derivada (Rentería-Alcántara, 2007; Gross y Olsen, 2010).

Hoy en día, los datos morfológicos y moleculares son útiles para establecer hipótesis de relaciones filogenéticas entre organismos, estimar la variación dentro y entre las poblaciones y probar hipótesis de adaptaciones ecológicas (Rentería-Alcántara, 2007). Estos estudios están cambiando nuestra comprensión del origen, el ritmo y la duración del proceso de domesticación (Gross y Olsen, 2010; Tang, *et al.*, 2010).

II.1.2. Conservación genética

El incremento en los avances tecnológicos, la mecanización de la agricultura, el uso de fertilizantes y pesticidas, los cambios en las preferencias y demandas alimentarias, el desplazamiento de agricultores, el cambio climático y la aplicación de la genética en las actividades de mejoramiento de las principales plantas domesticadas, han resultado, entre otros, en la sustitución significativa de variedades locales por variedades mejoradas genéticamente y en la pérdida de diversidad genética de las plantas cultivadas (Brush, 2004; Engels, *et al.*, 2006; van de Wouw, *et al.*, 2009; Heslop-Harrison y Schwarzacher, 2011).

La erosión genética en plantas domesticadas, es decir, la pérdida de su diversidad genética, ha conducido a la creación de programas dirigidos a la conservación de los recursos genéticos de las principales plantas cultivadas, a través de la conservación *ex situ* y la conservación *in situ* (Brush, 2000; van de Wouw, *et al.*, 2009).

La conservación *ex situ* consiste en la colecta y preservación sistemática de la variación genética de plantas cultivadas (almacenamiento de semillas o tejidos), incluyendo muestras de sus progenitores silvestres y parientes relacionados en bancos de germoplasma, jardines botánicos o centros dedicados a la investigación agrícola (Brown, 1978; Brush, 2000; Engels, *et al.*, 2006; van de Wouw, *et al.*, 2009). Por otra parte, la conservación *in situ* se refiere a la protección de las poblaciones cultivadas en sus hábitats naturales, jardines o en los campos de los agricultores, los cuales desempeñan un papel importante en la preservación de la diversidad genética de muchas plantas domesticadas (Brown, 1978; Brush, 2000; 2004; Engels, *et al.*, 2006).

Como en cualquier grupo de plantas o animales en la naturaleza, la futura evolución de las plantas cultivadas (la mayoría, aunque no en todas las plantas), dependerá en parte de la disponibilidad de la diversidad genética (Frankel, 1974; Ladizinskiy, 2012). Al menos una parte de esta diversidad es adaptativa, considerándose necesario determinar los procesos y parámetros que afectan la variación y la adaptación en términos de conservación, ya sea conociendo el flujo génico, la subestructuración poblacional, el tamaño de las poblaciones y/o la variación dentro y entre poblaciones (Frankel, 1974).

Los análisis de diversidad, medidas de similitud genética o diferencias entre especies de plantas cultivadas se convierten en información importante en la conservación y desarrollo de variedades (Romero, *et al.*, 2009; van de Wouw, *et al.*, 2009; Kalia, *et al.*, 2011). Estos análisis toman relevancia en países como México, el cual es centro de domesticación y diversificación de algunas plantas con altos niveles de diversidad genética, como el maíz, la calabaza, el frijol y el chile (Piñero, *et al.*, 2008a).

II.1.3. Marcadores moleculares, microsatélites (SSRs)

Los estudios de diversidad genética han empleado dos clases de marcadores: los fenotípicos y los genéticos. Los marcadores fenotípicos se han basado en caracterizar e identificar características fenotípicas de los organismos (la base genética de la mayoría de los caracteres morfológicos es asumida) que deben ser descubiertos y delimitados generalmente sin ningún criterio explícito para la selección o la codificación del carácter, por lo que tienen el potencial de ser arbitrarios (Azofeita-Delgado, 2006; Rentería-Alcántara, 2007). Por otro lado, los marcadores moleculares nos permiten caracterizar e

identificar a los organismos, a través de conocer cómo se encuentran las proporciones de alelos en las poblaciones naturales de manera indirecta, como con los análisis de proteínas, o de manera directa con los estudios de ADN (Rentería-Alcántara, 2007).

Los marcadores moleculares son capaces de detectar polimorfismos y pueden ser dominantes o codominantes (Simpson, 1997). Estos marcadores son clasificados como: i) marcadores basados en hibridación (polimorfismos de la longitud de los fragmentos de restricción, RFLPs); ii) marcadores basados en secuencias (polimorfismo de nucleótido único, SNPs); y iii) marcadores basados en PCR, como polimorfismos de ADN amplificados al azar (RAPDs), polimorfismos de longitud de los fragmentos amplificados (AFLPs), secuencias repetidas intersimples (ISSRs) y microsatélites o secuencias simples repetidas (SSRs) (Kalia, *et al.*, 2011).

Los microsatélites o SSRs son secuencias de ADN formadas de 1 a 6 pares repetidos de bases (mononucleótidos, dinucleótidos o tetranucleótidos). Estos se encuentran en regiones codificantes y no codificantes del ADN (Kalia, *et al.*, 2011; Rentería-Alcántara, 2007) y tienen la ventaja de ser altamente polimórficos como resultado de una alta tasa de mutación (10^{-3} o 10^{-4} por generación) (Allendorf y Luikart, 2009). Estos marcadores siguen el modelo de mutación paso a paso (S.M.M. por sus siglas en inglés) donde cada mutación crea un alelo nuevo, ya sea agregando o eliminando una unidad repetida del microsatélite (Balloux y Lugon-Moulin, 2002; Allendorf y Luikart, 2009). Además, los SSRs son específicos de una especie en particular o especies relacionadas y están distribuidos a través de todo el genoma nuclear, incluso se pueden encontrar en el genoma del cloroplasto y mitocondria. Este tipo de marcador ha ganado considerable importancia en la genética de plantas por muchos de sus atributos como la hipervariabilidad, naturaleza multialélica, herencia codominante,

reproducibilidad, abundancia relativa, gran cobertura del genoma, manejabilidad y el alto rendimiento del genotipado (Kalia, *et al.*, 2011). Sin embargo, las desventajas en analizar un locus por experimento los hace muy laboriosos, de la misma forma, deben identificarse y secuenciarse en regiones genómicas concretas, aunque una vez que se consigue pueden llegar a ser muy informativos (Azofeita-Delgado, 2006).

II.2. Mesoamérica: centro de origen y diversidad de plantas cultivadas

En 1926, el genetista ruso Nikolai Vavilov definió ocho centros de origen y diversidad de muchas plantas cultivadas. Estos centros presentan una gran diversidad genética dentro y entre especies, condiciones ecológicas y topográficas diversas, una distribución en regiones tropicales y subtropicales, y una distribución cercana de las plantas domesticadas con sus formas silvestres. Además, Vavilov reconoció que el trabajo humano era clave en la evolución de las plantas domesticadas e incorporó evidencia lingüística, histórica, arqueológica y antropológica para definir estos centros (Engels, *et al.*, 2006; Jardón-Barbolla, 2015).

Mesoamérica área cultural localizada entre el centro-norte de México y el norte de Costa Rica, es uno de estos centros de origen y diversidad, y es uno de los tres centros primarios de domesticación y origen de la agricultura en el mundo. Se utilizan cerca de 5, 000 a 7, 000 especies de plantas y se están domesticando 200 plantas nativas que coexisten cerca de sus parientes silvestres (Harlan, 1971; Casas, *et al.*, 2007). Estas especies domesticadas incluyen cereal (maíz, *Zea mays* L.), granos (amaranto, *Amaranthus spp.* y frijol común *Phaseolus vulgaris* L. y *P. lunatus* L.), especies y condimentos (chile, *Capsicum annum* L.), fibras (algodón, *Gossypium hirsutum* L., y

henequén, *Agave fourcroydes* Lem.), vegetales (calabaza, *Cucurbita spp.*, jitomate, *Solanum lycopersicum* L., nopalero, *Opuntia streptacantha* Lem., y tomate *Physalis philadelphica* Lam.), estimulantes (cacao, *Theobroma cacao* L.), raíces y tubérculos (jícama, *Pachyrhizus erosus* L.), y árboles frutales (papaya, *Carica papaya* L., y aguacate, *Persea spp.*, Engels, *et al.*, 2006; Pickersgill, 2016).

México es uno de los principales puntos de selección y domesticación de estas plantas. Ha dado al mundo cerca de 118 especies de plantas útiles como alimentos, textiles y colorantes. El 15% de las plantas destinadas como alimento en el mundo fueron originadas aquí (Mapes y Basurto, 2016). En México, estas plantas pueden encontrarse en estado cultivado, silvestre o arvense (gradiente de domesticación observado por Vavilov en sus investigaciones) y pueden estar sometidas a diversas formas de manejo humano (Jardón-Barbolla, 2015)

Las plantas arvenses se conforman de dos grupos: las toleradas y las fomentadas. Las plantas toleradas pueden crecer espontáneamente en ambientes antropogénicos, sin requerir ayuda del ser humano, adaptándose a las condiciones de manejo del policultivo, en cambio las fomentadas también son plantas que presentan adaptaciones a estos ambientes pero cuyas semillas son guardadas ocasionalmente por los agricultores para una posterior siembra (Rendón-Aguilar, *et al.*, 2017). Tanto las plantas silvestres como las arvenses pueden ser: silvestres, favorecidas, protegidas, semicultivadas, o incluso cultivadas por el ser humano. En algunos casos, se dice que están pasando por un proceso de domesticación incipiente. Este proceso puede estudiarse a través de las huellas genéticas en el gradiente de domesticación (Jardón-Barbolla, 2015; Mapes y Basurto, 2016).

II.2.1. Estudios de diversidad genética de plantas cultivadas en México

México es un centro importante de domesticación y diversidad de numerosas plantas. Esta diversidad no es solo el resultado de factores ambientales y biológicos, sino también de la interacción del ser humano y la planta. A raíz de sus necesidades, intereses, prácticas y conocimientos, los seres humanos han ido conformando y manteniendo esta diversidad en sus campos, solares, huertos, jardines y plantaciones (Bellón, *et al.*, 2009).

Los datos moleculares mediante el uso de isoenzimas, RAPDs, AFLPs y microsatélites confirman esta gran diversidad. Los estudios se han enfocado en el origen, flujo génico, hibridación y gradientes de variación de la planta cultivada, brindándonos datos importantes para su conservación y el estudio de la evolución. Tal es el caso del maíz, los frijoles, las calabazas, los chiles, los nopales, los agaves, los aguacates, la vainilla y el algodón. Sin embargo, en plantas nativas como la guayaba, el camote, la jícama, el jitomate, el cacao y la papaya son pocos los estudios realizados (Piñero, *et al.*, 2008a; Bellón, *et al.*, 2009).

Los estudios de diversidad genética se han centrado en el maíz, la planta cultivada más importante en México. Esta planta forma parte de un sistema de policultivo conocido como milpa, principal sistema agrícola para la producción de alimentos en Mesoamérica. En la milpa el maíz se planta en filas, permitiendo su asociación con otras plantas domesticadas como el frijol, la calabaza y plantas arvenses como los quelites y el chile (Mapes y Basurto, 2016).

En el caso de las calabazas (*Cucurbita*) existen pocos trabajos que estudien su diversidad genética (Piñero, *et al.*, 2008a). Sin embargo, diversos estudios de variabilidad del maíz (*Zea mays*) y el frijol (*Phaseolus spp.*), tanto morfológicos como genéticos,

muestran que existe una gran diversidad en poblaciones mexicanas. Hoy, esta diversidad persiste en los sistemas de agricultura tradicional, aunado a la existencia de sus parientes silvestres en poblaciones mexicanas (Bellón, *et al.*, 2009).

En el Chile *Capsicum spp.*, los estudios de diversidad apuntan a que ha sido domesticado aparentemente en diversos lugares y de distintas formas, sugiriendo que su domesticación ha tomado distintas direcciones (Loaiza Figueroa, *et al.*, 1989; Oyama, *et al.*, 2006; Perry y Flannery, 2007; Aguilar-Meléndez, *et al.*, 2009; González-Jara, *et al.*, 2011; Kraft, *et al.*, 2014). En México, se ha observado que los sistemas de cruzamiento y la distribución geográfica de las poblaciones de Chile influyen en su estructura genética (Oyama, *et al.*, 2006; Aguilar-Meléndez, *et al.*, 2009; González-Jara, *et al.*, 2011).

II.3. El género *Capsicum*

El género *Capsicum*, miembro de la familia Solanaceae, se originó en América del Sur e incluye cerca de 30 especies de las cuales 22 son endémicas del continente americano abarcando desde el sur de Estados Unidos hasta el norte de Argentina (McLeod, *et al.*, 1983, Loaiza Figueroa, *et al.*, 1989, Hernández-Verdugo, *et al.*, 2001; Oyama, *et al.*, 2006). El género incluye cinco especies domesticadas (*C. annuum* L., *C. chinense* Jacq., *C. frutescens* L., *C. baccatum* L. y *C. pubescens* R. y P.), de las cuales cuatro son cultivadas en México (Pickersgill, 1971; 1997; McLeod, *et al.*, 1983, Loaiza Figueroa, *et al.*, 1989; Hernández-Verdugo, *et al.*, 1999, 2001; Oyama, *et al.*, 2006).

Los chiles domesticados y sus parientes silvestres forman un grupo conocido como chiles verdaderos que de manera informal son divididos en dos grandes grupos: el grupo de flores moradas (*C. pubescens*) y el grupo de flores blancas (*C. annuum*, *C. chinense*,

C. frutescens y *C. baccatum*) (Hernández-Verdugo, et al., 1999). Estas especies fueron domesticadas en diferentes partes de los trópicos de América antes de la conquista europea. *C. baccatum* fue domesticada en Bolivia y *C. pubescens* en la región andina. El grupo cercanamente relacionado formado por *C. annuum*, *C. chinense* y *C. frutescens* evolucionó en los bajos trópicos de América y el Caribe con *C. annuum* dominando México, *C. frutescens* el Caribe y *C. chinense* el Amazonas (Long-Solís, 2012; Pickersgill, 2007; 2016; Eshbaugh, 2012).

En México la distribución de las especies de *Capsicum* es amplia. La especie *C. frutescens* es el principal ingrediente de la salsa Tabasco y se localiza al sur del país. El chile habanero *C. chinense* se adapta a climas cálido húmedos en la península de Yucatán y la especie *C. pubescens* conocido como chile manzano o perón se cultiva en el centro del país. Por otro lado, la especie *C. annuum* es la mejor adaptada a tierras altas mexicanas, pero más propensa al ataque de plagas en climas húmedos tropicales. Es el chile económicamente más importante a escala mundial y, por consecuencia, es la especie más estudiada en su diversidad, domesticación y genética (Pickersgill, 2007; 2016).

II.3.1. La especie *Capsicum annuum*

La especie *C. annuum* es una planta anual, perenne de zonas tropicales, con flores blancas, anteras azules o moradas y un cáliz dentado. Su fruto crece en posición colgante y es de color verde a amarillo cuando se encuentra fresco y amarillo, rojo o café cuando está maduro. Sus semillas son redondeadas y varían de tonos en color amarillo a crema (Long-Solís, 2012). Los parientes silvestres (*C. annuum* var. *grabriusculum*) son

herbáceas perennes de vida corta, erectas o hierbas trepadoras de 4 m de alto con una sola flor blanca por nodo y un cáliz rudimentario con o sin hendiduras. Sus frutos son más pequeños, erectos, deciduos, rojos y picantes (Hernández-Verdugo, *et al.*, 1999, 2001). Se diferencian de las domesticadas en la posición, tamaño, color, peso del fruto y las semillas son más pequeñas que en plantas cultivadas (Long-Solís, 2012; ver Fig. 1).

C. annuum es considerada una especie autocompatible a diferencia de otras solanáceas como el tomate (*Lycopersicon peruvianum* Mil.), la papa (*Solanum tuberosum* L.) y el tabaco (*Nicotiana tabacum* L.), las cuales presentan sistemas de autoincompatibilidad que conducen a la exogamia y a mantener la variabilidad genética de las especies (Matton, *et al.*, 1994; Onus y Pickersgill, 2004; Pickersgill, 1997). En cambio, en el género *Capsicum* suele predominar la endogamia (McLeod, *et al.*, 1983; Pickersgill, 1997).

Los sistemas reproductivos de *C. annuum* varían entre la autogamia (autopolinización) y la alogamia (polinización cruzada). En las formas no domesticadas y en las formas domesticadas de fruto pequeño, generalmente el estigma sobresale por encima de las anteras, por lo que domina la alogamia, siendo polinizadas por aves o insectos. Dentro de las variedades cultivadas de fruto grande, globoso, largo y grueso, el estilo suele ser más corto, dominando la autogamia (Nuez, *et al.*, 2003).

C. annuum está ampliamente distribuida en todo México con poblaciones silvestres localizadas, del oeste, de Sonora a Chiapas, y al este, de Tamaulipas a la Península de Yucatán (Oyama, *et al.*, 2006; Pickersgill, 2016). Al norte del país los chiles silvestres se pueden encontrar asociados a plantas de acebuche (*Celtis pallida* Torr.), mezquite (*Prosopis sp.*), o cactáceas columnares. Al sur del país, las plantas silvestres pueden ser encontradas como maleza o plantas ruderales en mayor frecuencia cerca de paisajes

perturbados por el ser humano tales como remanentes de bosque tropical caducifolio, caminos, huertos, jardines, pastizales y alrededor de los campos de cultivo (Oyama, *et al.*, 2006; Kraft, *et al.*, 2014).

Los chiles silvestres son conocidos localmente como *chile Piquín*, *chile de Monte*, *Chiltepín* y *Chiltepec*, y han sido registrados a más de 2, 000 msnm pero regularmente los encontramos en bajas altitudes (Pickersgill, 2016). Generalmente crecen a una temperatura entre 20°C y 26°C. Los lugares más fríos son tierras altas del centro de México y los lugares más cálidos son las regiones costeras del sur de México, siendo los más adecuados para su crecimiento el sureste, y las costas occidental y oriental de México (Kraft, *et al.*, 2014).

En México existe una gran diversidad morfológica y genética de chile *C. annum*, al encontrarse aproximadamente 60 tipos de chiles distribuidos a lo largo y ancho de todo el país en formas cultivadas y silvestres. La diversidad morfológica y el amplio rango de adaptabilidad de la planta se deben a sus características de picor, sabor, aroma y color, lo que los hace tener una gran variedad de usos alimenticios, religiosos, medicinales e industriales. Solo algunos tipos de chile son ampliamente reconocidos en el país (Serrano, Jalapeño, Ancho, Guajillo y los chiles dulces o pimientos), ya que la mayor parte de la diversidad solo es conocida y utilizada a nivel regional o local (Aguilar-Rincón, *et al.*, 2010; Long-Solís, 2012; Mapes y Basurto, 2016).



Figura 1. Características de *C. annuum*. a) Planta cultivada: arbusto pequeño (según el tipo de chile al que pertenezcan y las condiciones ambientales crecen entre 30 y 75 cm), el fruto es más grande que el silvestre, crece en forma de pendiente y oculto por las hojas. b) Flor de *C. annuum*: blanca, pentámera y con anteras azules o moradas. b) y c) Chile silvestre *C. annuum* var *glabriusculum*: el fruto crece erecto y pequeño separándose con facilidad del cáliz. La planta alcanza los 4 m de altura, crece como maleza y presenta mayor ramificación que la planta cultivada. Fotos tomadas por Andrea Moctezuma y Ana Laura Pérez.

II.3.2. Domesticación y diversificación de *C. annuum* en México

En Mesoamérica, más concretamente en México, el inicio de la domesticación de plantas está registrado arqueológicamente en las cuevas de Ocampo de la Sierra de Tamaulipas, en los yacimientos del Valle de Tehuacán, Puebla y en la Cueva de Guilá Naquitz,

Oaxaca. El chile ha sido registrado en las tres regiones y los restos más antiguos se han encontrado en Tehuacán, en la fase El Riego, fechada entre 7, 000 y 5, 000 antes del presente. Por lo tanto, el chile es una de las primeras plantas domesticadas en México y Mesoamérica (Nuez, *et al.*, 2003; Perry y Flannery, 2007; Long-Solís, 2012; Pickersgill, 2016).

La domesticación y diversificación del chile probablemente fue el resultado de una selección humana consciente sobre las características del fruto, tales como la presencia, ausencia o grado de picor; la variación en la forma del fruto y el color del fruto maduro o inmaduro para su posterior uso fresco o seco. Muchas de estas variantes son controladas por un solo gen (Pickersgill, 2016). Por otro lado, se cree que una de las primeras modificaciones del hombre al domesticar la planta pudo haber sido dificultar su dispersión natural, ya que las hojas en los chiles cultivados esconden el fruto de las aves (Long-Solís, 2012).

La diversificación de *C. annuum* en México se ha estudiado con varios marcadores moleculares. Análisis con isoenzimas y RAPDs reportan que la domesticación de *C. annuum* ha tomado diferentes direcciones (Hernández-Verdugo, *et al.*, 2001; Oyama, *et al.*, 2006). Uno de los primeros estudios con isoenzimas propone dos centros de domesticación del chile. El primero, al este de México, en el norte de Nuevo León y Tamaulipas, hasta el sur de Hidalgo y Veracruz. El segundo, al oeste-centro de México en el estado de Nayarit (Loaiza Figueroa, *et al.*, 1989). Por otro lado, un estudio con microsatélites (SSRs) y datos arqueobotánicos, paleolingüísticos y ecológicos propone a la región centro-este de México como la zona de domesticación incipiente de *C. annuum*. Esta área abarca desde el sur de Puebla y norte de Oaxaca al sur de Veracruz, rodeando la zona del Valle de Tehuacán (Kraft, *et al.*, 2014). Un tercer centro fue propuesto para

Yucatán mediante el análisis con microsatélites y secuencias de ADN (Aguilar-Meléndez, *et al.*, 2009; González-Jara, *et al.*, 2011).

Las poblaciones silvestres y domesticadas de *C. annuum* mantienen la mayor parte de la variación genética dentro y entre poblaciones (Oyama, *et al.*, 2006). Esta diversidad y estructura genética del chile están influenciadas por su distribución geográfica, sistemas de polinización, efectos fundadores, cuellos de botella y selección artificial actuando en diferentes direcciones (Loaiza-Figueroa, *et al.*, 1989; Pickersgill, 1997; Hernández-Verdugo, *et al.*, 2001; Oyama, *et al.*, 2006).

Las poblaciones de chiles silvestres y arvenses se han propuesto como reservorios genéticos (Hernández-Verdugo, *et al.*, 2001; Aguilar-Meléndez, *et al.*, 2009; González-Jara, *et al.*, 2011). Sin embargo, se ha observado que en las poblaciones silvestres existe una disminución de diversidad genética en comparación con las arvenses que mantienen o presentan una mayor diversidad (Aguilar-Meléndez, *et al.*, 2009; González-Jara, *et al.*, 2011). Esto sugiere que los sistemas de manejo en los que se encuentran las poblaciones arvenses deben ser tomados como sitios prioritarios de preservación genética, tomando en consideración que las poblaciones silvestres se ven afectadas por la fragmentación de hábitat y la deforestación de bosques (Hernández-Verdugo, *et al.*, 2001; Aguilar-Meléndez, *et al.*, 2009; González-Jara, *et al.*, 2011). Por lo tanto, los estudios de diversidad genética en poblaciones de chile que involucren el manejo y las prácticas de los agricultores son necesarios en la preservación de los recursos genéticos del chile en México, ante el panorama del incremento en la introducción de variedades mejoradas (tecnificadas) de muchas plantas cultivadas y la decisión de algunos agricultores de mantener sus variedades locales (Pickersgill, 2016).

II.3.3. *C. annuum* en Oaxaca

El estado de Oaxaca es conocido por su gran diversidad biológica y cultural. Es una de las regiones de México donde la interacción del hombre y la planta es más antigua y continúa hoy en día, alcanzando un alto grado de complejidad. Se le considera un área crítica de origen de la agricultura y domesticación de plantas en Mesoamérica como el maíz, el frijol, la calabaza y el chile al encontrarse registro arqueológico de éstas en la Cueva Guilá Naquitz de Ocampo, Oaxaca (Caballero, *et al.*, 2004; Long-Solís, 2012).

Oaxaca cuenta con alrededor de 20 variedades locales de chile *C. annuum* presentes en diversos sistemas de manejo que van desde monocultivos comerciales, traspatios, huertos familiares y milpas, a poblaciones silvestres localizadas alrededor de campos de cultivo, remanentes de bosques y caminos. Entre los chiles más conocidos en el estado podemos encontrar al chile de Agua, Tusta, Taviche, Huacle, Serrano, Paradito, Achilito, Costeño, Guiña, Mirasol, Nanchita y Soledad; y los chiles del tipo silvestre o arvense como Bolita, Piquín, Chilegole, De monte, Chiltepic o Pijita (Aguilar-Rincón, *et al.*, 2010). En la Fig. 2 y en el Anexo 1 se puede observar algunas de las variedades de chile *C. annuum* que podemos encontrar en el estado de Oaxaca.

La mayoría de los estudios de *C. annuum* en Oaxaca se han enfocado en su diversidad morfológica y de usos, por lo que es relevante conocer también cómo se encuentra distribuida la diversidad genética de sus poblaciones y los procesos que la conforman, ya que está sometido a distintos procesos de domesticación, diferentes sistemas de manejo humano bajo condiciones ambientales y topográficas diversas.



Figura 2. Algunos tipos de chiles *C. annuum* encontrados en el estado de Oaxaca. a) Morfología del chile silvestre, chile de monte, bolita, chilegole, piquín o arvense localizados en remanentes de bosque tropical caducifolio, alrededor de los campos de cultivo, huertos y traspatios. b) Chile paradito, se encuentra en traspatios o huertos familiares. c) y d) Chile Guiña Danhi del Istmo de Tehuantepec, Oaxaca, cultivado en línea en asociación a otras plantas como el tomate, el maíz y el camote. Se puede observar el cambio en el tamaño y la posición del fruto. Fotos de Ana Laura Pérez.

III. JUSTIFICACION

Capsicum annuum representa un sistema modelo ideal para abordar los patrones de variación genética asociados al proceso de domesticación, pues se le encuentra en diversos sistemas de manejo, incluyendo poblaciones silvestres (*C. annuum* var. *glabriusculum*) en zonas de bosque tropical caducifolio, traspatios, sistemas milpa y monocultivos comerciales. Lo anterior es notable en el estado de Oaxaca, pues probablemente sea el estado con mayor diversidad para este cultivo con 20 variedades locales de chile (Aguilar-Rincón, *et al.*, 2010).

Los estudios de diversidad genética sobre esta especie en el estado podrán aportar información sobre el nivel y tipo de diversidad genética de poblaciones de chile *C. annuum* que conservan los agricultores en sus parcelas y de las poblaciones silvestres especialmente en lo referente a huellas genéticas del proceso de domesticación del chile en Oaxaca.

IV. HIPOTESIS

- I. En el proceso de domesticación el manejo o las prácticas de los agricultores reducen el nivel de diversidad genética de las especies cultivadas (Engels, *et al.*, 2006; Gross y Olsen, 2010; Tang, *et al.*, 2010). Las poblaciones de Chile *Capsicum annuum* muestran una reducción de su diversidad genética cuando pasan de poblaciones silvestres a poblaciones cultivadas (Pickergill, 1971; 1997). Se espera encontrar menores niveles de diversidad genética en poblaciones cultivadas de Chile (*C. annuum*) y mayor diversidad en poblaciones donde el manejo humano es menor.

- II. Se asume que las plantas arvenses se encuentran en un proceso de domesticación incipiente empezando a reducir sus niveles de diversidad genética en relación a su pariente silvestre pero todavía con fenotipos muy similares a este (Clement, 1999; 2010; Mapes y Basurto, 2016). En *C. annuum*, las poblaciones arvenses son fenotípicamente similares a las poblaciones silvestres (Aguilar-Meléndez, *et al.*, 2009; González-Jara, *et al.*, 2011). En este contexto, se espera que las poblaciones arvenses de *C. annuum* sean genéticamente más similares a las poblaciones silvestres que a las domesticadas, pero con niveles de variación genética ligeramente menores.

V. OBJETIVOS

V.1. Objetivo General

Analizar la diversidad genética de poblaciones de chile (*Capsicum annuum*) en el estado de Oaxaca y evaluar si existen patrones de diversidad genética según el grado de domesticación que presentan (poblaciones silvestres, poblaciones arvenses y poblaciones de variedades cultivadas localmente).

V.2. Objetivos Particulares

- Analizar la diversidad genética de poblaciones de chile del estado de Oaxaca mediante la amplificación de microsatélites nucleares.
- Describir la estructura genética de las poblaciones de chile.
- Correlacionar los resultados obtenidos con los grados de domesticación que presentan las poblaciones (poblaciones silvestres, poblaciones arvenses y poblaciones de variedades cultivadas localmente).

VI. METODOLOGIA

VI.1. Sitio de estudio

El estado de Oaxaca se localiza al sur de México y en las zonas tropicales del continente americano. Cuenta con 26 tipos de vegetación que van desde praderas hasta bosques y selvas. Oaxaca se localiza entre las coordenadas extremas: al norte 18° 39´ y al sur 15° 39´ de latitud norte; al este 93° 52´ y al oeste 98° 32" de latitud oeste. Cuenta con una superficie de 95, 364 km² (representando el 4.8% del territorio nacional) y una regionalización económica que reconoce ocho regiones: Cañada, Costa, Istmo, Mixteca, Papaloapan, Sierra Norte, Sierra Sur y Valles Centrales (García-Mendoza, *et al.*, 2004).

VI.2. Obtención de las muestras

Se colectaron muestras de tejido vegetal de variedades locales, poblaciones arvenses y plantas silvestres de *Capsicum annuum* provenientes de 32 localidades del estado de Oaxaca. Estas localidades se ubicaron a lo largo de dos transectos, uno que recorre la región de la Costa de Oaxaca y otro que atraviesa el estado de norte a sur (los Valles Centrales y la Cañada) (ver Fig. 3). En el Tabla 1 se puede consultar la información de las localidades y en las Tablas 2 y 3 la variedad de chile colectada en cada localidad.

La colecta consistió principalmente de tejido de hoja. Sin embargo, para algunas localidades se colectaron frutos. Las semillas de los frutos fueron germinadas con el fin de obtener o aumentar el tejido de hoja y realizar posteriormente su extracción de ADN. Estas semillas se germinaron en un invernadero del Instituto de Ecología de la UNAM con las condiciones de temperatura y humedad necesarias para su crecimiento.

Tabla 1. Información de las localidades muestreadas en el estado de Oaxaca.

	Localidad	Clave	Región geográfica	Latitud	Longitud	Altitud	n
1	Abasolo	ABA	Valles Centrales	17.00	-96.59	1578	9
2	Baltazar	BAL	Valles Centrales	16.08	-96.79	1014	5
3	Paraje Coatequillas	COA	Valles Centrales	16.55	-96.72	1471	12
4	Coatecas Altas	COA	Valles Centrales	16.55	-96.68	1531	12
5	Ejutla	EJU	Valles Centrales	16.55	-96.72	1470	17
6	La Labor	LAB	Valles Centrales	16.73	-96.67	1540	8
7	La Lobera	LOB	Valles Centrales	16.94	-96.82	1675	16
8	Juan Gabriel Bautista Guelache	JGB	Valles Centrales	17.20	-97.00	1579	6
9	San Pablo Coatlán	SPA	Valles Centrales	16.23	-96.78	1451	4
10	San Cruz Mixtepec	MIX	Valles Centrales	16.79	-96.88	1543	10
11	Santa Cruz Nexila	SCR	Valles Centrales	16.64	-96.90	1461	10
12	Villa de ETLA	ETL	Valles Centrales	17.21	-96.79	1671	11
13	Zachila	ZAC	Valles Centrales	16.90	-97.00	1514	8
14	Los Cues	CUE	Cañada	18.03	-97.06	776	19
15	Cuicatlán	CUI	Cañada	17.80	-96.96	648	19
16	Añil	AL	Costa	16.28	-98.08	153	13
17	Lagartero	AL	Costa	16.25	-98.07	37	9
18	Copalita	COP	Costa	15.82	-96.06	58	3
19	Coyul	COY	Costa	15.92	-95.81	36	4
20	Huaxpaltepec	HUA	Costa	16.33	-97.93	226	6
21	Los Reyes	REY	Costa	15.82	-96.49	231	11
22	Paso de los Indios	PIN	Costa	15.83	-96.48	204	2
23	Lagunilla	PIN	Costa	15.81	-96.46	253	1
24	El Polvorín	POL	Costa	15.72	-96.45	144	9
25	Juan Diegal	JDI	Costa	15.72	-96.43	115	24
26	Rancho Llano	RAN	Costa	16.40	-95.16	25	19
27	Rosedal	ROS	Costa	15.78	-96.91	23	33
28	El Tomatal	TOM	Costa	15.79	-96.92	29	9
29	Santo Domingo del Ingenio	SDO	Costa	16.60	-94.80	48	13
30	San Lorenzo Tezoluca	TEZ	Costa	16.40	-97.88	254	4
31	Saachilac	TOS	Costa	15.98	-95.53	17	8
32	La Tortolita	TOS	Costa	15.98	-95.61	50	7

n= número de individuos procesados. Altitud (msnm), latitud y longitud (grados decimales)

Tabla 2. Variedades de chile *C. annuum* colectadas en 32 localidades del estado de Oaxaca.

	Localidad (Acrónimo)	Variedades de chile
1	Abasolo (ABA)	Agua
2	Baltazar (BAL)	Tusta
3	Paraje Coatequillas (COA)	Agua
4	Coatecas Altas (COA)	Agua
5	Ejutla (EJU)	Taviche
6	La Labor (LAB)	Agua
7	La Lobera (LOB)	Agua
8	Juan Gabriel Bautista Guelache (JGB)	Nanchita, Paradito, Arvenses y <i>C. frutescens</i>
9	San Pablo Coatlán (SPA)	Taviche
10	San Cruz Mixtepec (MIX)	Agua
11	Santa Cruz Nexila (SCR)	Agua
12	Villa de Etla (ETL)	Árbol, Nanchita, Paradito y Agua
13	Zachila (ZAC)	Nanchita, Paradito, Arvenses y <i>C. frutescens</i>
14	Los Cues (CUE)	Árbol y Serrano Criollo
15	Cuicatlán (CUI)	Achilito, Chilcostle Rojo y Huacles
16	Añil (AL)	Costeño Rojo y Costeño Amarillo
17	Lagartero (AL)	Costeño Rojo y Árbol
18	Copalita (COP)	Mirasol y Silvestres
19	Coyul (COY)	Jalapeño y Arvenses
20	Huaxpaltepec (HUA)	Arvenses y <i>C. frutescens</i>
21	Los Reyes (REY)	Costeño Rojo, Tusta y <i>C. frutescens</i>
22	Paso de los Indios (PIN)	<i>C. frutescens</i>
23	Lagunilla (PIN)	<i>C. frutescens</i>
24	El Polvorín (POL)	Mirasol, Arvenses y Silvestres
25	Juan Diegal (JDI)	Tusta y Arvenses
26	Rancho Llano (RAN)	Guiña Danhi y <i>C. frutescens</i>
27	Rosedal (ROS)	Costeño Rojo, Costeño Puya, Arvenses y Silvestres
28	El Tomatal (TOM)	Tusta, Mirasol y Arvenses
29	Santo Domingo del Ingenio (SDO)	Arvenses, Silvestres y <i>C. chinense</i>
30	San Lorenzo Tezoluca (TEZ)	Tusta
31	Saachilac (TOS)	Mareño y Silvestres
32	La Tortolita (TOS)	Guajillo, Arvenses y <i>C. chinense</i>

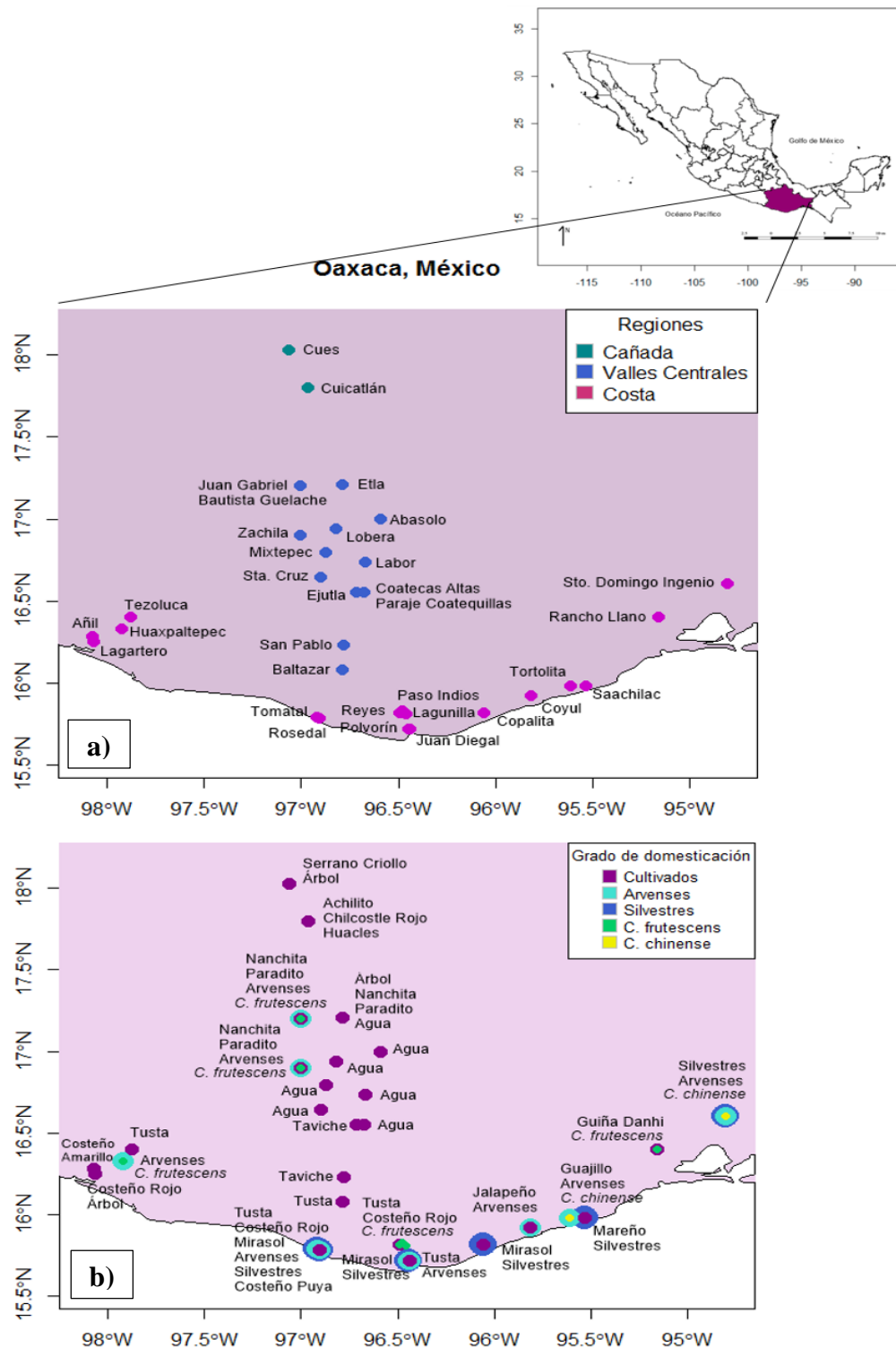


Figura 3. Localidades muestreadas en el estado de Oaxaca. a) Dos transectos fueron trazados, uno a lo largo de la costa de Oaxaca (puntos rosas) y otro que cruzó el estado de norte a sur (puntos verdes y azules). b) Cada localidad presenta de una hasta cuatro variedades de chile *C. annuum* en formas cultivadas, arvenses y silvestres.

Tabla 3. Información de las variedades de chile *C. annuum* colectadas en el estado de Oaxaca. Se incluyeron 16 individuos de *Capsicum frutescens* y 3 individuos de *Capsicum chinense* como grupos externos y colectados también en el estado de Oaxaca.

	Variedad	Localidad (Acrónimo)	n
Cultivadas	1 Agua	ETL, ABA, LOB, LAB, MIX, SCR, COA	79
	2 Costeño Rojo	ROS, AL, REY	45
	3 Tusta	TEZ, BAL, TOM, REY, JDI	35
	4 Taviche	EJU, SPA	21
	5 Guiña Danhi	RAN	17
	6 Huacles	CUI	12
	7 Árbol	CUES, AL	12
	8 Paradito	ETL, ZAC JGB	10
	9 Serrano criollo	CUE	9
	10 Nanchita	ETL, ZAC, JGB	8
	11 Chilcostle Rojo	CUI	4
	12 Mirasol	TOM, POL, COP	4
	13 Achilito	CUI	3
	14 Costeño Amarillo	AL	2
	15 Mareño	TOS	2
	16 Costeño Puya	ROS	1
	17 Guajillo	TOS	1
	18 Jalapeño	COY	1
19 Arvenses	JGB, ZAC, HUA, ROS, TOM, JDI, POL, COY, TOS, SDO	33	
20 Silvestres	ROS, TOM, POL, COP, TOS, SDO	23	
Grupos externos			
21 <i>Capsicum frutescens</i>	RAN, REY, ZAC, HUA, JGB, PIN	16	
22 <i>Capsicum chinense</i>	TOS, SDO	3	

VI.3. Extracción de ADN

El ADN de las muestras de *C. annuum* se extrajo siguiendo el protocolo CTAB-miniprep (Doyle y Doyle, 1987), modificado en algunos pasos (ver Anexo 2). El material extraído fue almacenado a -20°C para evitar su degradación.

VI.4. Amplificación de microsatélites (SSRs)

Se realizaron pruebas preliminares de 20 loci de microsatélites para determinar cuáles se utilizarían para el estudio. Los microsatélites con los que se trabajó fueron 13 de los 20 probados (ver Tabla 4). Estos fueron reportados por Shirasawa *et al.*, (2013) y Nagy *et al.*, (2007) y son específicos de *C. annuum*. Los Primers Reverse estaban marcados con fluorescencia.

Tabla 4. Información reportada en la literatura de cada microsatélite. N= n° de alelos.

SSR	Motivo	Tamaño	Cromosoma	N	Colorante	Referencia
EPMS310	(CAT) ₁₃	140-172	8	20	PET	Nagy, <i>et al.</i> , 2007
EPMS386	(CA) ₁₅	122-170	3	18	FAM	Nagy, <i>et al.</i> , 2007
EPMS419	(AAT) ₆	224-248	9	13	VIC	Nagy, <i>et al.</i> , 2007
CaES2332	AAG	160	4	3	PET	Shirasawa, <i>et al.</i> , 2013
CaES4192	ATC	166	5	2	PET	Shirasawa, <i>et al.</i> , 2013
CaES0425	AAC	255	2	2	VIC	Shirasawa, <i>et al.</i> , 2013
CaES3862	GGA	209	1	3	PET	Shirasawa, <i>et al.</i> , 2013
CaES4787	ACT	132	3	3	NED	Shirasawa, <i>et al.</i> , 2013
CaES1003	AATC	176	1	3	VIC	Shirasawa, <i>et al.</i> , 2013
CaES5392	GGA	201	6	3	VIC	Shirasawa, <i>et al.</i> , 2013
CaES4584	AAC	108	7	4	NED	Shirasawa, <i>et al.</i> , 2013
CaES4597	GGA	244	10	4	FAM	Shirasawa, <i>et al.</i> , 2013
CaES3538	AGC	191	4	2	NED	Shirasawa, <i>et al.</i> , 2013

En la reacción de amplificación por PCR se retomó y modificó el protocolo de Shirasawa *et al.* (2013) y Nagy *et al.* (2007). Se realizaron mezclas incluyendo tres o cuatro primers por reacción. Para las mezclas de Shirasawa, *et al.*, (2013), se agregó a cada tubo 3 μ l de ADN (1 ng/25 μ l), 5 μ l de buffer PCR 1X, 3 μ l de MgCl₂ (3 mM), 0.08 μ l de Taq Polimerasa (0.4 μ l/25 μ l), 2 μ l de dNTPs (0.8 mM), 0.75 μ l de cada primer (0.3 μ M) y 7.42 μ l de agua ultra pura para un total de 25 μ l por reacción. Para las mezclas de Nagy *et al.*, (2007), se agregó a cada tubo 3 μ l de ADN (25 ng/25 μ l), 5 μ l de buffer PCR 1X, 2.5 μ l de MgCl₂ (2.5 mM), 0.08 μ l de Taq Polimerasa (0.4 μ l/25 μ l), 2 μ l de dNTPs (0.8 mM), 0.5 μ l de cada primer (0.2 μ M) y 9.42 μ l de agua ultra pura para un total de 25 μ l por reacción.

El programa con el que se realizó la amplificación por PCR para los primers de Nagy, *et al.*, (2007) fue: 94°C por 3 minutos de desnaturalización inicial, 35 ciclos de: 94°C por 30 segundos de desnaturalización, 60°C por 30 segundos de alineamiento, 72°C por 30 segundos para elongación; y una extensión final de 72°C por 2 minutos.

El programa de amplificación de Shirasawa, *et al.*, (2013) retoma el protocolo de Sato *et al.*, (2005), el cual, incluye una modificación conocida como "Touchdown", por lo que el programa se realizó con 94°C por 3 minutos de desnaturalización inicial, tres rondas con una desnaturalización de 94°C por 30 segundos y una temperatura de alineamiento decreciendo 1°C cada 2 ciclos hasta 56°C. La primera ronda constó de 3 ciclos de 94°C por 30 segundos y 68°C por 30 segundos. La segunda ronda fue: 10 ciclos de 94°C por 30 segundos, 66°C cada 30 segundos y 72°C por 30 segundos de elongación. La tercera ronda fue: 19 ciclos de 94°C por 30 segundos, 56°C por 30 segundos y 72°C por 30 segundos de elongación. La extensión final fue de 72°C por 3 minutos.

Se hicieron geles de agarosa al 2% (1.2 g de agarosa/60 ml de buffer TAE 0.5X), al cual se le agregó EpiQuick DNA Stain para visualizar las bandas y comprobar el éxito de la amplificación. Los geles se corrieron en una cámara electroforética y se observaron con un transiluminador UV. Posteriormente, se llenaron placas con las diluciones correspondientes según el patrón de bandeo obtenido en el gel de agarosa. Las placas fueron enviadas a la UIUC CORE SEQUENCING FACILITY de la Universidad de Illinois (<https://unicorn.biotech.illinois.edu/>) para el análisis del tamaño del fragmento mediante el método de capilaridad.

Para las lecturas del tamaño de fragmento obtenido se utilizó el programa Peak Scanner versión 1 (Applied Biosystems) y se generó una matriz de datos en Excel para los análisis de diversidad y estructura genética.

VI.5. Análisis de diversidad genética

De las 32 localidades del estado de Oaxaca que se muestrearon se agrupó a Tortolita y Saachilac; Paso de los Indios y Lagunilla; Coatecas Altas y Paraje Coatequillas; y Añil y Lagartero, dándonos un total de 28 poblaciones. Estas poblaciones fueron agrupadas por su cercanía geográfica (se encontraban a menos de 8.9 km de distancia) y por un bajo número de individuos colectados (menores a 10 individuos), para fines prácticos de los análisis de diversidad genética realizados.

Un total de 341 individuos fueron analizados de las 28 poblaciones: 322 individuos de *C. annuum*, 16 individuos de *C. frutescens* y 3 individuos de *C. chinense*. Los análisis estadísticos se realizaron por población y posteriormente agrupando cada una de las localidades con base en dos criterios: por región geográfica (Costa, Valles Centrales y

Cañada) y por grado de domesticación (cultivado, arvense y silvestre). La Costa se conformó de 14 poblaciones y 175 individuos genotipados; los Valles Centrales de 12 poblaciones y 128 individuos genotipados, y la Cañada de 2 poblaciones y 38 individuos genotipados.

En la evaluación de la diversidad genética según el grado de domesticación, se analizaron 317 individuos de Chile *C. annuum* y se clasificaron en silvestres, arvenses y cultivadas. En el caso de las cultivadas se excluyó a las variedades de Chile Mareño, Costeño Puya, Guajillo y Jalapeño, reduciendo las variedades de Chile de 18 a 14 (ver Tabla 3). Estas se excluyeron por presentar solo un individuo y en el caso de Mareño, uno de sus dos individuos no amplificó para 5 loci de microsatélites.

En la categoría cultivada se agruparon un total de 261 individuos de 14 variedades locales de Chile *C. annuum* (primeras 14 variedades de la Tabla 3), mientras que en la categoría silvestre se agrupó 23 individuos de Chile conocidos localmente como Chile de Monte (*C. annuum* var. *glabriusculum*). En la categoría arvense se agrupó 33 individuos de Chile creciendo como plantas toleradas, en sistemas de manejo humano como traspatios, jardines, milpas o huertos familiares. Estos Chile se observaron fenotípicamente muy similares a las poblaciones silvestres.

Se realizó un primer análisis de diversidad genética por locus de microsatélite con el programa GENEPOP versión 4.6. También, se realizó la prueba exacta de Hardy-Weinberg que está basada en el método de cadenas de Markov con 10, 000 desmemorizaciones, 20 batches y 5, 000 iteraciones por batch y posteriormente la prueba de desequilibrio de ligamiento basada en el mismo método con 10, 000 desmemorizaciones, 100 batches y 5, 000 iteraciones por batch (Rousset, 2008). Los

valores de χ^2 (xi-cuadrada) que se obtienen permiten identificar qué tanto se separa lo observado de lo esperado y si se encuentran o no en equilibrio de Hardy-Weinberg.

Se calculó la frecuencia de alelos nulos y la F_{ST} con intervalos de confianza en el programa FREENA (Chapuis y Estoup, 2007) con el fin de conocer si la presencia de alelos nulos es significativa y si es necesario eliminar algún locus de los análisis. Este programa sigue el algoritmo de máxima expectación (EM) y se realizó con 1, 000 repeticiones.

Por otra parte, se estimó el número de loci polimórficos, el promedio del número de alelos totales y el número de alelos privados (alelos únicos en la población). El promedio del número de alelos efectivos, es decir, los alelos que tienen la capacidad de pasar a la siguiente generación (Kimura y Crow, 1964) fueron calculados mediante el programa R versión 3.4.1 (<https://cran.r-project.org/>), utilizando el paquete GSTUDIO (ver Anexo 3). La riqueza alélica promedio fue calculada con el programa FSTAT versión 2.9.3.2 (Goudet, 2002) por locus y población basándose en el tamaño mínimo de muestra de un individuo diploide, por el método de rarefacción (Petit, *et al.*, 1998). Además, se realizó otro método de rarefacción implementado por el programa ADZE versión 1.0 (Szpiech, *et al.*, 2008) estandarizándolo con el máximo tamaño de muestra.

Los niveles de diversidad genética se estimaron con el programa ARLEQUÍN versión 3.5.2.2. (Excoffier, *et al.*, 2005). La heterocigosis esperada (H_E) es definida como la probabilidad de que dos alelos elegidos al azar sean diferentes en la muestra y fue calculada siguiendo la ecuación de Nei, 1987:

$$H_E = \frac{2n}{2n - 1} \left(1 - \sum_{i=1}^k p_i^2 \right)$$

donde n es el número de copia de genes en la muestra, p_i es la frecuencia del alelo i en la muestra y K el número de alelos.

Para conocer si los valores de H_E y H_o son significativamente diferentes entre grupos se realizó una prueba ANOVA con el programa R versión 3.4.1 utilizando el paquete RPSYCHI.

VI.6. Análisis de varianza molecular

Los estadísticos F de Wright (F_{IS} , F_{IT} y F_{ST}) y la distribución de la diversidad genética entre y dentro de las poblaciones fueron calculados mediante el análisis de varianza molecular (AMOVA) con el método de Weir y Cockerham (1984). Este análisis se realizó en el programa ARLEQUÍN versión 3.5.2.2 con una significancia de 0.05, 1, 000 permutaciones y una distancia molecular R_{ST} (Slatkin, 1995; Excoffier, *et al.*, 2005).

La R_{ST} fue propuesta por Slatkin, (1995) como una medida de estructuración genética basada en el modelo de mutación paso a paso (S.M.M.), el cual siguen los microsatélites. Esta medida es análoga a la F_{ST} , sin embargo, la R_{ST} se puede calcular a partir de las varianzas en el tamaño de los alelos, mientras que la F_{ST} , típicamente es derivada de las varianzas en las frecuencias alélicas. La R_{ST} es definida como:

$$R_{ST} = \frac{S - S_W}{S}$$

donde S es el promedio de las diferencias cuadradas en el tamaño del alelo entre todos los pares de alelos, y S_W , el promedio de la suma de las diferencias cuadradas en el tamaño del alelo en cada subpoblación (Slatkin, 1995; Balloux y Lugon-Moulin, 2002).

Dos análisis AMOVA se realizaron: uno excluyendo a individuos de *C. frutescens* y *C. chinense* de las poblaciones, y otro que los incluía. Esto con el fin de conocer si estos individuos afectan los niveles de diferenciación poblacional F_{ST} .

Para conocer si las diferencias en los valores individuales de F_{IS} son significativas se realizó una prueba ANOVA en XLSTAT PLUS 2017 para Excel (Microsoft).

VI.7. Distancias genéticas y aislamiento por distancia

VI.7.1. Distancias genéticas

Mediante el programa POPULATIONS versión 1.2.32 (Langella, 1999), se obtuvo un dendograma de Neighbor-Joining para las poblaciones y variedades basado en las distancias $\delta\mu^2$ (Goldstein, *et al.*, 1995). El método sigue la ecuación:

$$(\delta\mu)^2 = (m_x - m_y)^2$$

donde m_x y m_y son los promedios de los tamaños de los alelos en la población x y y respectivamente (Cooper, *et al.*, 1999). Estas distancias fueron desarrolladas específicamente para marcadores de microsatélites bajo el mismo modelo de mutación paso a paso en el cual se basa la R_{ST} . Se calcularon los valores de soporte estadístico de cada nodo utilizando 1, 000 repeticiones.

En el dendograma por población se omitieron los individuos pertenecientes a *C. frutescens* y *C. chinense* (ver Tabla 2). Los individuos pertenecientes a Paso de los Indios (PIN) corresponden a *C. frutescens*, por lo que fue considerada como el grupo externo. En el caso de las variedades se analizaron solo 8 loci de microsatélites debido a que 2 loci (CaES3862 y CaES4584) no amplificaron para Costeño Amarillo, Mirasol y Serrano

Criollo. Estas distancias son más sensibles al tamaño de muestra (Takezaki y Nei, 1996; Cooper, *et al.*, 1999). Se utilizó a *C. frutescens* como grupo externo.

En el mismo programa se realizó otro dendograma de Neighbor-Joining para las poblaciones y variedades esta vez basado en las distancias D_A (Nei, *et al.*, 1983) reportada como una distancia más adecuada en el análisis de microsatélites (Takezaki y Nei, 1996; 2008). El método sigue el modelo de alelos infinitos (I.A.M. por sus siglas en inglés) y la ecuación:

$$D_A = 1 - \frac{1}{r} \sum_j^r \sum_i^{m_j} \sqrt{x_{ij}y_{ij}},$$

donde x_{ij} son las frecuencias del alelos i en el locus j en la población X y Y , respectivamente, y m_j es el número de alelos en el locus j (Takezaki y Nei, 1996; 2008). Se calcularon los valores de soporte estadístico de cada nodo utilizando 1, 000 repeticiones.

La única variante para el dendograma por D_A fue para el caso de las variedades donde se analizaron los 10 loci de microsatélites en lugar de los ocho utilizados para el dendograma por $\delta\mu^2$. Se visualizaron los árboles mediante el programa TREEVIEW versión 1.6.6 (Page, 1996).

VI.7.2. Aislamiento por distancia

Para conocer si existe aislamiento por distancia en las poblaciones se realizó una prueba de Mantel con 10, 000 permutaciones. Esta prueba fue realizada en XLSTAT PLUS 2017 para Excel (Microsoft). Las distancias geográficas fueron obtenidas con el programa R

versión 3.4.1 utilizando el paquete GGMAP (ver Anexo 3). Esta prueba busca una correlación entre una matriz de distancias geográficas y una matriz de distancias genéticas (R_{ST} pareadas). Si existe una correlación significativa entre las distancias geográficas y genéticas, hay aislamiento por distancia (Slatkin, 1995; Piñero, *et al.*, 2008b).

Se realizaron tres pruebas de Mantel adicionales. La primera comparó la distancia genética (R_{ST} pareadas) y geográfica de la región de la Costa. La segunda comparó el transecto norte-sur que incluye a Valles Centrales y Cañada. Estas pruebas se realizaron con el fin de conocer si existe una correlación genética y geográfica a nivel de transecto. Las pruebas de Mantel a nivel geográfico se realizaron con y sin individuos de *C. frutescens* y *C. chinense*

La tercera prueba de Mantel se realizó entre la distancia genética (R_{ST} pareadas) de las variedades de *C. annuum* y su distancia de domesticación. La distancia de domesticación fue asignada de 3 para cultivada, 2 para arvense y 1 para silvestre. En la matriz de comparaciones pareadas entre variedades se realizaba una resta según el número asignado. Por ejemplo, dos variedades cultivadas quedaron a una distancia igual a 0 ($3-3= 0$), y una arvense y cultivada a una distancia de 1 ($3-2= 1$). Esta prueba se realizó para conocer si existe una correlación entre la distancia genética y la distancia de domesticación.

VI.7.3. Barreras geográficas

Con la finalidad de identificar barreras geográficas entre las poblaciones analizadas se aplicó un algoritmo conocido como Monmonier, el cual se basa en la conectividad de

redes (triangulación Delaunay) entre poblaciones o individuos, construyendo las barreras con base en las distancias mayores entre los nodos de la red (Monmonier, 1973; Garrido-Garduño y Vázquez-Domínguez, 2013). Este análisis se realizó en el programa R versión 3.4.1 utilizando los paquetes ADEGENET (ver Anexo 3) con una matriz de distancias genéticas R_{ST} pareadas (Slatkin, 1995), y una base de datos geográficos (latitud y longitud) de las localidades muestreadas.

Con el fin de definir grupos de poblaciones genéticamente homogéneas incorporando la información geográfica e identificar barreras significativas al flujo entre grupos, se realizó un análisis espacial de varianza molecular (SAMOVA) con el programa SAMOVA versión 2.0 (Dupanloup, *et al.*, 2002). Este método identifica el agrupamiento de poblaciones que maximiza la proporción de la varianza genética total (F_{CT}) debido a las diferencias entre grupos de poblaciones geográficamente cercanos. Para el análisis se utilizaron los parámetros determinados por el programa con una K entre 2 y 15 y un método de distancia molecular basada en R_{ST} . Posteriormente, con la estructura obtenida, se realizó la prueba AMOVA en ARLEQUÍN versión 3.5.2.2 (Excoffier, *et al.*, 2005) con 1, 000 permutaciones y un método de distancia molecular R_{ST} con el fin de encontrar la K y la estructura genética que maximizara F_{CT} . Ambos análisis se realizaron con y sin individuos pertenecientes a *C. frutescens* y *C. chinense*.

VI.8. Estructura poblacional

VI.8.1. Análisis multivariado

Se realizó un análisis discriminante de componentes principales (DAPC, por sus siglas en inglés), el cual es un método que describe la diferenciación genética entre los grupos

partiendo de un análisis de componentes principales previo a un análisis discriminante (Jombart, *et al.*, 2010). El análisis fue realizado a partir de los grupos formados por población, por región geográfica, por variedad de Chile *C. annuum* y por categoría de domesticación en el programa R versión 3.4.1 utilizando el paquete ADEGENET (ver Anexo 3). En el caso de región geográfica y población los análisis se realizaron con y sin individuos pertenecientes a *C. frutescens* y *C. chinense*.

VI.8.2. Análisis bayesiano

La estructura genética de las poblaciones de Chile se infirió mediante un análisis bayesiano de estructura poblacional con el programa STRUCTURE versión 2.3.4 (Pritchard, *et al.*, 2000). Este análisis se basa en un modelo de agrupamiento de datos genotípicos para inferir la estructura poblacional y asignar individuos a las poblaciones. El método asume que hay K poblaciones (puede ser desconocida), las cuales son caracterizadas por un conjunto de frecuencias alélicas para cada locus. Así, los individuos de la muestra son asignados probabilísticamente a las poblaciones, o, conjuntamente, a dos o más poblaciones, si sus genotipos indican admixtura (Pritchard, *et al.*, 2000).

Para el análisis se utilizó un modelo de admixtura, suponiendo correlación entre frecuencias alélicas e infiriendo el valor de α . Se utilizó una K de 1 a 10 (número de grupos inferidos) con 10 repeticiones para cada valor de K . En cada corrida se utilizaron 250,000 *burn-in chains* y 1,000,000 cadenas de Markov tipo Monte Carlo. Para conocer el valor de K más óptimo, se realizó la prueba de Evanno *et al.*, (2005) con STRUCTURE HARVESTER (Earl y von Holdt, 2012) disponible en: <http://taylor0.biology.ucla.edu/structureHarvester/>.

Independientemente del resultado de la prueba de Evanno *et al.*, (2005) en los gráficos se colocaron los resultados obtenidos para diferentes valores de K . Se ha reportado que cuando el muestreo no es uniforme en poblaciones o en niveles jerárquicos de la estructura de la población, el análisis posterior al STRUCTURE (por ejemplo, la prueba de Evanno) puede ser inexacto e identificar un número incorrecto de grupos (Gilbert, 2016) o bien otros valores de K pueden reflejarnos distintos procesos demográficos y por lo tanto garantizar la interpretación de los datos (Meirmans, 2015).

Los resultados gráficos obtenidos de STRUCTURE fueron editados y modificados en el programa R versión 3.4.1 utilizando los paquetes GGLOT2 y RESHAPE2.

VI.8.3. Análisis gráfico poblacional

Las gráficas poblacionales son una representación teórica-gráfica de la covarianza genética multilocus, que describe la forma en la cual la variación genética está distribuida en el paisaje y cómo las características espaciales, ecológicas y topográficas de dicho paisaje, influyen en la conectividad y la estructura genética. Además, nos permiten inferir el aislamiento por distancia y el flujo génico en las poblaciones (Dyer, 2004; 2015).

En el gráfico de redes que genera este análisis las poblaciones son representadas como nodos unidos por aristas. El tamaño del nodo es proporcional al componente de la varianza genética dentro de las poblaciones dándole volumen a partir del centroide de la población. Estos centroides son conectados por una arista (medidas de covarianza) cuya longitud se define por la variación entre todos los pares de poblaciones (Dyer, 2004; 2015).

Los gráficos poblacionales ofrecen una forma diferente de conocer si existe aislamiento por distancia (IBD) entre las poblaciones analizadas. La prueba no utiliza las distancias genéticas tradicionales F_{ST} o R_{ST} , sino que utiliza la distancia genética condicional derivada de las redes poblacionales (*cGD*). Esta distancia es estimada como la longitud de la más corta trayectoria (geodésica) que conecta los pares de poblaciones, y se basa en el cálculo de las diferencias de la covarianza genética asociada con la conectividad directa e indirecta (flujo de genes) entre poblaciones, haciéndolo adecuado para su uso en el modelado genético del paisaje (Dyer *et al.*, 2010).

En este análisis se utilizaron 26 poblaciones y 12 variedades cultivadas, la categoría arvense y la silvestres de Chile *C. annuum*. Las poblaciones PIN y COP y las variedades Costeño Amarillo y Achilito fueron omitidas al presentar menos de 4 individuos. En el caso de las variedades, este análisis solo fue representado con el fin de conocer su covarianza, ya que una variedad de Chile *C. annuum*, puede pertenecer a una o más de las localidades muestreadas (ver Tablas 2 y 3).

Se realizó el gráfico de poblaciones en el programa R versión 3.4.1 utilizando los paquetes POPGRAPH, GSTUDIO e IGRAPH (ver Anexo 3). Esta prueba utiliza una base de datos genéticos y una base de datos geográficos. Posteriormente, se realizó una gráfica de aislamiento por distancia, la cual utiliza una matriz de distancia geográficas de cada población y una matriz de distancia genética condicional (*cGD* o distancia gráfica). Un gráfico adicional fue realizado con la distancia altitudinal (m) con el fin de conocer si existe aislamiento por altitud. Las pruebas de Mantel (correlación entre las matrices de distancia) se realizaron en el programa R versión 3.4.1 con el paquete ADE4 (Ver Anexo 3).

VI.8.4. Cuellos de botella y tamaño efectivo poblacional

Conocer si el tamaño de las poblaciones se expande o decrece permite obtener información acerca de la historia demográfica de los organismos y sus trayectorias futuras; así la existencia de cuellos de botella o una reducción en el tamaño efectivo N_e conduciría a un exceso de alelos comunes en microsatélites comparados a alelos raros que podrían esperarse bajo equilibrio (Putman y Carbone, 2014).

Para determinar reducciones en el tamaño efectivo de las poblaciones se calculó en el programa ARLEQUÍN versión 3.5.2.2 (Excoffier, *et al.*, 2005) el índice de Garza y Williamson, 2001 (M). Este índice está basado en la relación entre el número de alelos de microsatélites (K) y el rango en el tamaño del alelo (r , es decir, $M= K/r$), valores $M < 0.68$ indican una reducción en el tamaño poblacional (Excoffier, *et al.*, 2005; Peery, *et al.*, 2012). Durante un cuello de botella, se espera que el número de alelos disminuya más rápido que el rango en el tamaño del alelo, ya que la mayoría de los alelos perdidos serán, por azar, de tamaño intermedio para loci con al menos cinco alelos. En consecuencia, se espera que la relación M sea más pequeña en las poblaciones con cuello de botella que en las poblaciones en equilibrio (Peery, *et al.*, 2012).

Adicionalmente se realizó un análisis con el programa BOTTLENECK versión 1.2.02 (Piry, *et al.*, 1999). El programa incluye tres modelos de mutación de microsatélites: modelo de alelos infinitos (I.A.M., Maruyama y Fuerst, 1985), modelo de mutación paso a paso (S.M.M., Cornuet y Luikart, 1996) y modelo de mutación dos fases (T.P.M., Di Rienzo, *et al.*, 1994). Además, permite realizar el análisis mediante tres pruebas estadísticas: de significancia, de Wilcoxon y de diferencia estandarizada (Cornuet y Luikart, 1996; Luikart, *et al.*, 1998).

En este análisis se agruparon a las poblaciones en dos grupos, siguiendo los dos transectos principales de muestreo con el fin de obtener un análisis más robusto ya que 13 poblaciones presentan menos de 10 individuos. El primer grupo consiste de las poblaciones de la Costa (con 175 individuos), y el segundo, de las poblaciones de la Cañada y los Valles Centrales (con 166 individuos). En el caso de las variedades, se agruparon según su grado de domesticación en poblaciones cultivadas, arvenses y silvestres (con 261, 33 y 23 individuos respectivamente) con el fin de conocer si existen desviaciones significativas del equilibrio mutación-deriva en el gradiente de domesticación de *C. annuum*. Se utilizaron los tres modelos mutacionales y las pruebas estadísticas de significancia y Wilcoxon. La varianza utilizada en el modelo T.P.M. fue de 5% y la probabilidad para I.A.M. y S.M.M., fue de 95% con 1, 000 repeticiones.

Un análisis complementario sobre la estimación del tamaño efectivo de las poblaciones (N_e) se realizó con el programa NeEstimator versión 2.01 (Do, *et al.*, 2014), mediante el método de desequilibrio de ligamiento, asumiendo apareamientos al azar y valores críticos $P_c = 0.02$. Con el fin de obtener un análisis más robusto se agruparon a las poblaciones en dos grupos: las poblaciones de la Costa (con 175 individuos), y las poblaciones de la Cañada y los Valles Centrales (con 166 individuos). En el caso de las variedades, se agruparon según su grado de domesticación en poblaciones cultivadas, arvenses y silvestres (con 261, 33 y 23 individuos respectivamente).

VII. RESULTADOS

VII.1. Amplificación y análisis de microsatélites (SSRs)

Se amplificaron un total de 13 microsatélites. Sin embargo, para los análisis de diversidad solo fueron utilizados 10 debido a que CaES3538, CaES4597 y EPMS419 no amplificaron para 20 poblaciones. Los 10 loci de microsatélites analizados corresponden a CaES1003, CaES4787, CaES3862, CaES2332, CaES0425, CaES4192, CaES5392, CaES4584, de Shirasawa *et al.*, (2013); y EPMS386 y EPMS310, de Nagy *et al.*, (2007).

El locus EPMS386 fue polimórfico para las 28 poblaciones analizadas, mientras que el CaES4584 fue polimórfico solo para 6 poblaciones. El número de alelos totales se encontró en un intervalo de 3 a 21 con un promedio de 10. La heterocigosis esperada fue mayor en EPMS386 con 0.66 y menor en CaES4584 con 0.05 y los niveles de F_{IS} se encontraron en un intervalo de 0.001 a 0.99, sugiriendo niveles altos de endogamia (ver Tabla 5). En total se encontraron 36 alelos privados siendo EPMS386 el locus con mayor número de alelos (11 alelos) mientras que CaES2332 y CaES4192 no presentaron alelos privados.

Los resultados de la prueba exacta de Hardy-Weingberg por locus determinaron que los 10 loci de microsatélite presentan una desviación significativa del equilibrio ($p < 0.05$), posiblemente por los valores de F_{IS} altos y positivos (Tabla 5).

Tabla 5. Resultados del análisis de diversidad y la prueba exacta de Hardy-Weinberg de 10 microsatélites (SSRs) en 28 poblaciones del estado de Oaxaca. C= número de cromosoma, Na= n° de alelos, Ne= promedio del número de alelos efectivos, A= riqueza alélica promedio, Np= alelos privados, χ^2 = xi-cuadrada y g.l.= grados de libertad.

Nombre	C	Na	Ne	A	Np	H_o	H_E	F_{IS}	F_{ST}	F_{IT}	χ^2	g.l	p
CaES3862	1	3	1.4	1.19	2	0.12	0.18	0.35	0.07	0.40	52.36	26	0.0016*
CaES1003	1	4	1.1	1.07	2	0.03	0.07	0.60	0.03	0.61	31.23	10	0.0005*
CaES0425	2	4	1.5	1.28	4	0.02	0.25	0.99	-0.05	0.99	182.38	30	0.0000*
EPMS310	3	14	2.9	1.85	1	0.24	0.59	0.67	0.37	0.79	l	46	0.0000*
CaES4787	3	6	1.2	1.12	4	0.09	0.11	0.65	0.01	0.65	34.93	18	0.0090*
CaES2332	4	5	2.2	1.53	0	0.29	0.44	0.33	0.16	0.44	103.66	46	0*
CaES4192	5	3	1.7	1.52	0	0.11	0.33	0.62	0.44	0.79	211.84	44	0.0000*
CaES5392	6	8	1.2	1.08	8	0.05	0.07	0.001	0.06	0.06	37.20	14	0.0007*
CaES4584	7	4	1.5	1.06	4	0.03	0.05	0.55	-0.004	0.55	20.91	8	0.0074*
EPMS386	8	21	3.2	1.89	11	0.26	0.66	0.71	0.44	0.84	l	56	0.0000*

*valores significativos ($p < 0.05$), l= valores altos (infinitos)

La presencia de alelos nulos solo se detectó en los loci EPMS386 y EPMS310 con frecuencias de 0.23 y 0.20 respectivamente (ver Anexo 4). Se ha indicado que frecuencias mayores a 0.20 deben considerarse en la eliminación del locus. Sin embargo, la prueba de corrección ENA indicó que la presencia de alelos nulos no es un factor que disminuya la confianza de los resultados obtenidos, ya que los valores F_{ST} sin ENA (0.24) y F_{ST} con ENA (0.21) no son significativos y se encuentran dentro de los intervalos de confianza de 0.13 a 0.29 para F_{ST} sin ENA y de 0.15 a 0.26 para F_{ST} con ENA.

El análisis de desequilibrio de ligamiento detectó que cinco pares de loci se encuentran ligados (ver Anexo 5), estos pares se conforman por CaES4787 y CaES1003 ($p = 0.001$), CaES4787 y CaES3862 ($p = 0.013$), CaES1003 y CaES3862 ($p = 0.001$), CaES4584 y CaES5392 ($p = 0.004$) y EPMS386 y EPMS310 ($p = 0.002$). Los loci CaES1003 y CaES3862 fueron los únicos que se localizan en el mismo cromosoma y se encuentran ligados.

VII.2. Análisis poblacional

VII.2.1. Diversidad genética

Todas las poblaciones presentaron loci polimórficos excepto Ejutla (EJU) y Tortolita-Saachilac (TOS). El promedio de loci polimórficos fue 2.86 ($SD= 1.82$). Por otra parte, se detectaron 10 alelos privados correspondientes a los loci CaES3862, CaES0425 y EPMS386 y CaES5392 con 1, 2, 3 y 4 alelos respectivamente (ver Tabla 6).

La heterocigosis esperada (H_E) fue mayor a la observada (H_O) en todas las poblaciones analizadas, con una heterocigosis total (H_T) de 0.36 ($SD= 0.32$) (ver Tabla 6). La diversidad genética fue mayor en la región de la Costa con una H_O de 0.12 ($SD= 0.16$) y una H_E de 0.28 ($SD= 0.27$). Los niveles de diversidad genética fueron menores en los Valles Centrales con una $H_O = 0.09$ ($SD= 0.14$) y una $H_E = 0.21$ ($SD= 0.25$), mientras que en la Cañada los niveles de diversidad fueron de $H_O = 0.11$ ($SD=0.15$) y H_E de 0.24 ($SD= 0.27$). Las poblaciones Añi-Lagartero, Coatecas y Cues presentaron las heterocigosis más altas $H_E > 0.25$ en cada región geográfica (Costa, Valles Centrales y Cañada), mientras que Tezoluca, Cuicatlán y Santa Cruz presentaron los valores más bajos $H_E < 0.22$, en sus respectivas regiones (ver Tabla 6).

Por medio de una prueba ANOVA se determinó que las diferencias en el promedio de H_E y H_O por región geográfica no son significativas. Para H_E se obtuvo un valor de $F= 0.21$ y $p=0.81$ mientras que para la H_O los valores fueron de $F= 0.12$ y $p= 0.89$ (ver Anexo 6).

Tabla 6. Heterocigosis observada (H_o) y esperada (H_E) promedio en cada población. Los resultados son solo de loci polimórficos. n = número de individuos en la población, nv = número de variedades de *C. annuum* por población, N = promedio del número de alelos, N_e = promedio del número de alelos efectivos, A = riqueza alélica promedio, N_p = alelos privados, P = número de loci polimórficos, M = índice Garza y Williamson, 2001 y F_{IS} = Coeficiente de endogamia. Significancia igual a 0.05.

Población (Clave)	n	nv	P	N	Ne	A	Np	H_o (SD)	H_E (SD)	M	F_{IS}
Costa											
AL	22	3 ^C	5	4.0	2.1	1.40	3	0.24 (0.20)	0.40 (0.22)	0.30	0.78
RAN	19	1 ^{CF}	3	2.8	2.2	1.34	1	0.15 (0.17)	0.34 (0.28)	0.28	0.15
COY	4	2 ^{CA}	6	2.0	1.8	1.34	0	0.07 (0.11)	0.34 (0.28)	0.35	0.003
ROS	33	4 ^{CAS}	4	3.9	2.6	1.33	0	0.27 (0.27)	0.33 (0.33)	0.27	0.72
POL	9	3 ^{CAS}	4	2.2	1.8	1.32	0	0.16 (0.25)	0.32 (0.29)	0.32	0.18
SDO	13	2 ^{ASH}	7	2.5	1.6	1.29	1	0.12 (0.17)	0.29 (0.26)	0.41	-0.07
TOM	9	3 ^{CA}	4	2.2	1.8	1.27	0	0.04 (0.07)	0.27 (0.26)	0.28	0.99
COP	3	2 ^{CA}	4	1.7	1.6	1.27	0	0.07 (0.20)	0.27 (0.36)	0.35	1.00
REY	11	2 ^{CF}	1	1.9	2.0	1.26	0	0.07 (0.16)	0.26 (0.28)	0.39	-0.06
HUA	6	1 ^{AF}	4	1.9	1.8	1.26	0	0.09 (0.13)	0.26 (0.26)	0.26	0.61
JDI	24	2 ^{CA}	5	3.0	1.5	1.21	2	0.08 (0.05)	0.21 (0.16)	0.31	-0.15
PIN	3	0 ^F	1	1.4	1.8	1.21	0	0.03 (0.10)	0.21 (0.33)	0.28	0.86
TEZ	4	1 ^C	5	1.6	1.5	1.19	0	0.15 (0.23)	0.19 (0.23)	0.18	-0.004
Promedio			4	2.4	1.9	1.29	0.5	0.12 (0.16)	0.28 (0.27)	0.31	0.40
Cañada											
CUE	19	2 ^C	1	2.2	1.9	1.25	0	0.13 (0.17)	0.25 (0.25)	0.34	-0.20
CUI	19	3 ^C	1	2.3	1.9	1.22	1	0.09 (0.14)	0.22 (0.28)	0.35	1.00
Promedio			1	2.2	1.9	1.23	0.5	0.11 (0.15)	0.24 (0.27)	0.35	-0.15
Valles											
COA	24	1 ^C	1	2.3	2.0	1.29	0	0.05 (0.11)	0.29 (0.29)	0.35	0.69
LOB	16	1 ^C	3	2.2	1.6	1.27	0	0.12 (0.16)	0.27 (0.18)	0.30	0.05
MIX	10	1 ^C	2	1.9	1.8	1.26	1	0.18 (0.27)	0.26 (0.23)	0.35	1.00
JGB	6	3 ^{CAF}	3	1.9	1.6	1.24	0	0.10 (0.18)	0.24 (0.30)	0.33	1.00
ABA	9	1 ^C	3	1.8	1.6	1.24	0	0.04 (0.06)	0.24 (0.25)	0.23	0.70
ETL	11	4 ^C	3	2.0	1.6	1.22	1	0.12 (0.14)	0.22 (0.24)	0.29	-0.05
LAB	8	1 ^C	1	1.5	1.6	1.20	0	0.01 (0.04)	0.20 (0.28)	0.26	1.00
SPA	4	1 ^C	3	1.5	1.4	1.17	0	0.05 (0.10)	0.17 (0.29)	0.28	0.70
ZAC	8	3 ^{CAF}	2	1.7	1.5	1.15	0	0.10 (0.20)	0.15 (0.27)	0.21	0.33
BAL	5	1 ^C	1	1.4	1.5	1.14	0	0.13 (0.24)	0.14 (0.24)	0.37	-0.60
SCR	10	1 ^C	3	1.4	1.3	1.08	0	0.04 (0.07)	0.08 (0.13)	0.30	0.00
Promedio			2	1.8	1.6	1.20	0.2	0.09 (0.14)	0.21 (0.25)	0.30	0.43
Promedio Total			3	2.2	1.8	1.25	0.4	0.10 (0.15)	0.25 (0.26)	0.30	0.32

^C= cultivados, ^A= arvenses, ^S= silvestres, ^F= *C. frutescens* y ^H= *C. chinense*

Se realizó una correlación entre los valores obtenidos de heterocigosis esperada con la altitud en la que se ubican las poblaciones. Esta correlación puede observarse negativa (ver Fig. 4).

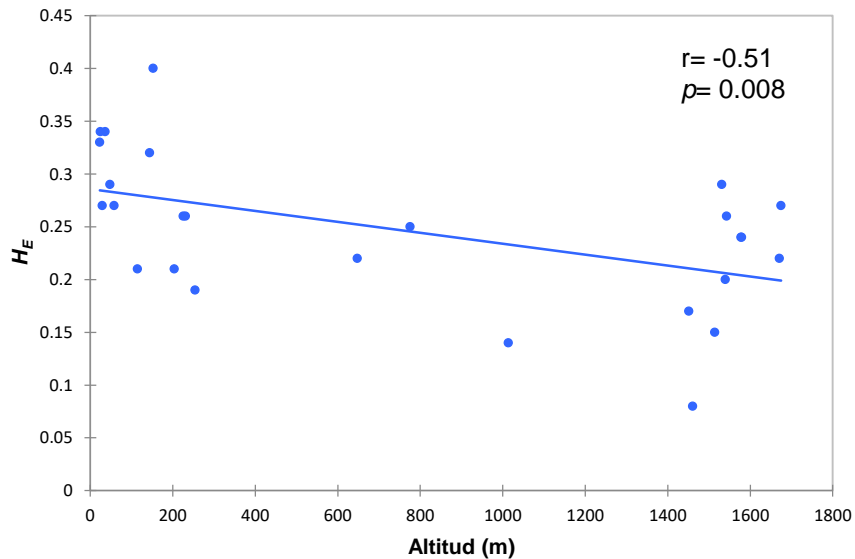


Figura 4. Correlación entre altitud y H_E donde se observa que los valores más altos de H_E (H_E entre 0.30 y 0.40) se encuentran a altitudes más bajas, mientras que valores más bajos de H_E corresponden a una mayor altitud.

Por otra parte, el análisis de rarefacción indicó que las poblaciones Rosedal, Añil-Lagartero, Juan Diegal y Tortolita-Saachilac presentaron los valores más altos de riqueza alélica (3.7, 3.6, 2.9 y 2.8 alelos promedio respectivamente). Estas poblaciones forman parte de la región de la Costa. La riqueza alélica para las otras 24 poblaciones fue menor a 2.5 alelos (ver Fig. 5), mientras que a nivel de región geográfica la riqueza alélica fue de 6.6, 4.7 y 2.6 para la Costa, los Valles Centrales y la Cañada.

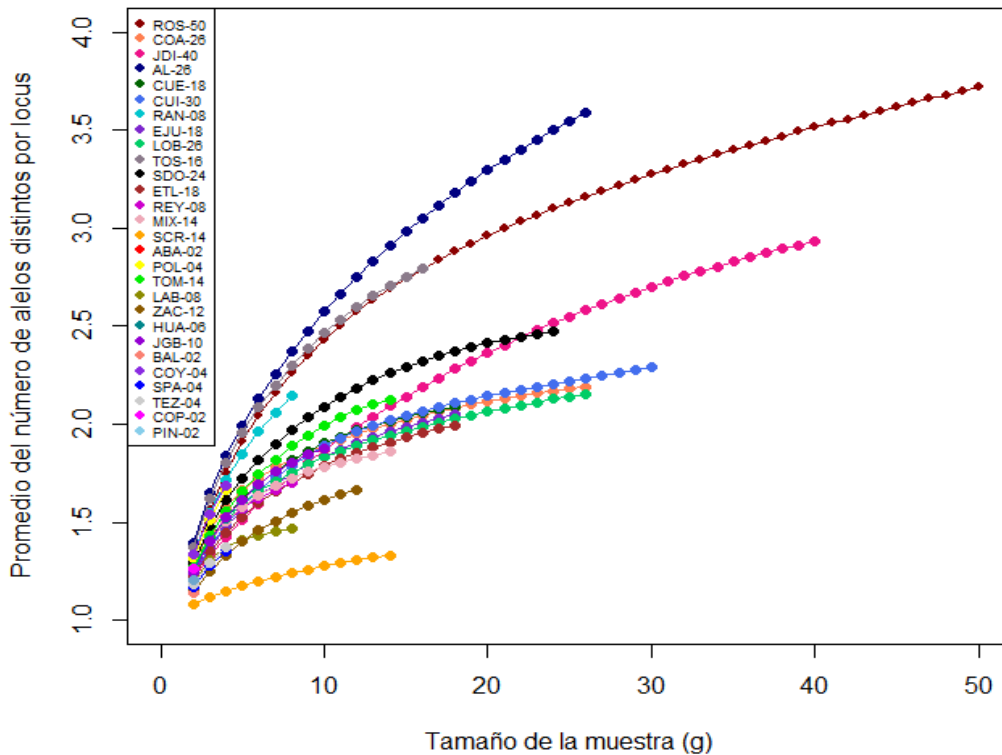


Figura 5. Enrarecimiento de muestra de 28 poblaciones de *C. annuum* en 10 loci de microsatélites (SSR). Se puede apreciar que el promedio de número de alelos máximo es de 3.7 alelos para Rosedal. El número al lado de la población indica el tamaño de la muestra (g).

VII.2.2. Análisis de varianza molecular

El análisis de varianza molecular (AMOVA) para las poblaciones se realizó formando tres grupos: la región de la Cañada, la región de la Costa y la región de los Valles Centrales (ver Tabla 7). Las dos AMOVAS realizadas (poblaciones con y sin individuos de *C. frutescens* y *C. chinense*) muestran que la mayor cantidad de varianza genética total se encuentra dentro de las poblaciones mientras que la menor varianza se da entre regiones (Costa, Valles Centrales y Cañada). Por otro lado, los índices de fijación $F_{ST} c/gpos= 0.10$ y $F_{ST} sin/gpos= 0.09$ nos indican una moderada diferenciación genética en las poblaciones (ver Tabla 8).

En cuanto a los niveles de endogamia F_{IS} que se muestran en la Tabla 6, podemos observar tanto valores positivos como negativos en las tres regiones geográficas donde los mayores niveles se encuentran en la región de los Valles Centrales, $F_{IS} = 0.43$, conformada principalmente de variedades cultivadas. Sin embargo, los valores F_{IS} por población no fueron significativamente diferentes ($F = 0.03$ y $p = 0.97$; ver Anexo 6).

Tabla 7. Resultados del análisis de varianza molecular (AMOVA) de loci de microsatélites por regiones geográficas (Costa, Valles Centrales y Cañada) para 28 poblaciones de *C. annuum* (distancia molecular R_{ST}). Significancia igual a 0.05. Se presentan los resultados para las poblaciones que incluyen individuos pertenecientes a grupos externos e individuos pertenecientes solo a *C. annuum*.

AMOVA para poblaciones con grupos externos				
Fuente de variación	g.l.	Suma de cuadrados	Componentes de la varianza	Porcentaje de variación
Entre regiones	2	4210.585	2.47 Va	0.66
Entre poblaciones dentro de las regiones	25	29934.309	37.12 Vb	9.86
Dentro de las poblaciones	654	220266.097	336.80 Vc	89.48
Total	681	254410.991	376.39	
AMOVA para poblaciones sin grupos externos				
Entre regiones	2	401.503	0.57 Va	2.09
Entre poblaciones dentro de las regiones	25	1653.896	1.96 Vb	7.22
Dentro de las poblaciones	617	15211.88	24.66 Vc	90.69
Total	643	17267.28	27.18535	

Tabla 8. Índices de fijación. La F_{ST} por región geográfica se obtuvo con el programa FSTAT (Goudet, 2002).

Índices de fijación con grupos externos	
$F_{ST}= 0.10$	$p= 0$
$F_{SC}= 0.09$	$p= 0.002$
$F_{CT}= 0.006$	$p= 0.27$
$F_{IT}= 0.38$	$p= 0$
Índices de fijación sin grupos externos	
$F_{ST}= 0.09$	$p= 0.04$
$F_{SC}= 0.07$	$p= 0.10$
$F_{CT}= 0.02$	$p= 0.10$
$F_{IT}= 0.38$	$p= 0$
F_{ST} por región geográfica	
Con grupos externos	Sin grupos externos
$F_{STcañada}= 0.002$	$F_{STcañada}= 0.005$
$F_{STvalles\ centrales}= 0.23$	$F_{STvalles\ centrales}= 0.23$
$F_{STcosta}= 0.06$	$F_{STcosta}= 0.07$

VII.2.3. Distancias genéticas y aislamiento por distancia

VII.2.3.1 Distancias genéticas

El dendograma Neighbor-Joining basado en las distancias genéticas de Goldstein *et al.* (1995), agrupó a las poblaciones en dos principales grupos. El primer grupo se integra por poblaciones de tierras altas (Valles Centrales y Cañada), Rosedal, Polvorín, Juan Diegal, Reyes, Añil-Lagartero y Rancho Llano de la Costa de Oaxaca. El segundo grupo se integra de las poblaciones de la Costa que se encuentran en los extremos del suroeste y sureste de esta región a excepción de Tomatal de la zona central de la Costa (Ver Fig. 6). El análisis con 1, 000 bootstaps indica que el sostenimiento de las ramas no es confiable. Paso Indios (PIN) representa al grupo externo con individuos de *C. frutescens*.

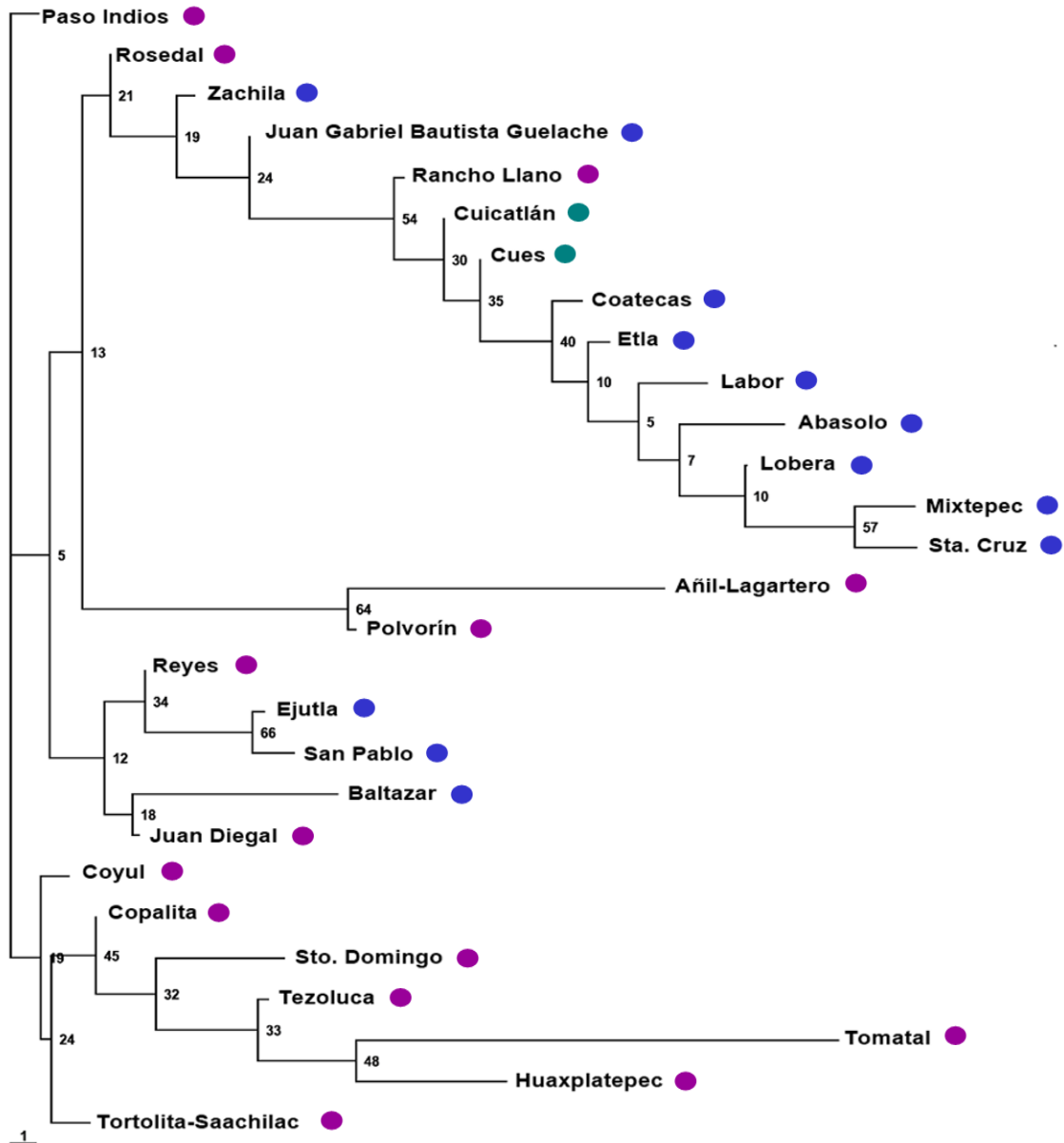


Figura 6. Dendrograma Neighbor-Joining basado en las distancias genéticas de Goldstein *et al.*, (1995) de 10 loci de microsatélites (SSR) en poblaciones de *C. annuum*. Los puntos morados ● indican poblaciones pertenecientes a la región de la Costa mientras que los puntos azules ● representan a los Valles Centrales y los verdes ● a la Cañada. Se presentan los bootstraps.

El dendrograma Neighbor-Joining basado en las distancias genéticas de Nei *et al.*, (1983) agrupó a las poblaciones en dos grupos: un grupo integrado por poblaciones de la región de la Costa y Baltazar, y otro grupo integrado de las poblaciones de los Valles

Centrales, la Cañada y las poblaciones Rosedal, Añil-Lagartero y Rancho Llano de la Costa de Oaxaca (ver Fig. 7). El análisis con 1, 000 bootstaps indica que el sostenimiento de las ramas no es confiable. Paso Indios (PIN) representa al grupo externo con individuos de *C. frutescens*.

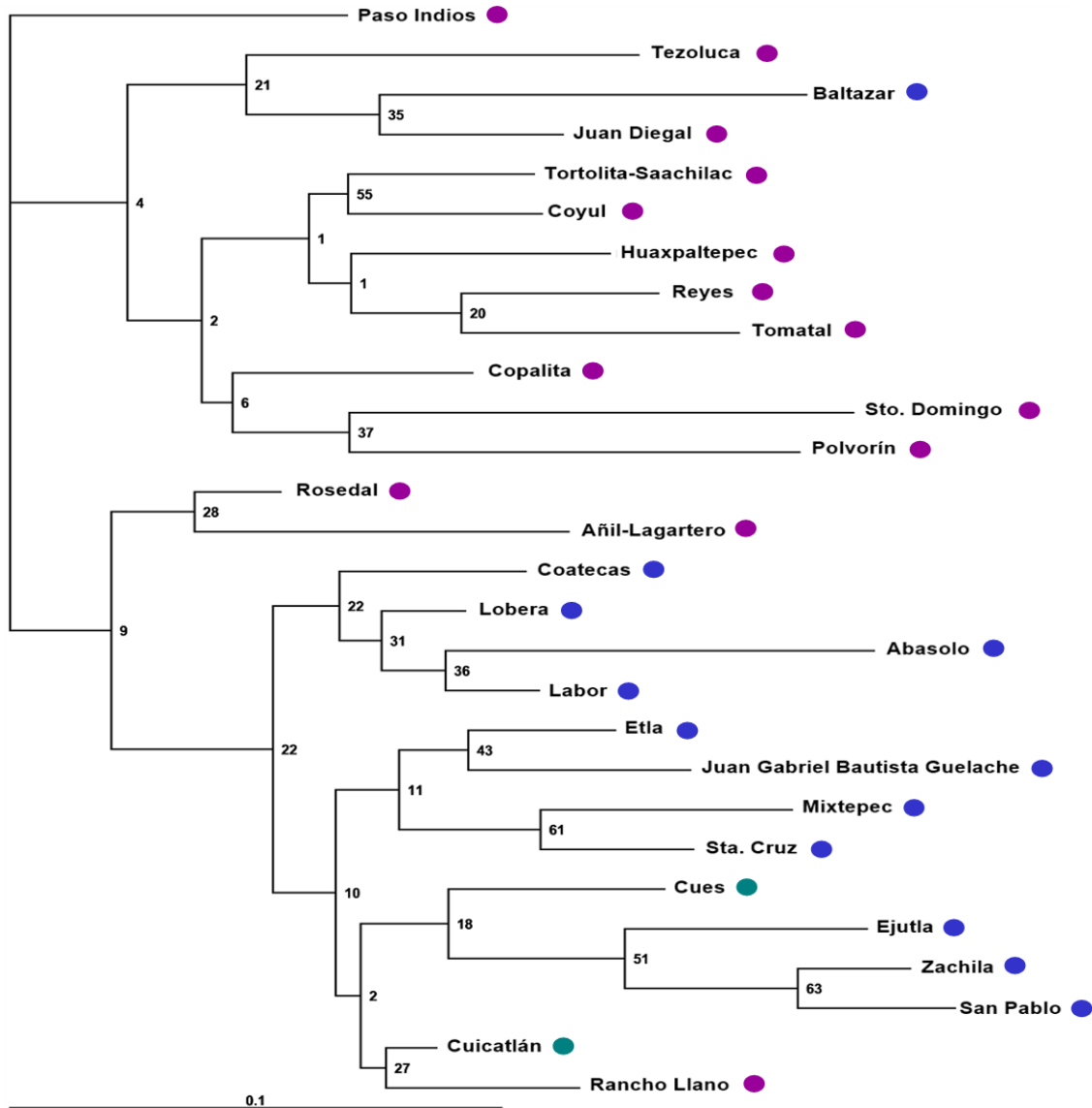


Figura 7. Dendrograma Neighbor-Joining basado en las distancias genéticas de Nei *et al.*, (1983) de 10 loci de microsatélites (SSR) en poblaciones de *C. annuum*. Los puntos morados ● indican poblaciones pertenecientes a la región de la Costa mientras que los puntos azules ● representan a los Valles Centrales y los verdes ● a la Cañada. Se presentan los bootstraps.

VII.2.3.2. Aislamiento por distancia

Las pruebas de Mantel incluyendo y sin incluir individuos de *C. frutescens* y *C. chinense*, nos indica que no existe aislamiento por distancia en las poblaciones. Sin embargo, al tomar en cuenta los individuos de *C. frutescens* y *C. chinense* la correlación se vuelve negativa mientras que la correlación es positiva incluyendo solo individuos de *C. annuum*. Es posible apreciar distancias R_{ST} negativas debido a un exceso de heterócigos (ver Fig. 8).

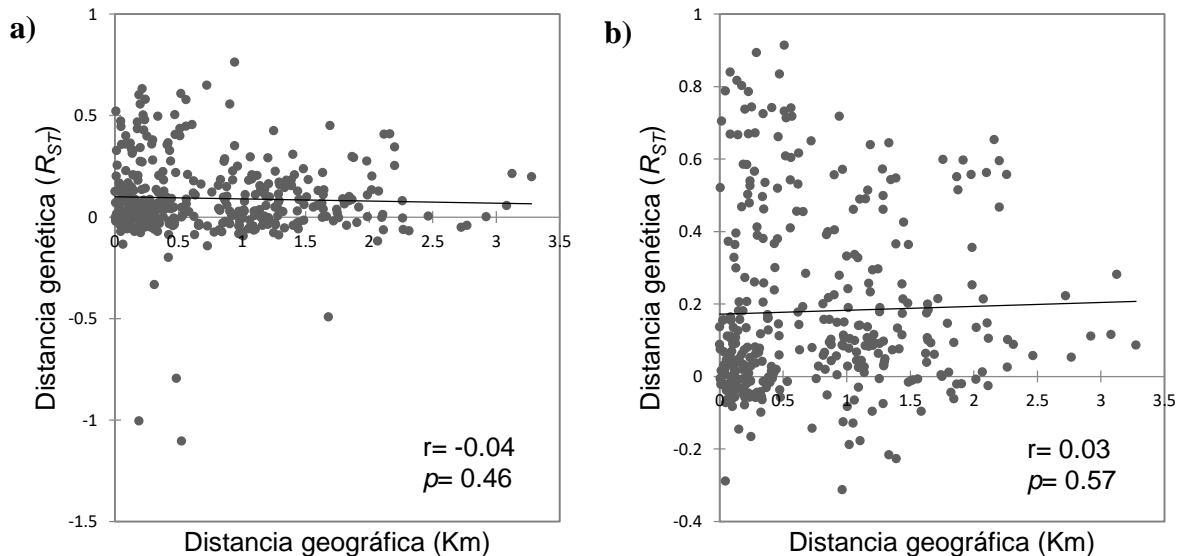


Figura 8. Resultados de la prueba de Mantel con distancia genética R_{ST} y distancias euclidianas (geográfica) muestran que no existe una correlación entre ambas. a) Resultados incluyendo individuos de *C. frutescens* y *C. chinense* en las poblaciones y b) Resultados solo para individuos de *C. annuum*.

Las pruebas de Mantel por transecto de muestreo, tierras altas (Valles Centrales y Cañada) y tierras bajas (Costa), no indican aislamiento por distancia. La correlación para tierras altas dio negativa en ambos casos (con y sin individuos de *C. frutescens* y *C.*

chinense). Sin embargo, la correlación para tierras bajas involucrando individuos de *C. frutescens* y *C. chinense* da positiva y con una $p= 0.052$ muy baja a comparación de la correlación para poblaciones con individuos únicamente de *C. annuum* (ver Fig. 9).

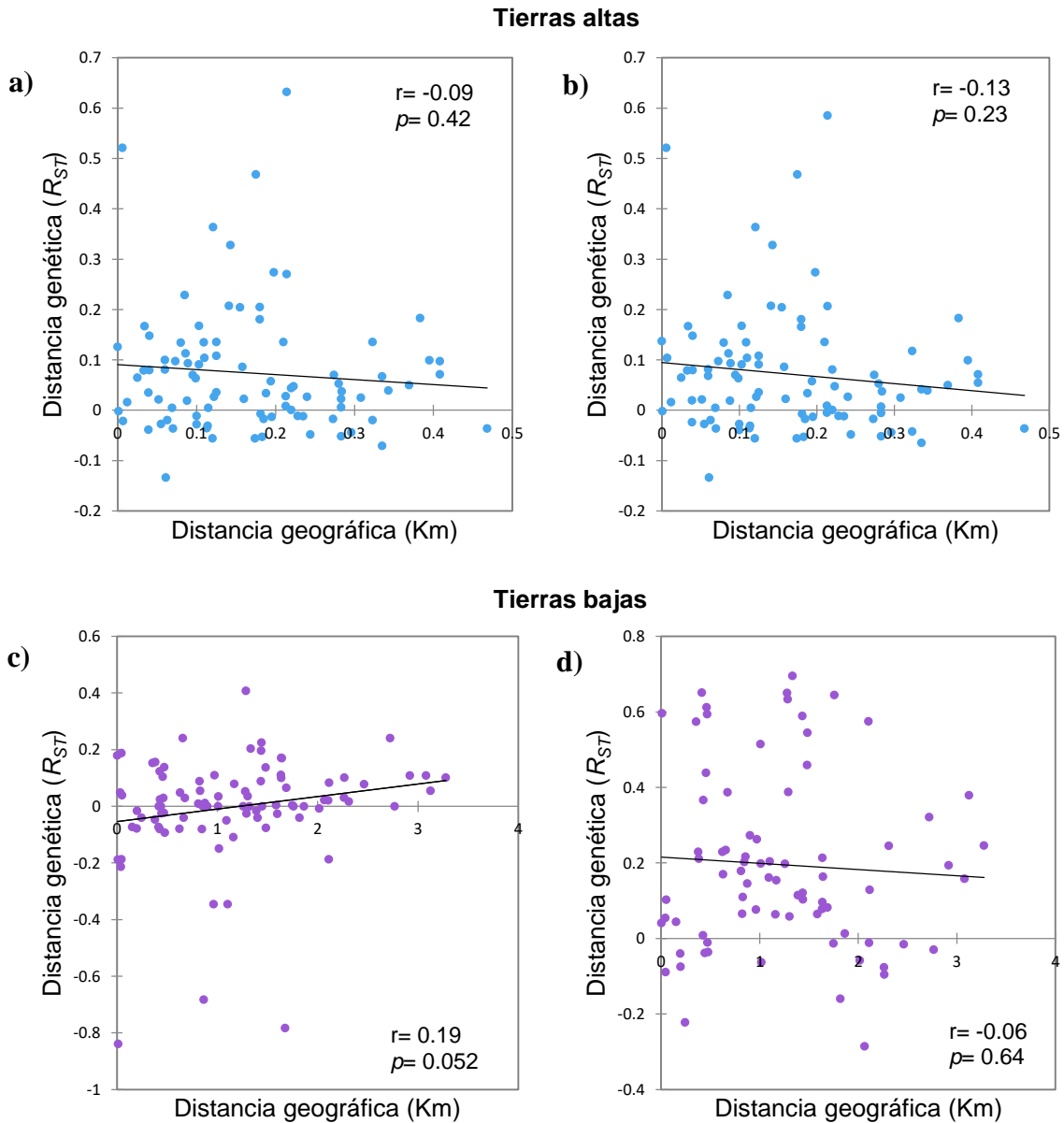


Figura 9. Resultados de la prueba de mantel por transecto con distancia genética R_{ST} y distancias euclidianas (geográfica) muestra que no existe una correlación entre ambas. a) y c) Resultados con grupos externos (*C. frutescens* y *C. chinense*) para tierras altas (Valles Centrales y Cañada) y bajas (Costa) respectivamente. b) y d) Resultados solo de individuos *C. annuum* para tierras altas y bajas respectivamente.

VII. 2.3.3. Barreras geográficas

Los resultados del análisis espacial de varianza molecular (SAMOVA) para poblaciones que incluyen individuos de *C. frutescens* y *C. chinense* obtuvo una $K=6$, $F_{CT}=0.07$ y $F_{ST}=0.11$; para poblaciones únicamente con individuos pertenecientes a *C. annuum* se obtuvo una $K=2$, $F_{CT}=0.05$ y $F_{ST}=0.11$ observándose que no existe una gran diferenciación entre las poblaciones en ambos casos (ver Tabla 9). En la Fig. 10 se puede apreciar los mapas con las agrupaciones geográficas hechas por SAMOVA. Los resultados individuales de F_{CT} para las 15 corridas se muestran en el Anexo 7.

Tabla 9. Resultados del análisis espacial de varianza molecular (SAMOVA) y la prueba AMOVA de loci de microsatélites para 28 poblaciones de *C. annuum* (distancia molecular R_{ST}). Se presentan los resultados para las poblaciones que incluyen individuos pertenecientes a grupos externos (*C. frutescens* y *C. chinense*) con una $K=6$ e individuos pertenecientes solo a *C. annuum* con una $K=2$.

AMOVA para poblaciones con grupos externos (K= 6)				
Fuente de variación	g.l.	Suma de cuadrados	Componentes de la varianza	Porcentaje de variación
Entre grupos	5	18092.516	25.64 Va	6.75
Entre poblaciones dentro de los grupos	22	16052.378	17.39 Vb	4.58
Dentro de las poblaciones	654	220266.097	336.80 Vc	88.67
Total	681	254410.991	379.82771	
AMOVA para poblaciones sin grupos externos (K= 2)				
Entre grupos	1	409.02	1.39 Va	5.01
Entre poblaciones dentro de los grupos	25	1646.379	1.76 Vb	6.34
Dentro de las poblaciones	617	15211.88	24.65 Vc	88.65
Total	643	17267.28	24.81153	

En los grupos obtenidos por SAMOVA con una $K=6$ podemos observar que hay 3 grupos bien definidos en la región de los Valles Centrales, dos de los cuales se conectan a la costa con la población Tezoluca y Tortolita-Saachilac. La Costa presenta solo un grupo homogéneo y las demás poblaciones comparten grupo con poblaciones de Valles Centrales y la Cañada. Entre los 4 grupos que se observan más homogéneos podría existir una barrera al flujo la cual pertenece a la Sierra Madre del Sur (ver Fig. 10a).

En los grupos obtenidos con una $K=2$ (sin incluir individuos de *C. frutescens* y *C. chinense*) las poblaciones de la parte central de los Valles Centrales se mantienen homogéneas compartiendo grupo con la población Tezoluca, Huaxpaltepec y Reyes de la región de la Costa (ver Fig. 10b).

Los resultados del análisis MONMONIER realizado con y sin individuos de *C. frutescens* y *C. chinense* se muestran en líneas moradas en la Fig. 10. Las barreras al flujo encontradas se ubican en la parte central de la región de la Costa y una de estas barreras afecta el flujo entre Baltazar y los Valles Centrales.

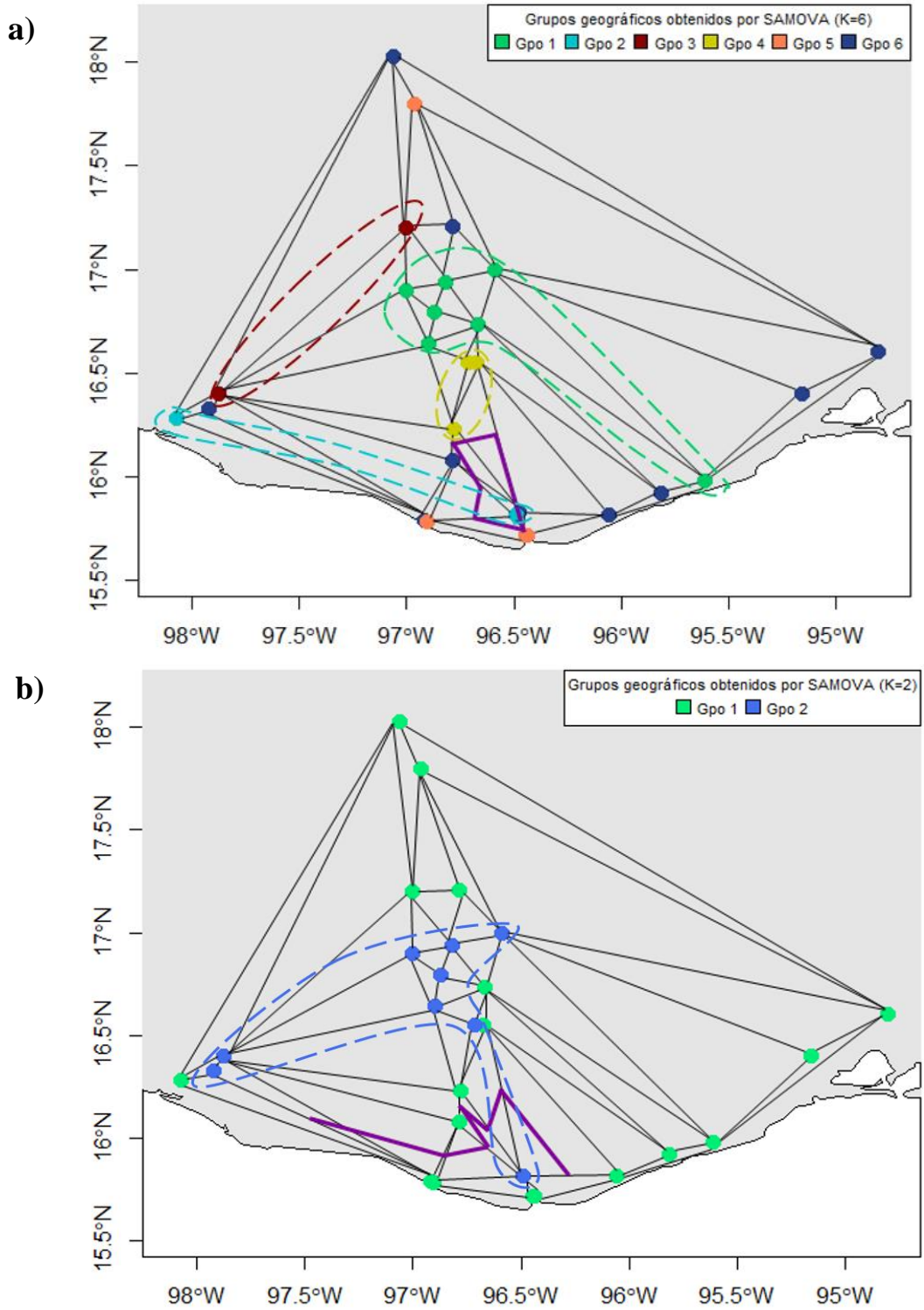


Figura 10. Mapa obtenido por SAMOVA y resultados del análisis MONMONIER. a) Mapa para 6 grupos (K= 6) y b) Mapa para 2 grupos (K= 2). En líneas moradas se aprecia las posibles barreras al flujo detectadas por el análisis MONMONIER en donde el flujo génico entre Valles Centrales y Baltazar es afectado por las barreras geográficas encontradas.

VII.2.4. Estructura poblacional

VII.2.4.1. Análisis multivariado (DAPC)

El análisis discriminante de componentes principales (DAPC) realizado por población no arrojó una clara distinción entre las poblaciones, siendo Santo Domingo la única población que se observa separada de las demás. Esta localidad presenta individuos arvenses y silvestres de *C. annuum* y en el caso de individuos de grupos externos presenta dos de *C. chinense* (ver Fig. 11). La probabilidad de asignación promedio por población fue de 30% (ver Anexo 8).

El DAPC a nivel regional muestra grupos más claramente definidos (ver Fig. 12). La probabilidad de asignación para los Valles Centrales y la Costa fue superior al 80% mientras que la probabilidad de la Cañada fue del 30%.

Los análisis se realizaron con y sin presencia de individuos de *C. frutescens* y *C. chinense*, sin embargo, no se observan grandes diferencias en el patrón de distribución de los datos. En el Anexo 8 se muestra la probabilidad de asignación a cada grupo y los eigenvalores.

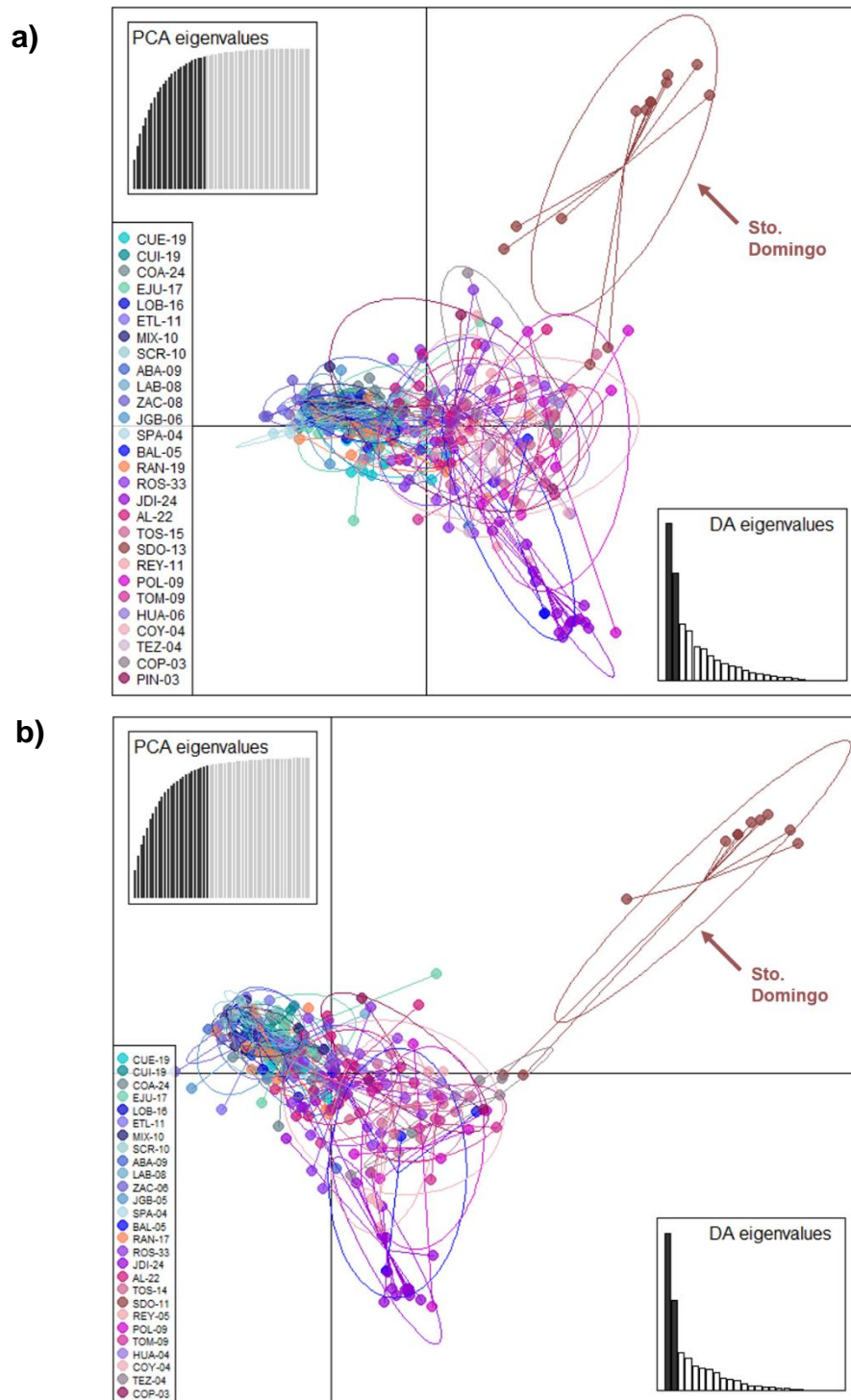


Figura 11. Análisis discriminante de componentes principales (DAPC) por población. a) Análisis incluyendo individuos de *C. frutescens* y *C. chinense*. b) Análisis solo de individuos *C. annuum*. En ambos casos se puede apreciar que la población SDO se aleja de las demás poblaciones que no pueden diferenciarse una de otra.

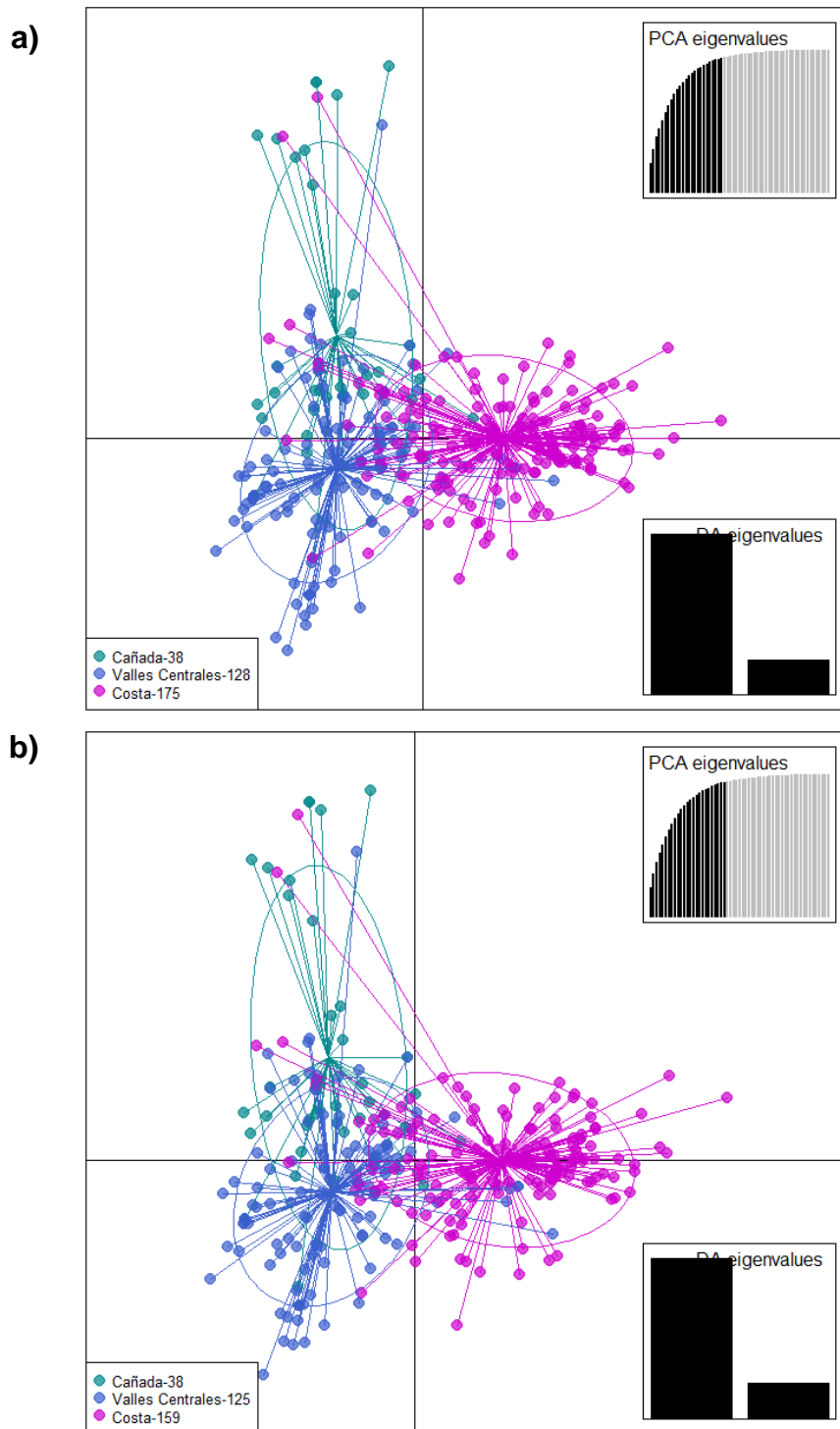


Figura 12. Análisis discriminante de componentes principales (DAPC) por región geográfica. a) Análisis incluyendo individuos de *C. frutescens* y *C. chinense* b) Análisis solo de individuos *C. annuum*. En ambos casos no se observa un cambio drástico en el patrón de dispersión de los datos. El número al lado indica el número de individuos analizados.

VII.2.4.2. Análisis bayesiano (STRUCTURE)

Los resultados del análisis bayesiano de estructura poblacional con el programa STRUCTURE con la corrección de Evanno *et al.*, (2005) nos despliega dos grupos ($K=2$) (ver Fig. 13) con un valor máximo de verosimilitud de $L = -3983.69$ y una probabilidad de 0.85 (ver Anexo 9).

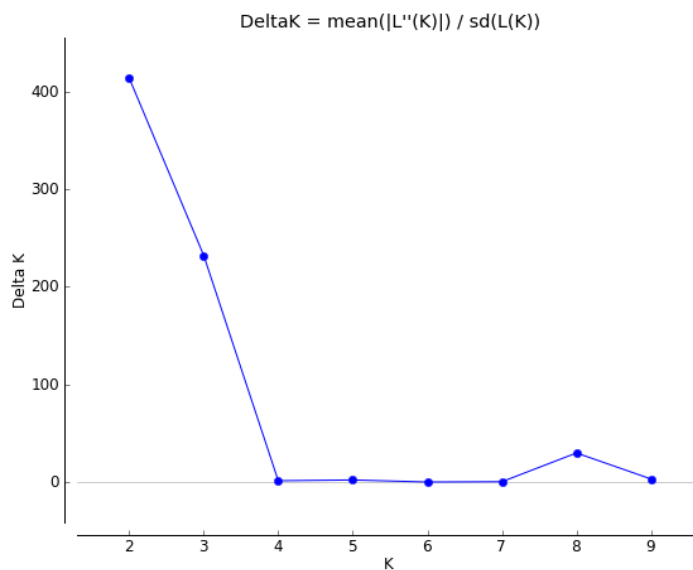


Figura 13. Gráfica de los resultados de la prueba de Evanno, *et al.*, (2005), $K=2$

Los resultados de la probabilidad de asignación a cada grupo se encuentran mayor a 0.5 (ver Anexo 9). Sin embargo, la población ROS muestra una probabilidad de asignación de 0.56 para el grupo 2, lo cual nos sugeriría una mezcla casi del 50% entre el grupo 1 y grupo 2. El primer grupo con una $K= 2$ está definido por 11 poblaciones de la región de los Valles Centrales, las 2 poblaciones de la Cañada y RAN de la región de

la Costa; el segundo grupo lo integran 13 poblaciones de la región de la Costa y BAL designada para los Valles Centrales.

En la Fig. 14 se observa el gráfico de los resultados del análisis STRUCTURE, el cual fue ordenado según la asignación a los grupos resultados del análisis y en el que se observa que los dos grupos definidos ($K= 2$) coinciden con la altitud en la que son cultivados los chiles. Estos dos grupos pertenecen a tierras bajas (0-600 msnm) y altas (600-1700 msnm) del estado de Oaxaca, a excepción de la población RAN, y BAL, sin embargo, está última presenta una localización geográfica cercana a la región de la Costa (ver Fig 3a).

Aunque la prueba de Evanno *et al.*, (2005) arrojó una $K= 2$, en la Fig. 14 se colocaron los resultados para una $K= 3$ y una $K= 4$. Si consideramos una $K= 3$ se aprecia que la asignación al grupo 1 (ver Fig. 14 color verde) mantiene un patrón más homogéneo que la Costa (ver Fig. 14 morado y azul) y la asignación a un tercer grupo parece coincidir con la ubicación de individuos silvestres (ver Fig. 14 puntos rojos). Considerando una $K= 4$ la asignación genética mantiene un patrón similar al obtenido con $K= 2$ y $K=3$. En la Fig. 15 se observan la proyección en mapas de los resultados de probabilidad de asignación de los individuos a los grupos.

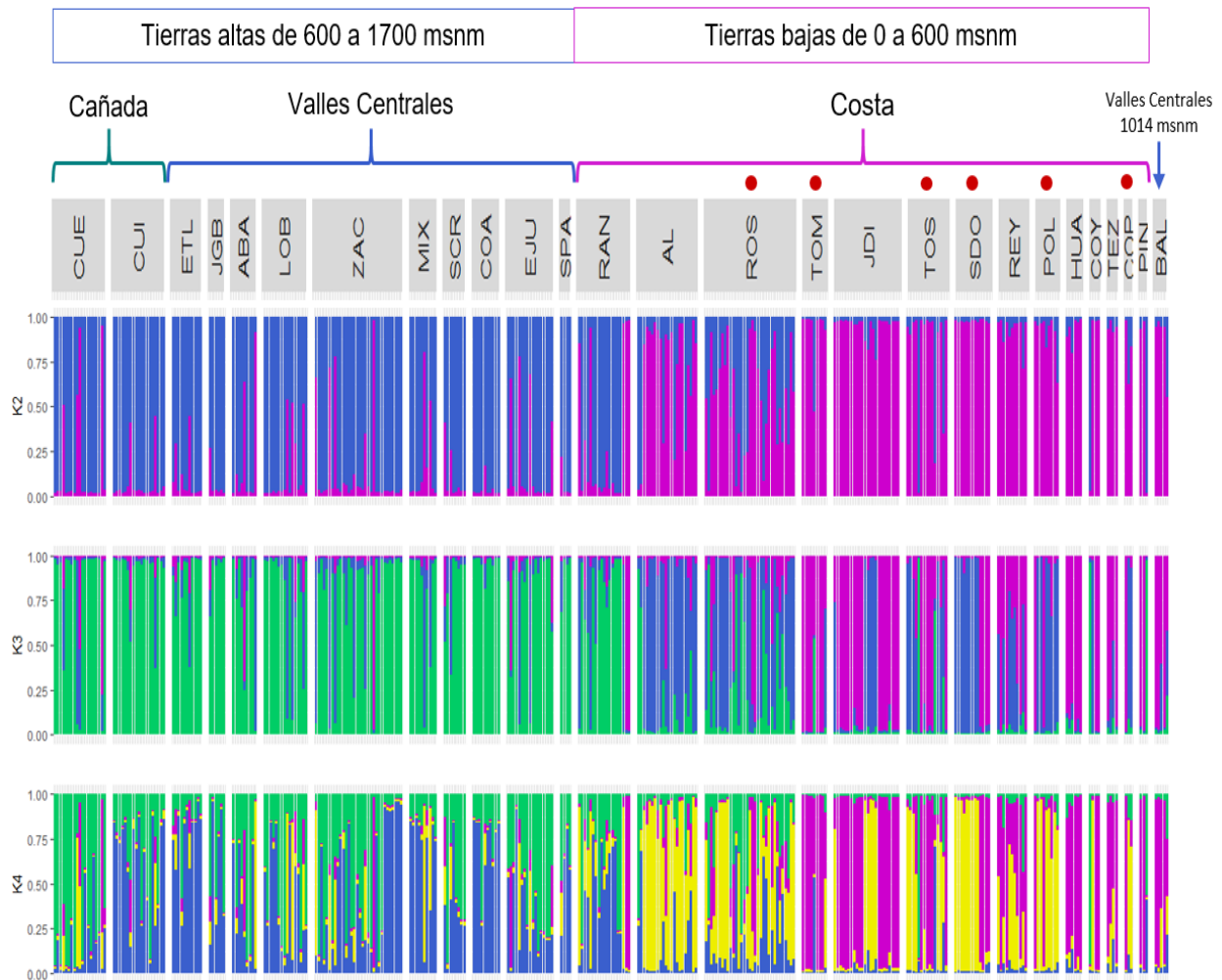


Figura 14. Resultados poblacionales obtenidos con el programa STRUCTURE. Se presentan $K=2$, $K=3$ y $K=4$. Prueba de Evanno, *et al.*, 2005 de $K=2$. El primer grupo coincide con las poblaciones pertenecientes a la Cañada y Valles Centrales a excepción de BAL, mientras que el segundo grupo coincide con las poblaciones de la Costa a excepción de RAN que adquiere un color más parecido al grupo que coincide con Valles Centrales y la Cañada. Los puntos rojos ● indican la presencia de individuos silvestres en las poblaciones.

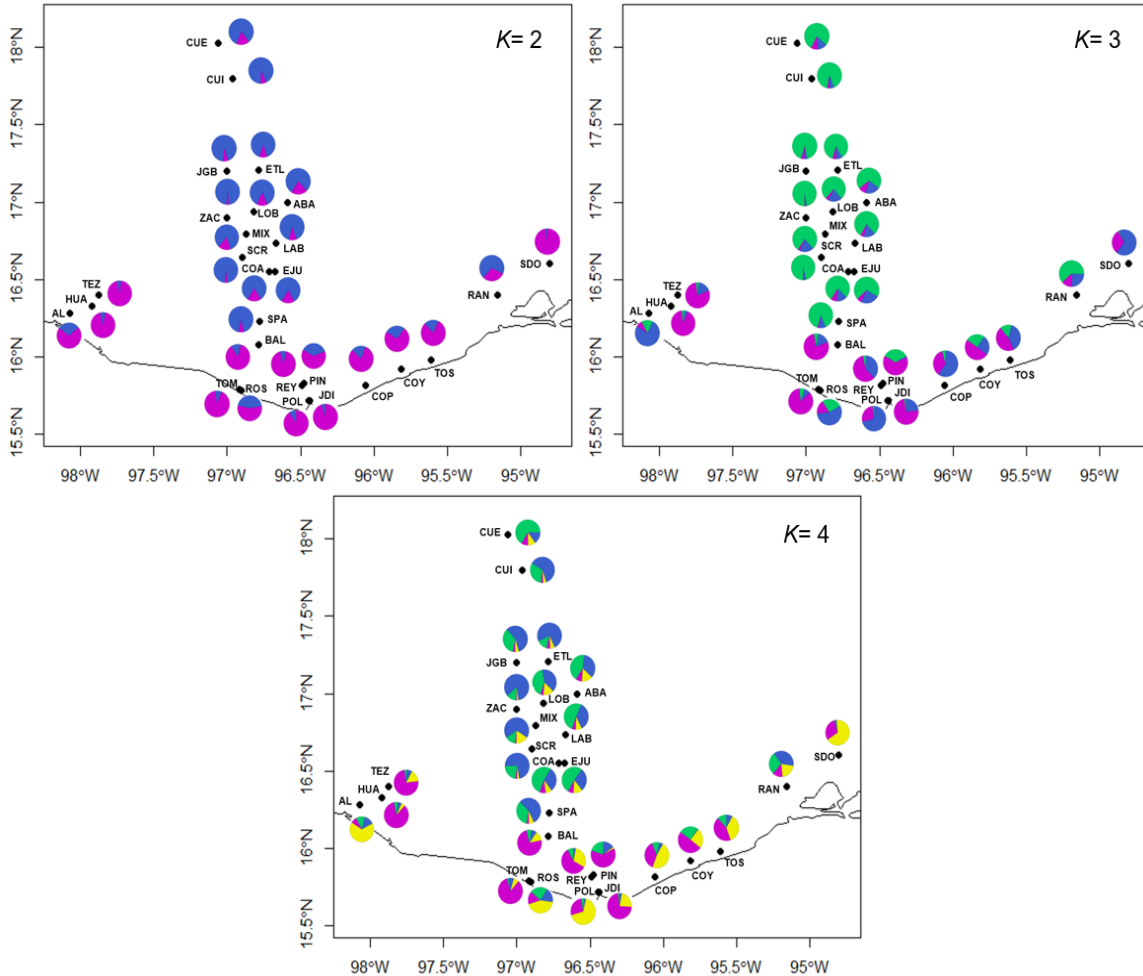


Figura 15. Mapas representando la asignación de los grupos hecho por STRUCTURE. Para $K=2$ se observa que la asignación al grupo 1 (color azul) se encuentra mayoritariamente en Valles Centrales y Cañada mientras que la asignación al grupo 2 (color morado) es mayoritaria en la región de la Costa. En $K=3$ y $K=4$ se observa que la asignación a grupos mantiene los grupos divididos entre tierras altas (Valles Centrales y Cañada) y tierras bajas (Costa).

Para comprobar si existe una relación entre la asignación de los individuos con la altitud en la que son cultivados los chiles se utilizaron los resultados de la asignación de los individuos al grupo 2 ($K=2$) para realizar una correlación con la altitud. Esta correlación fue negativa y el valor de $p=0.0001$ indica que si existe una correlación entre las variables (ver Fig.16).

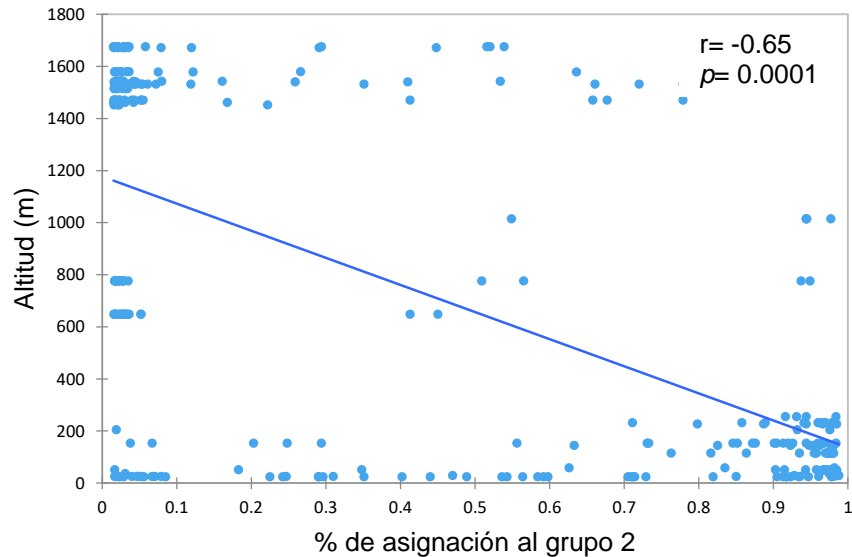


Figura 16. Correlación entre el porcentaje de asignación genética y la altitud de los 341 individuos analizados. Bajas altitudes se asignaron al grupo 2, mientras que altitudes altas presentó un menor porcentaje de asignación al grupo 2.

VII.2.4.3. Análisis gráfico poblacional

Los resultados obtenidos con el método de gráficas poblacionales, sin involucrar parámetros geográficos, nos indican niveles de covarianza similares entre las poblaciones (ver Fig. 17). Las conexiones entre SDO-JGB, SDO-TEZ, SDO-TOM y SDO-TOS presentaron la mayor covarianza. Los niveles más altos de covarianza al interior de las poblaciones se encontraron en LAB, BAL, COY, ETL, TOS y TEZ. La población HUA fue relacionada solo a TOM, siendo la población con menor número de aristas. La topografía mínima arrojó 26 nodos y 47 aristas (ver Fig. 17).

Los eventos de migración más largos entre poblaciones ocurrieron entre la región de la Costa y la Cañada, y, entre las mismas poblaciones de la Costa. En cambio, los Valles Centrales presentan niveles bajos de dispersión (ver líneas Fig. 18). La detección

de barreras al flujo génico se basa en la premisa de que en zonas con pocas líneas podría presentarse una barrera que impide el flujo de genes entre las poblaciones. Ante este panorama, las principales barreras se ubicaron en la zona suroeste de la costa de Oaxaca, la cual impediría el flujo entre las poblaciones TEZ y HUA con la zona noroeste de los Valles centrales y la Cañada. Es posible identificar una barrera más en la zona suroeste de la Costa hacia la región de la Cañada, y ésta correspondería a la Sierra de Tlaxiaco.

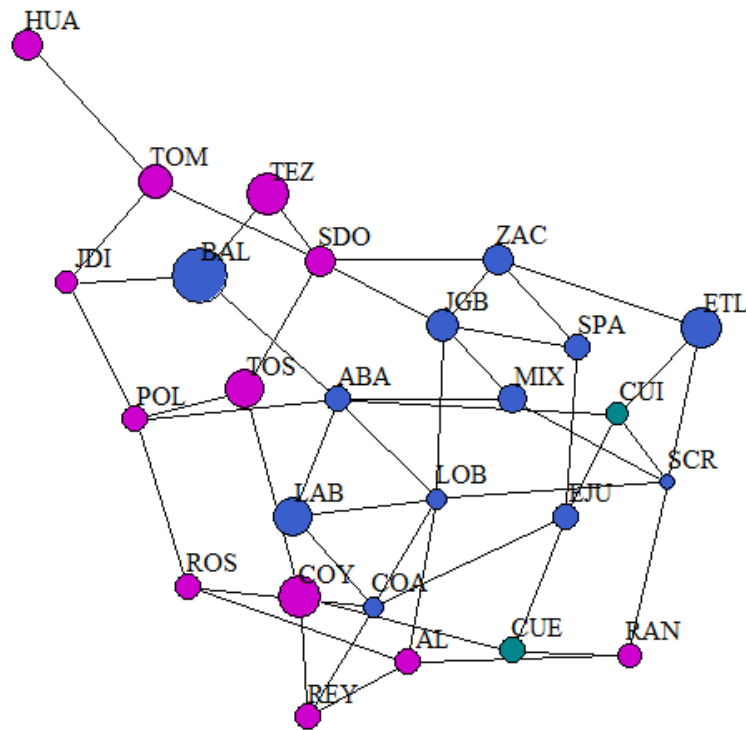


Figura 17. Gráfica poblacional Popgraph (Dyer, 2004; 2015) mostrando la topografía genética y las conexiones entre 26 poblaciones de *C. annuum* del estado de Oaxaca. Este gráfico solo muestra las relaciones y la varianza entre y dentro de las poblaciones sin contraste geográfico. El largo de la línea indica la covarianza entre poblaciones y el tamaño del nodo la varianza dentro de las poblaciones. Los puntos azules ● representan los Valles Centrales, los morados ● a la Costa y los verdes ● la Cañada.

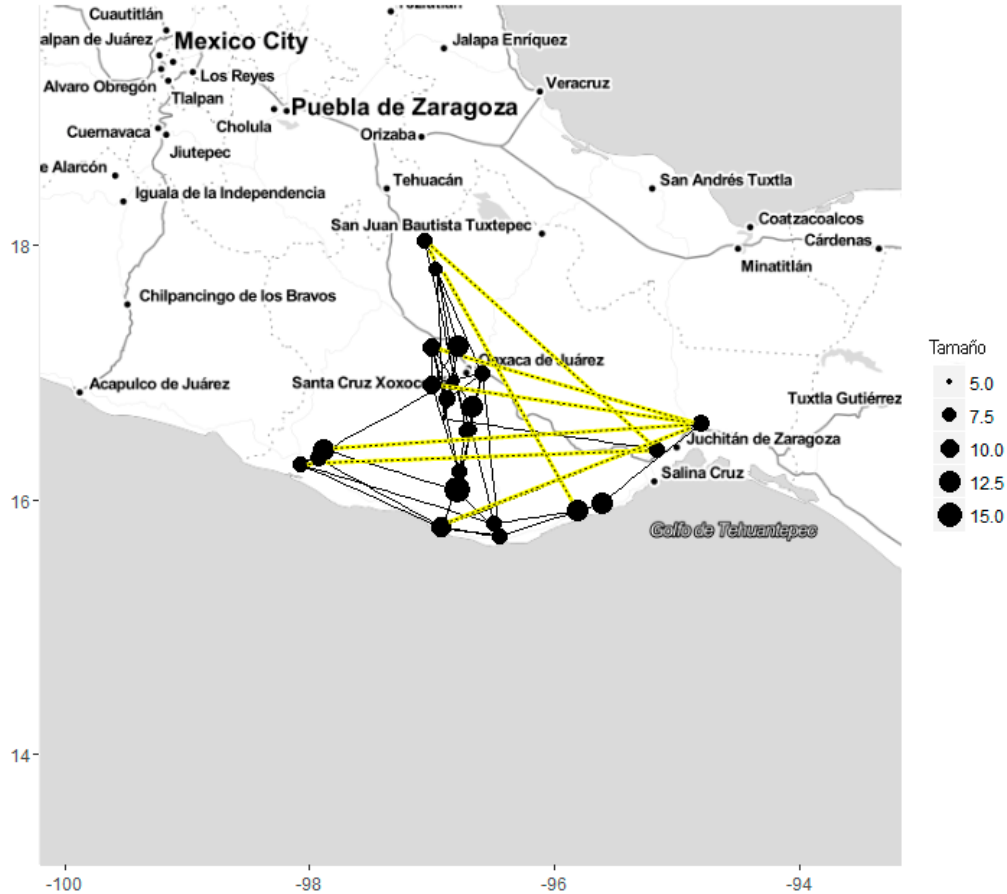


Figura 18. Gráfica poblacional Popgraph (Dyer, 2004; 2015) mostrando las conexiones entre 26 poblaciones de *C. annuum* del estado de Oaxaca. Este gráfico involucra las variables geográficas para detectar flujo génico. El largo de la línea indica la covarianza y la dispersión entre poblaciones. El tamaño del nodo la varianza dentro de las poblaciones. Las líneas amarillas representan los posibles eventos de migración más largos.

Tomando en cuenta la distancia genética condicional entre la conectividad de los nodos y la distancia geográfica de las poblaciones se obtuvo una gráfica de aislamiento por distancia y aislamiento por altitud (ver Fig. 19). Los resultados sugieren aislamiento a nivel de paisaje. Probablemente, la altitud en la que se cultivan los chiles esté influyendo en el aislamiento encontrado.

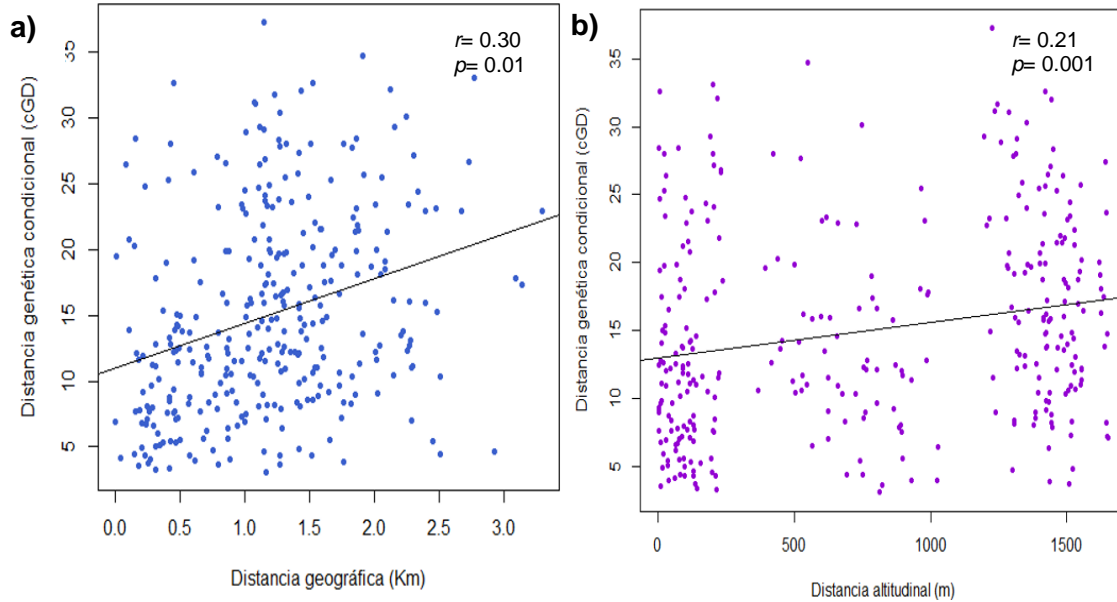


Figura 19. Gráfica representando aislamiento por distancia a nivel de paisaje. a) Aislamiento por distancia ($r = 0.30$ y $p = 0.01$). b) Aislamiento por altitud ($r = 0.21$ y $p = 0.001$).

Se realizó una prueba de Bonferroni con la finalidad de conocer si existía alguna correlación falsa positiva ajustando el valor de p . Los resultados fueron de $p = 0.02$ y $p = 0.002$ para aislamiento por distancia y por altitud respectivamente. Por lo tanto, ambas correlaciones indican aislamiento.

VII.2.4.4. Cuellos de botella y tamaño efectivo poblacional (N_e)

El índice de Garza-Williamson (M) promedio fue de 0.30 ($SD = 0.18$) para las poblaciones, Se considera que valores menores a 0.68 indican la presencia de cuellos de botella y el nivel más alto encontrado fue para la localidad de Santo Domingo ($M = 0.41$) mientras que el valor más bajo se encuentra en Tezoluca ($M = 0.18$). En general las poblaciones presentan niveles similares de M entre 0.20 a 0.40 (ver Tabla 5).

A nivel de región geográfica los valores M calculados son menores a 0.68 y muy similares entre regiones $M= 0.31$ ($SD= 0.17$), $M= 0.30$ ($SD= 0.20$) y $M= 0.35$ ($SD= 0.19$) para la Costa, Valles Centrales y Cañada (ver Tabla 5). En el análisis hecho en BOTTLENECK los eventos de cuello de botella por regiones geográficas se encuentran principalmente en las regiones de la Cañada y Valles Centrales (tierras altas) con los modelos de mutación de dos fases (T.P.M.) y mutación paso a paso (S.M.M.) y las pruebas de significancia y Wilcoxon (p y pW) (ver Tabla 10). La región de la Costa (tierras bajas) presenta evidencia de cuello de botella con valores de $p= 0.01$ para la prueba de Wilcoxon con dos colas en los modelos T.P.M. y S.M.M.

El tamaño efectivo en tierras bajas $N_e= 26.5$ fue mayor al de tierras altas $N_e= 12.4$ (ver Tabla 10). Los intervalos de confianza JackKnife para tierras bajas fue de 15.3 a 45.7 y para tierras altas fue de 3.1 a 31.5.

Tabla 10. Cuellos de botella y tamaño efectivo por región geográfica de chiles *C.annuum* en 10 loci de microsatélites (SSRs). I.A.M.= modelo de alelos infinitos, T.P.M.= modelo de mutación dos fases, S.M.M.= modelo de mutación paso a paso y N_e = estimación del tamaño efectivo de la población con un valor crítico $P_c= 0.02$. Varianza para T.P.M.= 5% y probabilidad para I.A.M. y S.M.M.= 95%. (*) = desviación significativa del equilibrio ($p < 0.05$), p = probabilidad obtenida de la prueba significativa y pW = probabilidad obtenida de la prueba Wilcoxon con dos colas.

Modelos	Tierras bajas (Costa)		Tierras altas (Valles Centrales y Cañada)	
	p	pW	p	pW
I.A.M.	0.23	0.70	0.47	1.0
T.P.M.	0.06	0.01*	0.003*	0.002*
S.M.M.	0.06	0.01*	0.003*	0.002*
N_e	26.5		12.4	

VII.3. Análisis por grado de domesticación

VII.3.1. Diversidad genética

El análisis de diversidad genética por categoría de domesticación (cultivados, arvenses y silvestres) indicaron que los niveles de diversidad H_E fueron similares entre las categorías siendo la categoría arvense ligeramente más alta a las otras dos categorías (ver Tabla 11). La heterocigosis total fue $H_T = 0.36$ (SD= 0.32).

La mayor cantidad de alelos privados se encontraron en los chiles cultivados y pertenecen a los loci CaES3862 (1), CaES5392 (3), CaES0425 (2) y EPMS386 (3).

Tabla 11. Heterocigosis observada (H_o) y esperada (H_E) promedio por grado de domesticación. n = número de individuos, P = número de loci polimórficos, N = promedio del número de alelos, N_e = promedio del número de alelos efectivos, A = riqueza alélica promedio basada en el tamaño mínimo de muestra de 15 individuos, N_p = número de alelos privados, M = índice de Garza y Williamson, 2001 y F_{IS} = Coeficiente de endogamia.

Grado de domesticación	n	P	N	Ne	A	Np	H_o (SD)	H_E (SD)	M	F_{IS}
Cultivados	261	4	6.4	2.7	3.39	7	0.12 (0.11)	0.33 (0.31)	0.26	0.64
Arvenses	33	2	4.0	2.3	3.61	1	0.12 (0.10)	0.37 (0.23)	0.31	0.17
Silvestres	23	4	3.5	2.2	3.30	1	0.19 (0.16)	0.35 (0.29)	0.34	0.25
Promedio		3.3	4.6	2.4	3.72	3.0	0.14 (0.12)	0.35 (0.28)	0.30	0.62

Evaluando la diversidad genética por cada variedad de chile cultivado (14 variedades), se encontró que todas las variedades presentaron loci polimórficos excepto el chile Huacle y que los niveles de H_E encontrados fueron bajos en cada variedad siendo Tusta y Mirasol los que presentaron la menor variación (ver Tabla 12).

Tabla 12. Heterocigosis observada (H_o) y esperada (H_E) promedio por variedad. n= número de individuos por variedad, P= número de loci polimórficos, N= promedio del número de alelos, Ne= promedio del número de alelos efectivos, A= riqueza alélica promedio, Np= número de alelos privados, M= índice de Garza y Williamson, 2001 y F_{IS} = Coeficiente de endogamia.

Variedad	n	P	N	Ne	A	Np	H_o (SD)	H_E (SD)	M	F_{IS}
Cultivados										
Costeño Rojo	45	6	4.6	2.3	1.36	1	0.23 (0.22)	0.36 (0.28)	0.28	0.48
Costeño Amarillo	2	5	1.7	1.9	1.35	2	0.30 (0.33)	0.35 (0.39)	0.32	0.29
Guiña Danhi	17	3	2.5	2.0	1.31	1	0.15 (0.18)	0.31 (0.26)	0.27	0.79
Chilcostle Rojo	4	4	1.6	1.9	1.28	0	0.07 (0.22)	0.28 (0.31)	0.43	1.00
Paradito	10	3	2.4	1.9	1.28	0	0.15 (0.19)	0.28 (0.31)	0.27	0.34
Agua	79	3	3.5	1.9	1.27	1	0.08 (0.11)	0.27 (0.23)	0.31	0.99
Árbol	12	1	2.3	1.8	1.27	0	0.13 (0.16)	0.27 (0.20)	0.28	0.02
Serrano Criollo	9	2	1.8	1.6	1.23	0	0.16 (0.22)	0.23 (0.25)	0.38	1.00
Taviche	21	3	2.4	1.9	1.23	0	0.04 (0.05)	0.23 (0.27)	0.34	0.78
Achilito	3	2	1.6	1.8	1.22	0	0.12 (0.24)	0.22 (0.37)	0.29	1.00
Nanchita	8	3	1.8	1.6	1.22	0	0.12 (0.20)	0.22 (0.26)	0.30	0.18
Tusta	35	1	3.3	1.6	1.19	1	0.08 (0.07)	0.19 (0.20)	0.24	-0.04
Mirasol	5	2	1.2	1.6	1.12	0	0.02 (0.07)	0.12 (0.26)	0.14	0.79
Huacles	12	0	1.9	1.7	1.18	1	0.00 (0.00)	0.00 (0.00)	0.29	0.00
Promedio		2.7	2.3	1.8	1.25	0.5	0.12 (0.16)	0.24 (0.26)	0.30	0.63
Arvenses	33	2	4.0	2.3	1.37	1	0.12 (0.10)	0.37 (0.23)	0.31	0.17
Silvestres	23	4	3.5	2.2	1.34	1	0.19 (0.16)	0.35 (0.29)	0.34	0.25
Promedio total		2.7	2.5	1.9	1.26	0.6	0.12 (0.16)	0.26 (0.27)	0.30	0.62

Por medio de una prueba ANOVA se determinó que las diferencias en el promedio de H_E y H_o por grado de domesticación no son significativas. Para H_E se obtuvo un valor $F= 0.19$ y $p=0.83$ mientras que para la H_o los valores fueron $F= 0.09$ y $p= 0.92$ (ver Anexo 6).

El enrarecimiento de muestra por variedad de chile indica que la variedad Costeño Rojo presenta la mayor riqueza alélica con 4.2 alelos y una tendencia a incrementar si el tamaño de muestra fuera más grande. En el caso de las demás variedades cultivadas el promedio del número de alelos es menor a 3.5, mientras que para arvenses el promedio del número de alelos por locus fue 3.9 y para Silvestres fue 3.3. Sin embargo, si se incrementará el tamaño de muestra de los individuos Silvestres, estos podrían igualar o rebasar el promedio del número de alelos encontrados para arvenses (ver Fig. 20). La riqueza alélica para la categoría cultivada fue de 6.2.

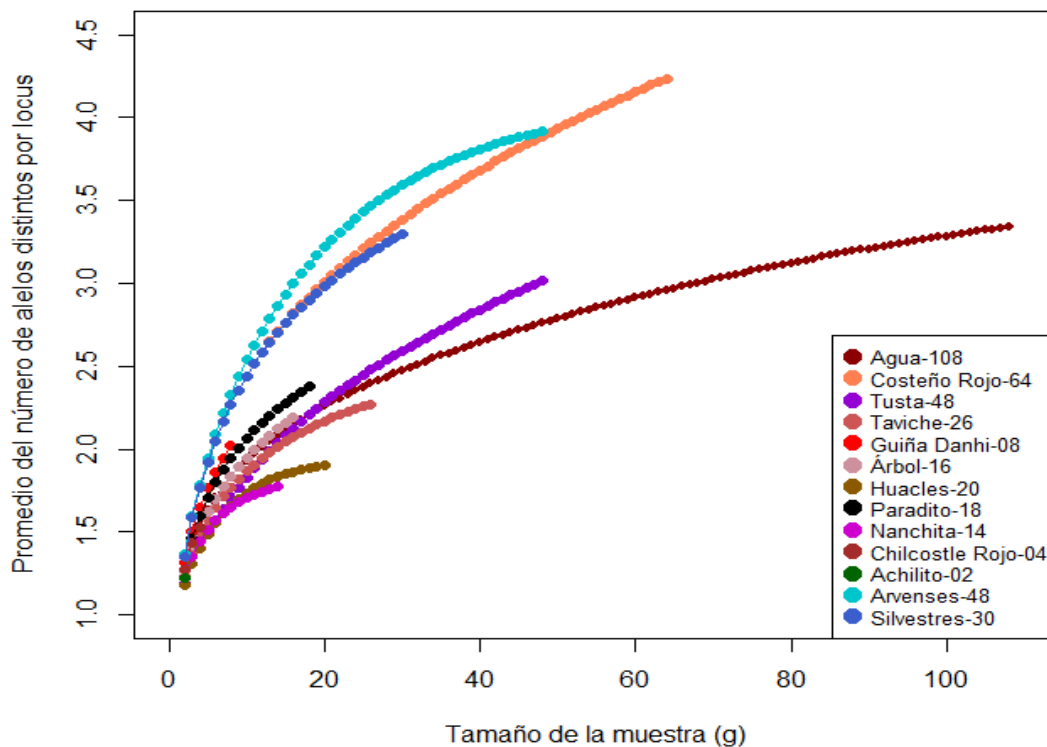


Figura 20. Enrarecimiento de muestra de variedades de chile *C. annuum* en 10 loci de microsatélites (SSR). Se puede apreciar que el promedio de número de alelos máximo es de 4.2 alelos para el chile Costeño Rojo. El número al lado de la población indica el tamaño de la muestra (g).

VII.3.2. Análisis de varianza molecular

El análisis de varianza molecular (AMOVA) para las variedades se realizó formando tres grupos por grados de domesticación: las 14 variedades de Chile cultivadas localmente, los arvenses y los silvestres. Los resultados de la AMOVA indicaron que el mayor porcentaje de variación (92.71%) se encuentra dentro de las variedades (ver Tabla 13).

Tabla 13. Resultados de la prueba de varianza molecular (AMOVA) por grado de domesticación de loci de microsatélites para 16 poblaciones de variedades de *C. annuum* (distancia molecular R_{ST}). Significancia igual a 0.05. Se analizó un grupo de 14 variedades de chiles domesticados, otro grupo de chiles arvenses y un grupo de chiles silvestres.

Fuente de variación	g.l.	Suma de cuadrados	Componentes de la varianza	Porcentaje de variación
Entre grupos	2	425.887	1.80 Va	6.30
Entre poblaciones dentro de los grupos	13	436.239	0.21 Vb	0.73
Dentro de las poblaciones	618	16393.155	26.53 Vc	92.71
Total	633	17255.281	28.53	

Los resultados para variedades individuales de F_{IS} se muestran en las Tablas 11 y 12, con un promedio de 0.62. Los valores de F_{IS} por variedad cultivada, arvenses y silvestres no fueron significativos ($F= 0.813$ y $p= 0.47$; ver Anexo 6). Los niveles $F_{ST}= 0.07$ indicaron moderada diferenciación genética (ver Tabla 14).

Tabla 14. Índices de fijación por grado de domesticación. La F_{ST} para variedades cultivadas se obtuvo con el programa FSTAT (Goudet, 2002)

Índices de fijación	
F_{ST}	0.07 $p= 0.22$
F_{SC}	0.007 $p= 0.63$
F_{CT}	0.06 $p= 0$
F_{IT}	0.64 $p= 0$
$F_{STcultivados}$ = 0.16	

VII.3.3. Distancias genéticas y aislamiento por distancia

VII.3.3.1. Distancias genéticas

En el dendograma Neighbor-Joining basado en las distancias genéticas de Goldstein *et al.*, (1995), realizado por variedad se observa que el chile Costeño Amarillo se agrupa separado de las demás variedades de *C. annuum*, las cuales se subdividen en Costeño Rojo y un grupo integrado por todas las demás variedades incluyendo a silvestres y arvenses que se relacionan cercanamente al chile Mirasol, Tusta, Taviche, Achilito y Serrano Criollo (ver Fig. 21a).

En cambio, el dendograma Neighbor-Joining basado en las distancias genéticas de Nei *et al.*, (1983), agrupa a las variedades en dos principales grupos: el primero se integra por los chiles Tusta, Mirasol, arvenses y silvestres, mientras que el segundo grupo se integra por las otras 12 variedades cultivadas (ver Fig. 21b).

En ambos casos se utilizó a *C. frutescens* como grupo externo y los bootstraps indicaron que las ramas se sostienen muy poco.

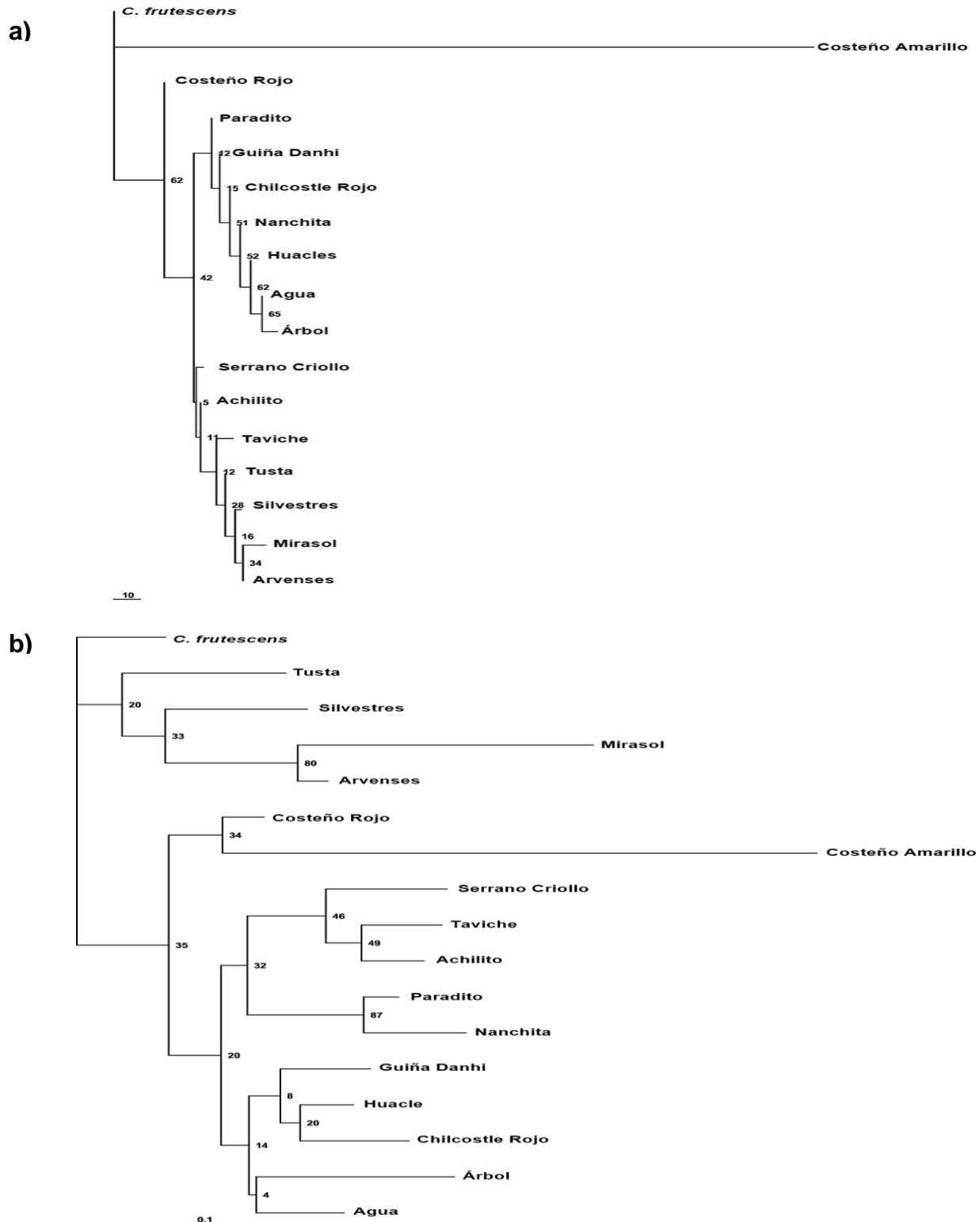


Figura 21. Dendrograma Neighbor-Joining de loci de microsatélites (SSRs) en variedades de *C. annuum*. a) Dendrograma basado en las distancias genéticas de Goldstein *et al.*, (1995) con 8 SSRs, b) Dendrograma basado en las distancias genéticas de Nei *et al.*, (1983) con 10 SSRs. Se presentan los bootstraps.

VII.3.3.2. Aislamiento por distancia

El resultado de la prueba de Mantel que compara la distancia genética con la distancia de domesticación no fue significativo $r= 0.05$; $p= 0.60$ (Fig. 22).

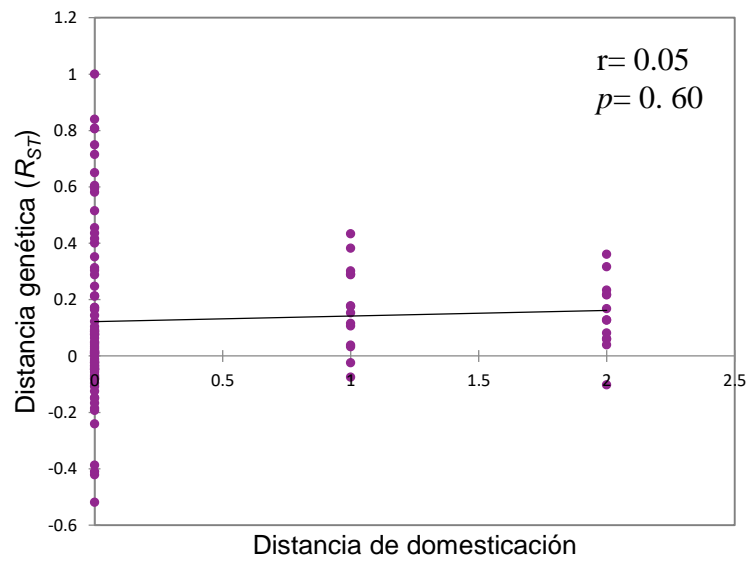


Figura 22. Resultado de la prueba de Mantel que correlaciona distancia genética (R_{ST}) y distancia de domesticación. No existe correlación entre ambas distancias.

VII.3.4. Estructura poblacional

VII.3.4.1. Análisis multivariado (DAPC)

El DAPC (ver Fig. 23) por grado de domesticación diferencia entre individuos cultivados, arvenses y silvestres. La probabilidad de asignación del grupo cultivado fue de 0.97, para los arvenses de 0.70 y los silvestres de 0.57. En el caso de los individuos de *C. frutescens* su probabilidad de asignación fue de 0 mientras que para *C. chinense* la probabilidad fue de 0.67. Así, los individuos de *C. frutescens* se asignaron a los grupos conformados por individuos de *C. annuum*.

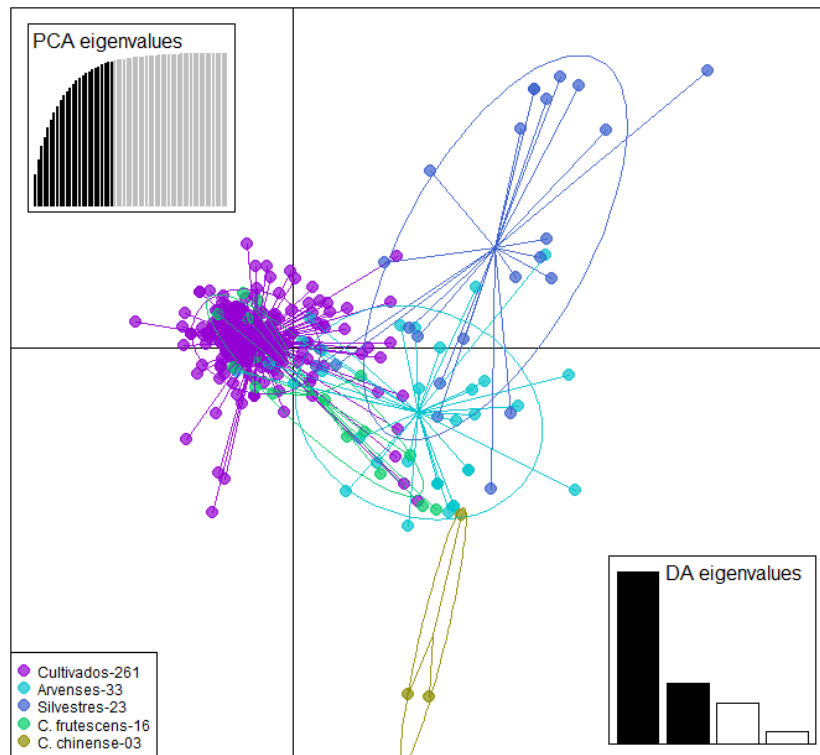


Figura 23. Análisis discriminante de componentes principales (DAPC) por grado de domesticación. Se puede observar que las tres categorías se diferencian una de otra. *C. frutescens* se localiza entre las variedades de *C. annuum* a diferencia de *C. chinense* que se dispersa más de este grupo. El número al lado de la población indica el número de individuos analizados en cada una.

En el DAPC por variedad de Chile se observó que las variedades consideradas como cultivadas se agrupan en un conjunto muy similar a su agrupación cuando las consideras como un solo grupo cultivado, sin embargo, las variedades de Chile Mirasol y Tusta se alejan de este conjunto de variedades cultivadas (ver Fig. 24). El Chile Mirasol es morfológicamente similar a un Chile arvense (ver Anexo 1) y en la Fig. 24 se observa muy cercano a este grupo. En el caso del Chile Tusta se observó que su agrupación la diferencia de las demás especies cultivadas incluso también de las arvenses y silvestres.

La probabilidad de asignación para el chile Tusta fue de 0.80, es decir 80% de sus individuos son exclusivos de Tusta, mientras que en caso de Mirasol su probabilidad fue de 0. El grupo arvense presentó una probabilidad de asignación del 0.61 y el grupo de los silvestres de 0.70. Las probabilidades de asignación y los eig-valores se pueden consultar en el Anexo 8.

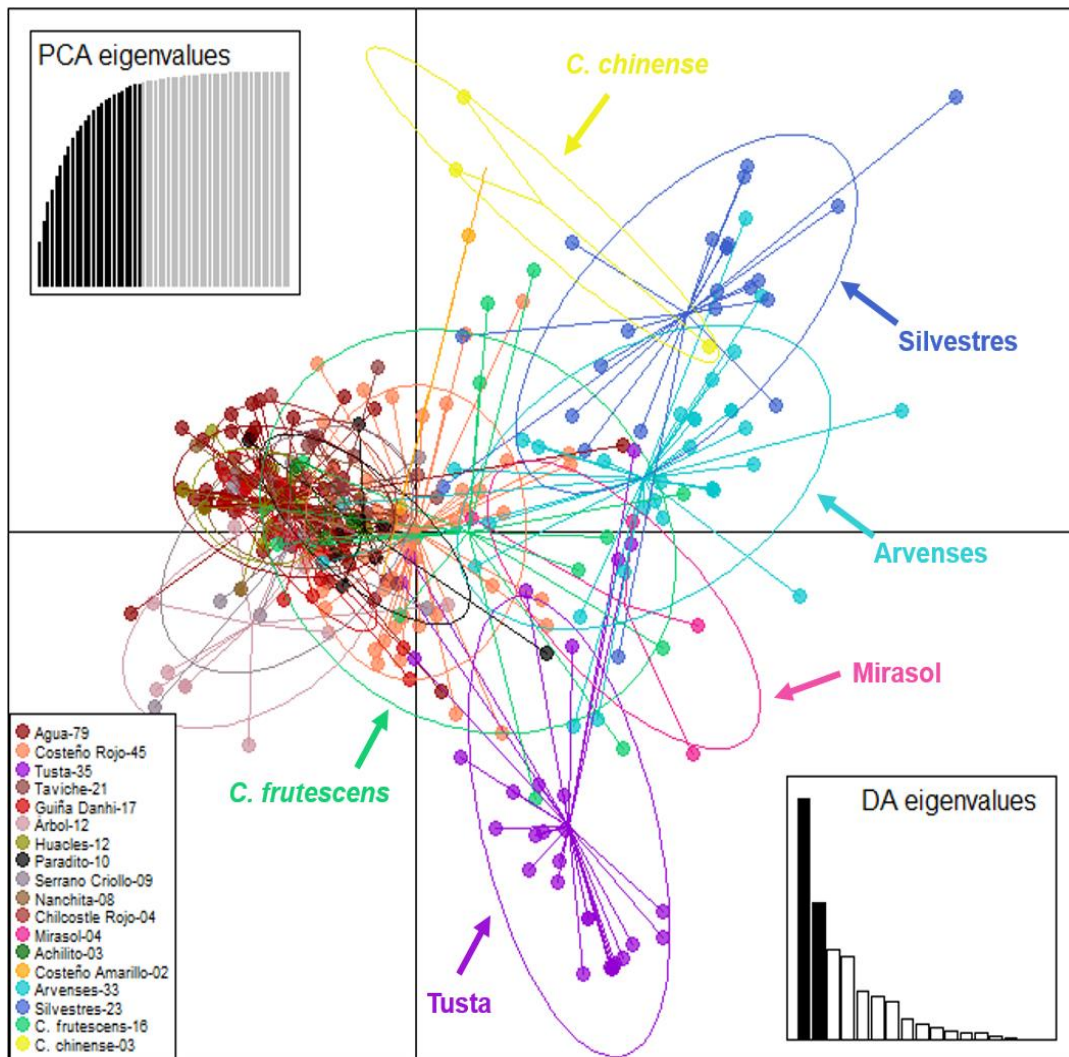


Figura 24. Análisis discriminante de componentes principales (DAPC) por variedad de chile. Se puede observar que el chile Tusta, los arvenses, los silvestres y el chile mirasol (color morado, azul y rosa) se diferencian de las demás variedades cultivadas de *C. annuum*. El número al lado de la población indica el número de individuos analizados en cada una.

VII.3.4.2. Análisis bayesiano (STRUCTURE)

Los resultados por variedad del análisis STRUCTURE y la corrección de Evanno, *et al.*, (2005) nos despliega dos grupos ($K=2$) (ver Fig. 25) con un valor máximo de verosimilitud de $L_n = -3916.05$ y una probabilidad de 0.96 (ver Anexo 9).

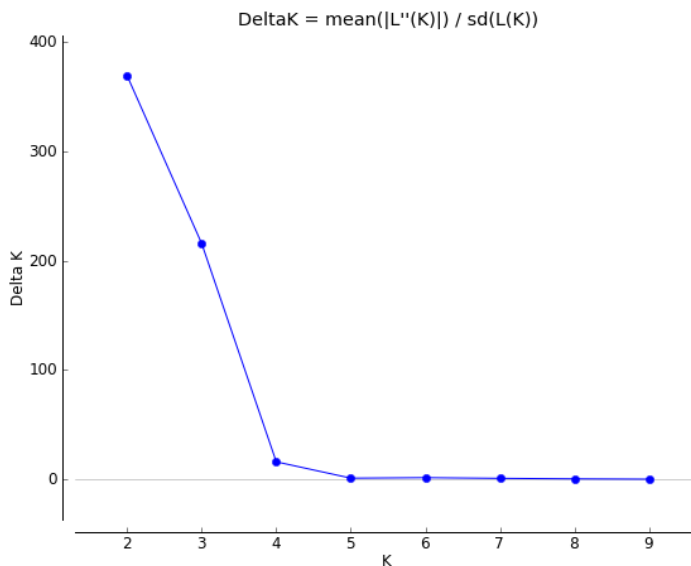


Figura 25. Gráfica de los resultados de la prueba de Evanno, *et al.*, (2005) por variedad. $K=2$

Los datos obtenidos de la probabilidad de asignación a cada grupo nos muestran que se encuentra mayor a 0.5 en los dos grupos (ver Anexo 9). La asignación a los grupos con una $K=2$ coincide en la mayoría de las variedades con la altitud y la categoría de domesticación en la que se ubican los chiles *C. annum*. En el caso del chile Guiña Danhi la probabilidad de asignación al grupo de chiles que coinciden en su cultivo en tierras altas es del 77% mientras que para el Costeño Rojo es del 40%. Por lo tanto, en el caso del chile Costeño Rojo podemos observar la presencia de dos grupos (ver Fig. 26).

En el análisis de una $K= 3$ la primera poza génica agrupa a todas las variedades que se cultivan principalmente en tierras altas junto al Guiña Danhi de la Costa, parte del Costeño Rojo y *C. frutescens*. La segunda poza génica la comparte Costeño Rojo, Arvenses y parte de las Silvestres mientras que el chile Tusta, Mirasol y parte de las Arvenses conforman la tercera poza génica (ver Fig. 26).

En una $K= 4$ podemos observar que las primeras 10 variedades cultivadas se agrupan en dos principales grupos. En los Arvenses y el Costeño Rojo se ubican dos grupos, mientras que los Silvestres y el Tusta se ubican como grupos más homogéneos y en el caso de Tusta se presenta diferente a las demás variedades cultivadas pero similar a Mirasol, Arvenses, *C. frutescens* y *C. chinense* (ver Fig. 26).

La Fig. 27 muestra un STRUCTURE realizado solo para las variedades cultivadas de *C. annuum*. En esta figura se observa que la inclusión de individuos silvestres, arvenses y grupos externos no alteran en gran medida la asignación a los grupos y que esta asignación continúa coincidiendo con la altitud en la que se cultivan los chiles a excepción de Guiña Danhi que se localiza como chile de la Costa, pero es genéticamente más parecido a chiles que se cultivan en tierras altas. En el caso del Costeño Rojo el porcentaje de asignación al grupo que coincide con tierras altas es menor. En el análisis de una $K= 4$ de la Fig. 27 se observa que Tusta y Mirasol se mantienen como un grupo diferente a las demás variedades cultivadas.

En el Anexo 9 se pueden consultar los datos de probabilidad de asignación a los grupos ($K= 2$) y la prueba de Evanno, *et al.*, (2005) para el análisis únicamente por variedad cultivada.

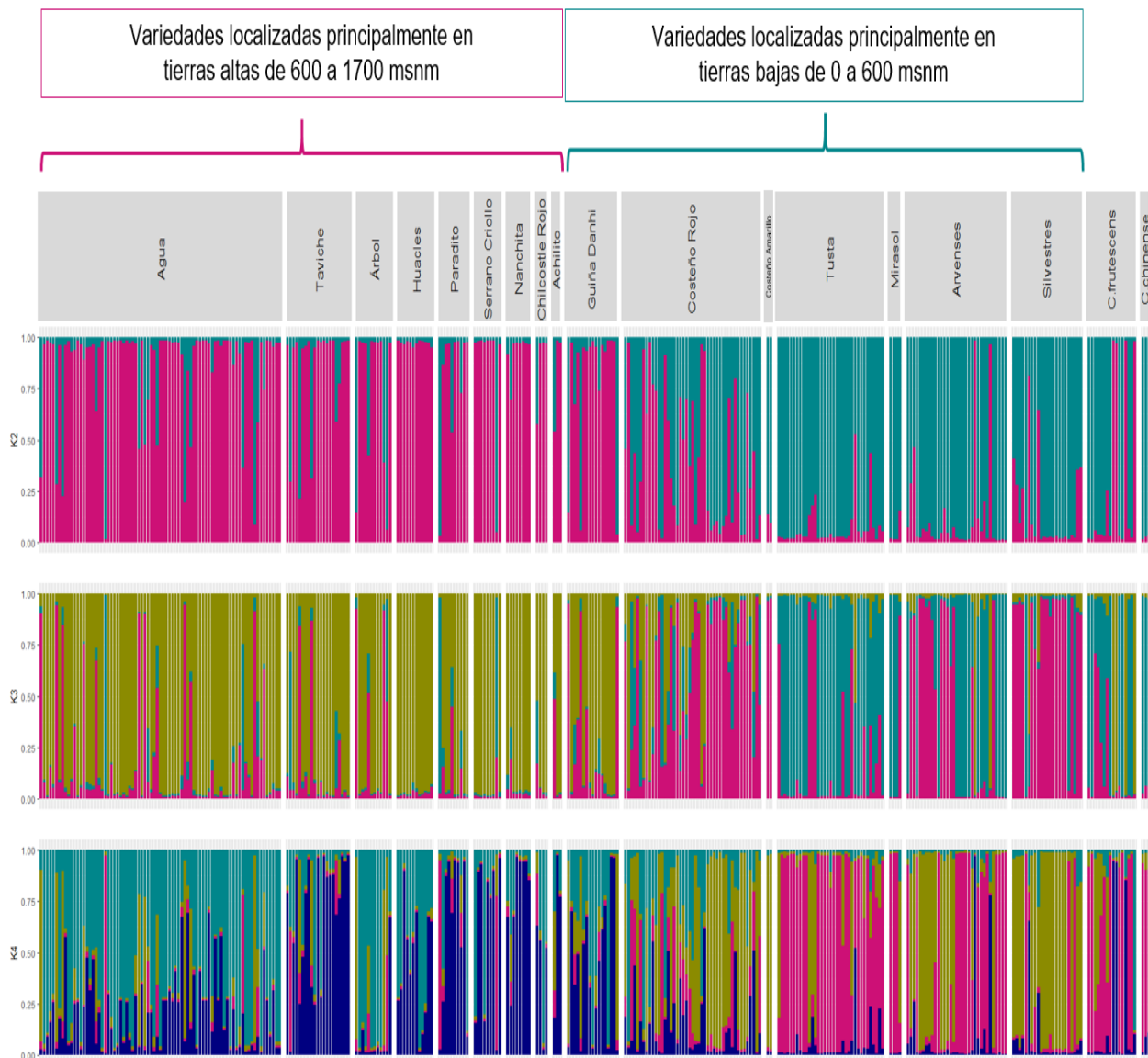


Figura 26. Resultados por variedad obtenidos con el programa STRUCTURE. Se presentan $K=2$, $K=3$ y $K=4$. La estructura de las variedades coincide en la mayoría de los casos con la altitud en la que se cultivan los chiles. Prueba de Evanno, *et al.*, 2005 con $K=2$. El chile Tusta con $K=4$ se va diferenciando de las demás variedades cultivadas y el chile Guña Danhi coincide en agrupación con tierras altas del estado de Oaxaca.

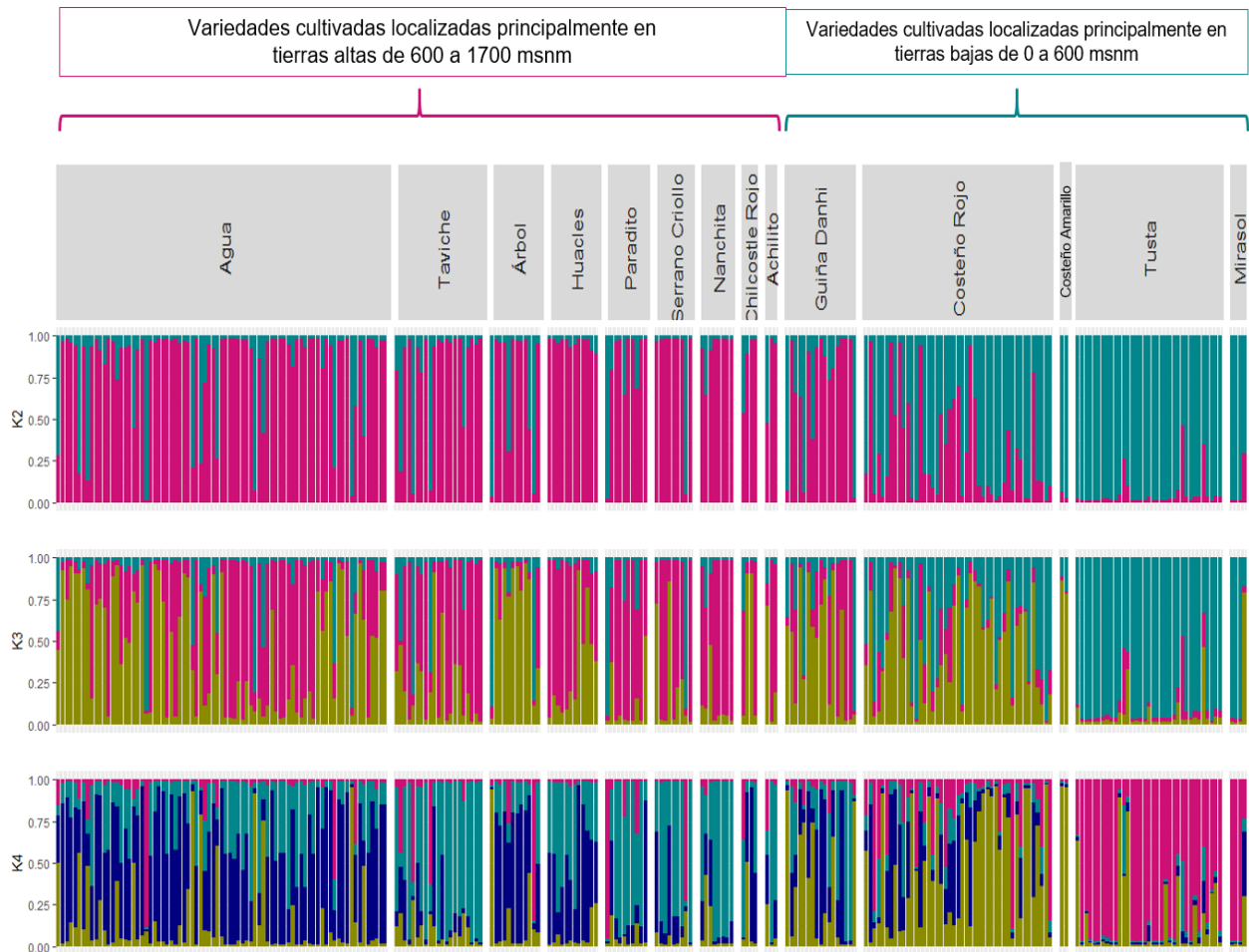


Figura 27. Resultados por variedad cultivada obtenido con el programa STRUCTURE. Se presentan $K=2$, $K=3$ y $K=4$. Se observa que la presencia de poblaciones arvenses y silvestres colectadas principalmente en la Costa no afecta en gran medida la asignación de los grupos que coinciden con la altitud en la que se cultivan los chiles. Prueba de Evanno, *et al.*, 2005 con $K=2$.

Se realizó una correlación entre la altitud en la que se cultiva el chile (a nivel individual) y la asignación de la variedad al grupo 2 ($K=2$). Los resultados arrojaron una correlación positiva entre las variables con una $p=0.0001$ (ver Fig. 28a). Además, es probable que se correlacione también con el grado de domesticación del chile *C. annum* ($p=0.0001$) (ver Fig. 28b).

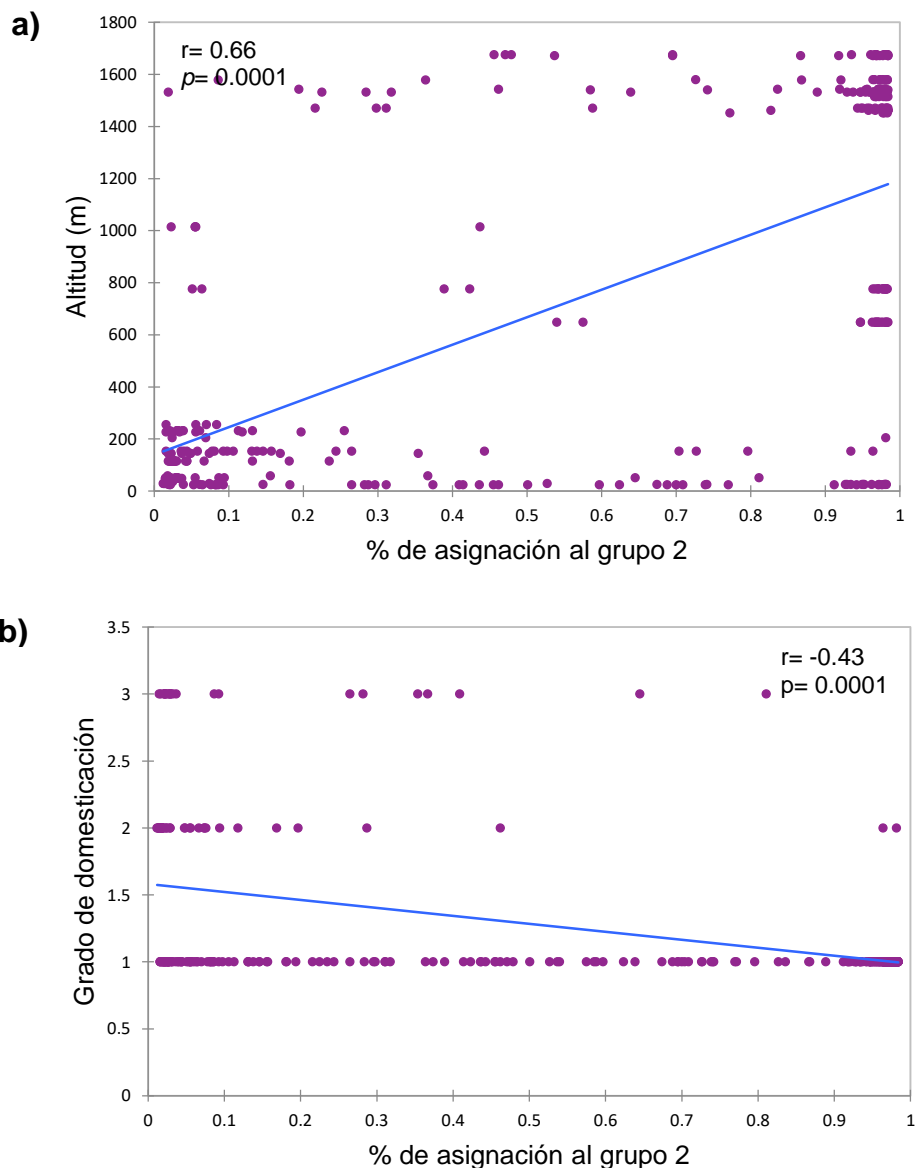


Figura 28. Correlaciones entre el porcentaje de asignación al grupo 2 con: a) la altitud y b) grado de domesticación: 1=cultivado, 2=Arvense y 3=Silvestre. Ambas correlaciones fueron significativas.

La prueba de Bonferroni arrojó valores de $p = 0.0003$ para las tres correlaciones entre la asignación genética al grupo 2 con la altitud de las poblaciones, de las variedades

y el grado de domesticación del chile. Por lo tanto, sí existe una correlación entre estas variables.

VII.3.4.3. Análisis gráfico poblacional

El gráfico poblacional con la mínima topología arrojó 15 nodos y 26 aristas. Los niveles más bajos de covarianza genética dentro de las variedades se encontraron en el chile Tusta y Agua, mientras que los niveles más altos se localizaron en las variedades Mirasol, Huacle, Nanchita y Paradito (ver Fig. 29). Los niveles más altos de covarianza entre poblaciones se encontraron entre el chile Mirasol con el Chilcostle Rojo y Árbol con Tusta. Los niveles más bajos se encontraron entre el chile de Agua con los chiles Huacles y Costeño Rojo.

En este caso no se agregaron parámetros geográficos, debido a que cada variedad puede pertenecer a una o más localidades (ver Fig. 29). Las variedades cultivadas en la región de la Costa como el Costeño Rojo, Tusta y Mirasol (en color rosa) junto con las poblaciones arvenses y silvestres están covariando entre sí, lo mismo ocurre para las variedades cultivadas en la región de la Cañada y los Valles Centrales (en color morado). La variedad Nanchita arrojó el menor número de aristas al relacionarse solo con la variedad Paradito. Solo se utilizó a *C. frutescens* como grupo externo al presentar más de 4 individuos.

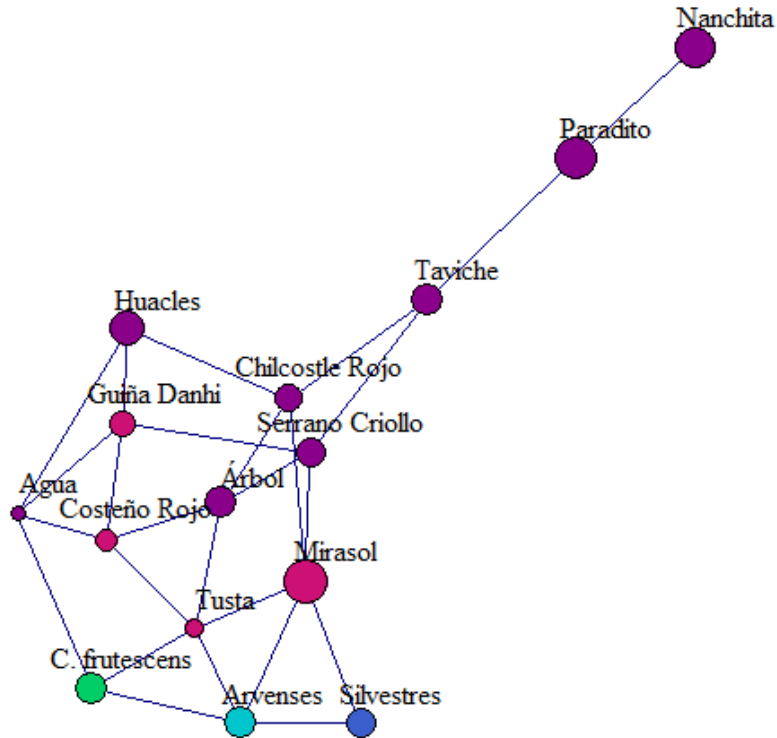


Figura 29. Gráfica poblacional Popgraph (Dyer, 2004; 2015) mostrando las conexiones entre 14 variedades de *C. annuum* en el estado de Oaxaca. El largo de la línea indica la covarianza entre poblaciones y el tamaño del nodo la varianza dentro de las poblaciones. Puntos morados ● representan variedades cultivadas en la Cañada y los Valles centrales; puntos rosas ● variedades cultivadas en la Costa del estado de Oaxaca.

VII.3.4.4. Cuellos de botella y tamaño efectivo poblacional (N_e)

El índice de Garza-Williamson (M) promedio fue de 0.30 ($SD= 0.13$). Las variedades cultivadas de Chile *C. annuum* se ven mayormente afectadas por los cuellos de botella en los modelos de mutación T.P.M y S.M.M (ver Tabla 15) y el índice Garza-Williamson $M= 0.26$ ($SD= 0.13$) (ver Tablas 11 y 12). La categoría silvestre con BOTTLENECK no indica valores para cuellos de botella (ver Tabla 15), sin embargo, el índice M fue de 0.34 ($SD= 0.12$) (ver Tabla 11 y 12). La población arvense tiene presencia de cuellos de botella

con valores de $p= 0.03$ y $p= 0.02$ para la prueba Wilcoxon con dos colas y los modelos de mutación dos fases (T.P.M.) y mutación paso a paso (S.M.M.) (ver Tabla 15) y el índice Garza-Williamson fue de $M= 0.31$ ($SD= 0.14$) (ver Tabla 11 y 12). Índices de Garza-Williamson menores a 0.68 indican presencia de cuellos de botella. Los cuellos de botella a nivel de variedades cultivadas indicaron que las variedades Tusta y Mirasol presentaron los niveles más bajos de M (0.24 y 0.14) mientras que el Chilcastle Rojo, Serrano Criollo y Taviche presentaron los niveles más altos ($M= 0.43$, $M= 0.38$ y $M= 0.34$), sin embargo, los tamaños de muestra son diferentes entre estas variedades (ver Tabla 12).

El tamaño efectivo de la población cultivada, arvense y silvestre fue de $N_e= 14.7$, $N_e= 28.2$ y $N_e= 16.0$ (Ver Tabla 15). Los intervalos de confianza JackKnife para cultivados es de 3.9 a 31.7, para arvenses fue valores mayores a 9.4 y para arvenses fue de 6.0 a 111.9.

Tabla 15. Cuellos de botella y tamaño efectivo según el grado de domesticación de chiles *C. annuum* en 10 loci de microsatélites (SSRs). I.A.M.= modelo de alelos infinitos, T.P.M.= modelo de mutación dos fases, S.M.M.= modelo de mutación paso a paso y N_e = estimación del tamaño efectivo de la población con un valor crítico $P_c= 0.02$. Varianza para T.P.M.= 5% y probabilidad para I.A.M. y S.M.M.= 95%. (*) = desviación significativa del equilibrio ($p < 0.05$), p = probabilidad de la prueba significativa y pW = probabilidad obtenida de la prueba Wilcoxon de dos colas.

Modelos	Cultivados		Arvenses		Silvestres	
	p	pW	p	pW	p	pW
I.A.M.	0.27	1.0	0.54	0.77	0.41	0.77
T.P.M.	0.002*	0.002*	0.08	0.03*	0.25	0.28
S.M.M.	0.002*	0.002*	0.08	0.02*	0.25	0.28
N_e	14.7		28.2		16.0	

VIII. DISCUSION

La diversidad y estructura genética de las poblaciones de Chile *C. annuum* en México han sido analizadas con distintos marcadores moleculares (Loaiza-Figueroa, *et al.*, 1989; Hernández-Verdugo, *et al.*, 2001; Oyama, *et al.*, 2006; Aguilar-Meléndez, *et al.*, 2009; Contreras, *et al.*, 2011; González-Jara, *et al.* 2011; Pacheco-Olvera, *et al.*, 2012; Toledo-Aguilar, *et al.*, 2016). Sin embargo, se ha recalcado la importancia de estudios de diversidad genética que involucren el manejo y las prácticas de los agricultores para fines de preservación de los recursos genéticos del Chile *C. annuum* (Pickersgill, 2016). En este estudio se analizó la diversidad y estructura genética de 10 loci de microsatélites (SSRs) en 28 poblaciones y 16 variedades de Chile *C. annuum* en el estado de Oaxaca, lugar donde la especie ha sido sometida a diversos sistemas de manejo y a condiciones ambientales y topográficas diversas (Caballero, *et al.*, 2004; Jardón-Barbolla, 2017).

VIII.1. Diversidad genética

El promedio del número de alelos encontrados en este estudio con el método de rarefacción de la muestra es similar al rango reportado (3.4 a 22.8 alelos) para la especie *C. annuum* mediante el uso de microsatélites SSRs, a excepción de la región de la Cañada con un promedio de 2.6 alelos (Contreras, *et al.*, 2011; González-Jara, *et al.* 2011; Pacheco Olvera, *et al.*, 2012; Toledo-Aguilar, *et al.*, 2016; ver Tabla 16). Los estudios de González-Jara *et al.*, (2011) y Pacheco Olvera *et al.*, (2012), han involucrado el análisis no solo de individuos cultivados, sino también de poblaciones silvestres. En ambos casos las poblaciones silvestres presentan un mayor número de alelos en comparación a su contraparte cultivada. Sin embargo, en el presente estudio se encontró

que la variedad Costeño Rojo presentó un promedio del número de alelos mayor (4.7 alelos) a las silvestres y arvenses.

Los SSRs se han descrito como ampliamente exitosos en explorar la diversidad genética de *Capsicum* (Zhang, *et al.*, 2016). En este estudio se encontró una diversidad genética total para las poblaciones de *C. annuum* de $H_T = 0.36$ por región geográfica y grado de domesticación. Estos valores fueron más altos que los obtenidos por Oyama *et al.*, (2006) en la especie *C. annuum* mediante el análisis de RAPDs ($H_T = 0.17$ para silvestres y $H_T = 0.13$ para cultivados). Por otro lado, Hernández-Verdugo *et al.*, (2001), analizó la diversidad genética de *C. annuum* con 9 isoenzimas obteniendo niveles más altos de H_T a los encontrados en este estudio ($H_T = 0.50$ para silvestres y $H_T = 0.53$ para cultivados) mientras que Loaiza-Figueroa *et al.*, (1989) encuentra niveles más bajos ($H_T = 0.28$ para silvestres, $H_T = 0.08$ para semi-silvestres y $H_T = 0.18$ para cultivados), sin embargo, su análisis incluye diferentes especies del género *Capsicum*. Esto parece indicar que el uso de microsatélites (SSRs) e isoenzimas parecen ser prácticos en la exploración de la diversidad genética de la especie *C. annuum*.

La diversidad genética promedio para las poblaciones de *C. annuum* fue $H_E = 0.25$ por regiones geográficas, $H_E = 0.25$ para las variedades y de $H_E = 0.35$ por grado de domesticación. Estos valores fueron más bajos a los obtenidos en otros trabajos con microsatélites (Contreras, *et al.*, 2011; González-Jara, *et al.* 2011; Pacheco-Olvera, *et al.*, 2012; Yumman, *et al.*, 2012; Nicolai, *et al.*, 2013; Toledo-Aguilar, *et al.*, 2016) y al valor de $H_E = 0.61$ reportado para plantas por Nybom, (2004). Sin embargo, es necesario tener en cuenta que los trabajos con microsatélites que exploran la diversidad genética de la especie *C. annuum* en México se han realizado en diferentes regiones y estados del país (ver Tabla 16). En cambio, este trabajo fue realizado solo para un estado en tres regiones

de las ocho reportadas para Oaxaca y con poblaciones relativamente cercanas entre sí. Por lo tanto, los valores de H_E obtenidos pueden considerarse como una aproximación a la diversidad genética que se localiza en el estado.

Tabla 16. Estudios de diversidad y estructura genética para la especie *C. annuum* en México. M= marcador genético, ISO= isoenzimas, GD= grado de domesticación, S= silvestres, SS= semisilvestres, C= cultivados, npb= número de poblaciones, N= número de alelos promedio, H_E = heterocigosis esperada, F_{IS} = coeficiente de endogamia, F_{ST} = diferenciación genética.

Referencia	M	Región geográfica	GD	npb	N	H_E	F_{IS}	F_{ST}
Hernández-Verdugo, <i>et al.</i> , 2001	9 ISO	Noroeste de México	S	10	2.7	0.45	0.34	0.04
			SS	1	2.5	0.46	-	-
			C	3	2.6	0.41	0.35	0.21
Oyama, <i>et al.</i> , 2006	60 RAPDs	Sonora, Sinaloa y Nayarit	S	15	-	0.07	-	-
			C	3	-	0.08	-	-
Contreras, <i>et al.</i> , 2011	19 SSRs	Puebla y Zacatecas	C	63	5.5	0.40	-0.15	0.11
González-Jara, <i>et al.</i> , 2011	9 SSRs	Yucatán, Sierra Madre Oriental, Antiplano Zacatecano Potosino, Costa del Pacífico y Sonora	S	11	22.8	0.80	0.62	
			SS	6	15.7	0.75	0.53	0.43
			C	10	14.0	0.79	0.78	
Pacheco-Olvera, <i>et al.</i> , 2012	7 SSRs	Noroeste de México (Nayarit, Sonora y Sinaloa)	S	12	3.6	0.47	0.41	0.40
			C	3	3.4	0.42	0.54	0.30
Toledo-Aguilar, <i>et al.</i> , 2016	24 SSRs	Guanajuato, Zacatecas, Durango, San Luis Potosí, Puebla, Oaxaca y Yucatán	C	38	9.2	0.36- 0.59	-0.02- (0.49)	-0.50- (0.17)

El mayor nivel de diversidad genética H_E se encontró en la región de la Costa ($H_E=0.28$) a comparación de lo obtenido para Valles Centrales ($H_E= 0.21$) y la Cañada ($H_E=0.24$). Por otra parte, la diversidad genética para variedades cultivadas ($H_E= 0.24$) fue menor a la obtenida para los silvestres y arvenses ($H_E= 0.35$ y $H_E= 0.37$ respectivamente). Sin embargo, en ambos casos los valores de H_E obtenidos no difirieron significativamente, sugiriendo que la H_E es igual en las tres regiones y en las tres categorías de domesticación.

Se ha planteado que la deriva génica ha sido una de las principales causas de reducción de la diversidad genética de muchas plantas cultivadas que suelen contener solo una pequeña fracción de la variación existente en su progenitor silvestre (Brush, 2004; Ladizinsky, 2012). En el caso del chile *C. annuum* se han reportado valores de reducción de la diversidad del 8% al 17% en relación a su pariente silvestre (Hernández-Verdugo, *et al.*, 2001; Oyama, *et al.*, 2006; Aguilar-Meléndez, *et al.*, 2009; Pacheco-Olvera, *et al.*, 2011). En este estudio se encontró una reducción de la diversidad genética (H_E) del 6% en relación a los parientes silvestres y una reducción del 11% en relación a las poblaciones arvenses. Estos valores son bajos si se considera que especies como el maíz llegan a reducir aproximadamente el 40% de su diversidad genética en relación a su pariente silvestre más cercano, el teocintle (Doebley, *et al.*, 2006). Como discutiré más adelante, la conexión entre las poblaciones silvestres y las poblaciones cultivadas a través de las poblaciones de traspatio, podría estar jugando un papel importante en el patrón observado de los datos.

Al analizar las reducciones de diversidad genética por cada variedad podemos observar que en relación a su pariente silvestre, los chiles Costeños no presentan reducciones de su diversidad. Las reducciones genéticas para las variedades son: 11%

Guiña Danhi, 20% Chilcostle Rojo y Paradito, 23% Agua y Árbol, 34% Serrano criollo y Taviche, 37% Achilito y Nanchita, 46% Tusta y 66% Mirasol. En relación a los individuos arvenses las reducciones fueron: 3% Costeño Rojo, 5% Costeño Amarillo, 16% Guiña Danhi, 24% Chilcostle Rojo y Paradito, 27% Agua y Árbol, 38% Serrano criollo y Taviche, 41% Achilito y Nanchita, 49% Tusta y 68% Mirasol. Sin embargo, es necesario tomar en cuenta que el tamaño de la población es diferente para cada variedad, por ejemplo, las variedades de Chile Costeño Amarillo, Chilcostle Rojo, Paradito, Serrano Criollo, Achilito, Nanchita y Mirasol presentan menos de 10 individuos.

La reducción de alelos en relación al tamaño mínimo de muestra (riqueza alélica) y el número de alelos totales fue hasta del 50% para los análisis por regiones geográficas y variedades cultivadas. Estas reducciones en la diversidad alélica han sido observadas para poblaciones cultivadas de *C. annuum* por González-Jara, *et al.*, 2011, atribuyendo esta reducción a fuertes cuellos de botella actuando en las poblaciones. Por otra parte, los chiles arvenses y silvestres presentaron menores niveles de reducción en la diversidad alélica. Este resultado concuerda con lo obtenido del análisis de cuellos de botella (BOTTLENECK), el índice de Garza-Williamson (2001), y el tamaño efectivo de la población, al encontrar que las poblaciones silvestres y arvenses se encuentran menos afectadas por la deriva génica.

El número de alelos privados totales que se encontró fue bajo ($N_p = 11$). En estudios con microsatélites se han reportado entre 9 y 59 alelos privados en *C. annuum* (Contreras, *et al.*, 2011; González-Jara, *et al.* 2011; Toledo-Aguilar, *et al.*, 2016). Este bajo número de alelos puede deberse al muestreo y tamaño de las poblaciones, los cuales se ha indicado afectan la probabilidad de muestreo de alelos raros (Mohammadi y Prassanna, 2003). En este estudio 13 poblaciones y 6 variedades presentaron menos

de 10 individuos por población, mientras que Mohammadi y Prassanna (2003) indican que un número entre 48 y 59 individuos son suficientes para tener un 95% de probabilidad de detectar todos los alelos.

A pesar de lo anterior, la distribución espacial de los alelos privados parece ser informativa respecto a los procesos que subyacen en términos del manejo. Los alelos privados registrados se encontraron en la población silvestre de Santo Domingo Ingenio, en la arvense de Juan Diegal y en las variedades cultivadas de chiles Costeños, Guiña Danhi y Tusta en las poblaciones Añil-Lagartero, Rancho Llano, y Juan Diegal de la región de la Costa. El alelo privado de los Valles Centrales se localizó en el chile de Agua de Mixtepec. La población Cuicatlán de la Cañada, presenta solo un alelo privado en el chile Huacle aunque este no presentó loci polimórficos. Los alelos privados han sido atribuidos al manejo y selección artificial de la región, y a la distancia geográfica y al bajo intercambio de semillas o morfotipos que permiten los alelos sean preservados en cada región (Toledo-Aguilar, *et al.*, 2016). Los chiles domesticados en México no suelen ser intercambiados en un amplio rango geográfico (Aguilar-Meléndez, *et al.*, 2009).

En Oaxaca, los chiles que presentaron alelos privados son característicos de cada región, por ejemplo, el chile de Agua es considerado endémico de los Valles Centrales y el Huacle de la Cañada (Montaño-Lugo, *et al.*, 2014; García-Gaytán, *et al.*, 2017), indicando que existe poco intercambio de semillas o morfotipos entre las regiones analizadas y coincide con la moderada diferenciación genética que se encontró a nivel geográfico ($F_{ST} \text{ c/gpos} = 0.10$ y $F_{ST} \text{ sin/gpos} = 0.09$) y por grado de domesticación ($F_{ST} = 0.07$) con el análisis de varianza molecular AMOVA. Los valores de F_{ST} entre 0.05 a 0.15 son considerados como moderada diferenciación genética, mientras que valores

superiores se consideran alta diferenciación entre poblaciones (Conner y Harl, 2004). La diferenciación con el análisis SAMOVA también fue moderada ($F_{ST}= 0.11$).

La mayor cantidad de variación genética ocurre en las poblaciones, conduciendo a altos niveles de endogamia con una $F_{IS}= 0.32$ por región geográfica y $F_{IS}= 0.62$ para variedades. Una alta cantidad de diversidad genética dentro de las poblaciones sugiere menos divergencia permitiéndoles una mayor flexibilidad y resiliencia a resistir situaciones de estrés y variabilidad ambiental (Brush, 2004). A nivel de región geográfica es posible encontrar menores niveles de endogamia. Esto es resultado de la interacción de diferentes variedades de Chile *C. annuum* en una sola población. La subdivisión de una población es una de las principales causas de un aparente déficit de heterocigos en las poblaciones, lo que se conoce como efecto Wahlund (Allendorf y Luikart, 2009; Hedrick, 2011). Este efecto puede estar presente tanto para las poblaciones por región geográfica como para las variedades de Chile.

Los niveles de endogamia $F_{IS}= 0.64$ localizados en plantas cultivadas son consecuencia de su sistema de apareamiento autogámico que conduce a un limitado intercambio de genes entre poblaciones (Pacheco-Olvera, *et al.*, 2012). Todas las especies de *Capsicum* son facultativamente autocompatibles al presentar tasas de alogamia con valores del 7 al 90% (MacLeod, *et al.*, 1983; Pickersgill, 1997; Djian-Caporalino, *et al.*, 2006), este sistema potencialmente autogámico parece exacerbarse en las poblaciones cultivadas. La excepción es *Capsicum cardenasii* una especie silvestre del noreste de Bolivia relacionada cercanamente a *C. pubescens* (grupo de flores moradas) y considerada como autoincompatible (MacLeod, *et al.*, 1983; Onus y Pickersgill, 2004; Long-Solís, 2012).

En la especie *C. annuum* se han reportado tasas de entrecruza de 1 a 46% y dependerá de si coexiste con sus polinizadores, la localidad, el ambiente, el espacio entre plantas y la variedad de chile (Franceschetti 1971; Nuez, *et al.*, 2003). La moderada diferenciación genética que se encontró puede deberse al factor de la polinización cruzada al observarse niveles de endogamia más bajos para silvestres $F_{IS}= 0.25$ y arvenses $F_{IS}= 0.17$. Aunque los valores de F_{IS} no fueron significativamente diferentes, se ha reportado que los individuos silvestres de *C. annuum* son polinizados principalmente por aves e insectos (Hayano-Kanashiro, *et al.*, 2016). El flujo de genes por aves puede ocasionar la pérdida de diferenciación genética (Aguilar-Meléndez, *et al.*, 2009). Este constante flujo de genes llevado a cabo por las aves y el bajo nivel de intercambio de semillas y morfotipos pueden llevar a niveles intermedios de diferenciación entre las regiones analizadas.

Los dendogramas de Neighbor-Joining obtenidos para las poblaciones y variedades con las distancias $\delta\mu^2$ y D_A difirieron en sus topologías. Para el caso de las poblaciones la topología obtenida mediante la distancia D_A separó casi en su totalidad poblaciones de la Costa y poblaciones de los Valles Centrales y la Cañada, mientras que la topología obtenida mediante $\delta\mu^2$ muestra un grupo integrado principalmente por poblaciones arvenses, silvestres y algunos individuos de chile Mirasol y Tusta. Estas poblaciones se localizan geográficamente en los extremos suroeste y sureste de la Costa. Probablemente, la polinización cruzada de los chiles silvestres llevada a cabo por aves puede estar ocurriendo en esta dirección, ocasionando que la similitud genética entre estas poblaciones sea mayor.

La topología obtenida para las variedades de Chile mediante D_A separa claramente entre variedades cultivadas (a excepción de Tusta y Mirasol) de los chiles silvestres y arvenses. Aunque la topología obtenida con $\delta\mu^2$ relaciona a todos los chiles *C. annuum* como un grupo y los Costeños como otros grupos, en ambas topologías las relaciones cercanas entre Tusta, Mirasol, arvenses y silvestres son muy similares. Esta relación (que retomaré más adelante), resulta interesante pues Tusta es una variedad cultivada con una morfología diferente al silvestre, caso contrario al Mirasol que presenta un fruto muy similar al silvestre (ver Anexo 1).

Las distancias genéticas $\delta\mu^2$ son más robustas a fluctuaciones en el tamaño poblacional (que observamos en este estudio) y son útiles estimando los tiempos de divergencia (Takezaki y Nei, 1996; Cooper, *et al* 1999). Aunque fueron desarrolladas especialmente para el análisis de microsatélites siguiendo el modelo de mutación S.M.M., se ha reportado que presentan muy poca eficiencia en la construcción filogenética por su gran error de muestreo, mientras que D_A se considera más eficiente sobre todo cuando se tiene un bajo número de microsatélites (10 loci), ya que se ha estimado que se necesitarían más de 700 loci para que $\delta\mu^2$ sea eficiente (Takezaki y Nei, 2008). En este estudio se observó que la topología obtenida por la distancia D_A parece ser más informativa tanto a nivel geográfico como a nivel categoría de domesticación. Esta topología, como discutiré más adelante, resulta muy similar a la estructuración genética obtenida mediante el programa STRUCTURE.

Si bien los valores de bootstraps que obtuvimos fueron menores al 50% es necesario tomar en cuenta que la reconstrucción de filogenias mediante microsatélites tiene limitaciones importantes relacionadas con irregularidades y asimetrías en el proceso

de mutación, pues varía considerablemente la tasa de mutación de un locus a otro, lo cual ocasiona un incremento en la varianza de las distancias genéticas (Takezaki y Nei, 1996; Goldstein y Pollock, 1997).

Por otra parte, las pruebas de Mantel con R_{ST} no indican aislamiento por distancia ni a nivel poblacional ni a nivel de transecto representado por tierras altas (Valles Centrales y Cañada) y tierras bajas (Costa) incluyendo, o no, a individuos pertenecientes a *C. frutescens* y *C. chinense*. Sin embargo, las pruebas de Mantel con R_{ST} indicaron probabilidades del 23%, 43%, 46%, 57% y 64% de rechazar aislamiento por distancia cuando ésta existe, mientras que una probabilidad de 0.52% para la Costa de Oaxaca al incluir individuos de *C. frutescens* y *C. chinense* (Valles Centrales y Cañada) indicaría que el aislamiento por distancia parece no existir. No obstante, las pruebas basadas en las distancias genéticas R_{ST} pareadas presentan problemas que limitan su poder al estudiar la genética del paisaje, al no tomar en consideración que el flujo génico no es lineal y que puede estar influenciado por variables ecológicas (Dyer, *et al.*, 2010). Lo anterior indica la necesidad de emplear diferentes herramientas que también involucren variables ecológicas permitiéndonos tener un mejor entendimiento de la evolución y el proceso de domesticación de las plantas cultivadas (Milla, *et al.*, 2015).

Dyer *et al.*, (2010), propone las distancias de covarianza genética cGD como medidas más sensibles para entender cómo el flujo de genes interactúa con variables espaciales y ecológicas a nivel de paisaje. Estas distancias genéticas nos indicaron que sí existe aislamiento por distancia en las poblaciones analizadas. El aislamiento por distancia está asociado a efectos de deriva génica, pequeños tamaños efectivos y distancias de migración limitadas (Bradburd, *et al.*, 2013; Sexton, *et al.*, 2014). Estos

efectos fueron observados principalmente en tierras altas de Oaxaca (Valles Centrales y la Cañada).

El aislamiento por ambiente nos permite entender qué factores ambientales (altitud, humedad, hábitat, tipo de sustrato, temperatura, precipitación o densidad) influyen en el flujo de genes y la estructura poblacional, casi siempre relacionados a la selección al promover la adaptación local (Sexton, *et al.*, 2013; Wang y Bradburg, 2014). En este estudio la variable ambiental que analizamos fue la altitud en la que se localizan las poblaciones de Chile sugiriendo que posiblemente está influenciando el aislamiento por distancia y la diferenciación entre las poblaciones.

La altitud parece ser un factor que determina los patrones de diversidad genética de los chiles *C. annuum* en México. En 2011 González-Jara, y colaboradores, encuentran disminuciones en la diversidad genética de zonas bajas hacia zonas altas del país. Por otro lado, Toledo Aguilar, *et al.*, (2016), encuentran que los chiles que se localizan en la zona Sur y Centro de México se diferencian de las poblaciones del Norte, probablemente por la distancia geográfica y su morfología. En este estudio no se encontraron niveles de diversidad genética significativamente diferentes entre las regiones. Sin embargo, se encontró una tendencia negativa entre la correlación de la diversidad genética con la altitud a la que se cultivan los chiles, lo cual nos indica que niveles altos de diversidad genética (H_E) se asocian a una baja altitud, mientras que una mayor altitud se asocia a niveles más bajos de diversidad genética. También, es posible relacionar una mayor variación en zonas donde el cultivo es más antiguo, es decir, cerca del área de distribución de su pariente silvestre (Vavilov, 1926; Harlan, 1975; Jardón-Barbolla, 2015), tal es el caso de la región de la Costa donde fue posible encontrar el mayor número de individuos silvestres en relación a Valles Centrales y la Cañada. Estos individuos por

medio del flujo de genes llevado a cabo por las aves, contrarrestaría los efectos de la deriva génica y la selección (Slatkin, 1987; Montes y Eguiarte, 2002) y conduciría a disminuir su diferenciación poblacional.

Las mayores reducciones poblacionales estimadas mediante BOTTLENECK se encontraron en las poblaciones cultivadas de *C. annuum*, principalmente de la Cañada y Valles Centrales (tierras altas), que podrían estar involucradas en el aislamiento detectado por el método de Gráfica de poblaciones (Dyer, *et al.*, 2010). Además, se ha reportado a la región de los Valles Centrales como parte de las zonas más afectadas en la reducción del bosque tropical caducifolio, ecosistema característico del estado de Oaxaca (Meave, *et al.*, 2012), conduciendo a que el tamaño efectivo de los silvestres sea mucho menor y a que exista un alto efecto de cuello de botella, endogamia y un reducido flujo de genes en las variedades cultivadas (Young, *et al.*, 1996). Pequeños tamaños efectivos aumentan el aislamiento por distancia, especialmente en el caso de la fragmentación del hábitat resultando en erosión genética (Young, *et al.*, 1996; González-Jara, *et al.*, 2011). El flujo génico entre regiones contribuiría a disminuir el efecto de la fragmentación del hábitat y la extinción de variedades locales en la zona (González-Jara, *et al.*, 2011). Los Valles Centrales y la Cañada se conforman principalmente de variedades cultivadas que se han adaptado a las condiciones ambientales propias de estas regiones (climas menos estacionales y de mayor altitud).

VIII.2. Estructura genética

El análisis discriminante de componentes principales (DAPC) por poblaciones nos indicó que la variación genética existente es muy similar entre las diferentes poblaciones. La

única excepción es para Santo Domingo Ingenio, ya que su variación genética es diferente a las otras 27 poblaciones sugiriendo que la poza génica de los individuos silvestres y arvenses que se localizan en esta localidad no se parece a las demás poblaciones (la probabilidad de asignación de los individuos a la población fue del 80%). El resultado de STRUCTURE poblacional con una $K= 3$ y $K= 4$ muestra un patrón similar en esta población. Su asignación genética coincide en mayor parte con la presencia de individuos silvestres en las poblaciones. Es importante señalar que, a diferencia de otras localidades, en Santo Domingo Ingenio se encontraron solamente chiles silvestres y un arvense y que los reportes de los informantes locales señalan que las variedades domesticadas se comenzaron a cultivar solo recientemente a una escala relativamente reducida. Ello explicaría parte de la especificidad genética de la población.

Santo Domingo Ingenio se localiza en la zona del Istmo de Tehuantepec de la Costa de Oaxaca. Esta zona junto a la cuenca del río Tehuantepec, y la región suroeste de la Costa se han propuesto como centros prioritarios de conservación del bosque tropical caducifolio al contar con una de las mayores extensiones de remanentes de este bosque (Meave, *et al.*, 2012). En estas zonas es posible ubicar chiles silvestres en fragmentos de vegetación original (bosque tropical caducifolio) que son intencionalmente dejados por los campesinos formando cercas vivas relativamente anchas entre los campos de cultivo. Sin embargo, estas poblaciones no solo se ven afectadas por la fragmentación y deforestación de bosques, los huracanes y sequías (Meave, *et al.*, 2012; Narez-Jiménez, *et al.*, 2014), sino también por el impacto ambiental que genera la introducción de parques eólicos en la zona del Istmo de Tehuantepec (Jara, 2011).

Tras la llegada de los parques eólicos al Istmo la percepción pública sobre el índice de aves que mueren a causa de las aspas de los aerogeneradores se ha elevado (Jara,

2011; Agatón, *et al.*, 2016). Esto debe tomarse en consideración, ya que los chiles silvestres (*Capsicum annuum* var *glabriusculum*) son principalmente dispersados por aves (Hayano-Kanashiro, *et al.*, 2016), provocando que la poza génica localizada en Santo Domingo probablemente disminuya su variación genética al perder su principal dispersor de semillas o reduzca la distancia de dispersión ocasionando que la población se diferencie genéticamente de las otras poblaciones. Las poblaciones silvestres de *C. annuum* var. *glabriusculum* constituyen la mejor reserva de germoplasma útil para el sostenimiento natural de la especie y una fuente genética apreciable (Bran, *et al.*, 2007).

Por otra parte, el DAPC por región geográfica nos indicó que el 80% de los individuos analizados de cada región son exclusivos de la Costa y de Valles Centrales, mientras que para el caso de la Cañada el 30% de asignación supone que esta región comparte gran parte de su variación genética con las otras dos regiones. Probablemente, las condiciones ambientales (suelo, precipitación, temperatura) característicos de las regiones pueden estar influyendo en la variación genética de cada región. Sin embargo, para el caso de la Cañada es necesario tener en cuenta que se analizó una menor cantidad de individuos a diferencia de las otras dos regiones.

Los factores ecológicos, su forma de vida, rango geográfico, sistema de apareamiento, dispersión de semillas y el estado sucesional tienen efectos importantes en la distribución de la variación genética dentro de las poblaciones de plantas, especialmente cuando los factores ecológicos afectan la reproducción y dispersión de los individuos (Loveless y Hamrick, 1984; Nybom, *et al.*, 2004). Estas influencias ecológicas han sido observadas también para el maíz en México (Bellón, *et al.*, 2009).

La estructura obtenida mediante STRUCTURE a nivel poblacional y por variedad con una $K= 2$ nos arrojó una estructura genética correspondiente en la mayoría de los

casos con la altitud: aquellas poblaciones y variedades que se cultivan en tierras altas de Oaxaca de 600-1,700 msnm y que corresponden a la Cañada y Valles Centrales; y aquellas que se localizan en tierras bajas (0-600 msnm) que corresponden a la región de la Costa de Oaxaca. Estos dos grupos fueron los mismos que se mostraron en el árbol Neighbor-Joining mediante la distancia D_A , lo cual corroboraría que esta medida de distancia es más informativa conforme a las relaciones que existen entre poblaciones y variedades. Por ejemplo, ambos análisis ubican a la población de San Baltazar (chile Tusta) como parte de la Costa, mientras que a Rancho Llano (Guiña Danhi) la ubica como parte de tierras altas más parecida a Cuicatlán de la Cañada (Chilcostle Rojo y Huacles). Este último parecido podría estar asociado a la fuerte estacionalidad de secas-lluvias que presentan ambas regiones.

Por otra parte, la asignación genética para los individuos por variedad se encuentra relacionada también a su categoría de domesticación. Las variedades de chiles Tusta, Costeños y Mirasol (morfológicamente muy similar al chile silvestre, ver Anexo 1) son cultivadas principalmente en tierras bajas y fueron asignadas en un grupo ($K= 2$) junto a las poblaciones arvenses y silvestres. En el grupo 2 se agrupan todas las demás variedades cultivadas en tierras altas, incluidos el Guiña Danhi de la Costa de Oaxaca. No obstante, cuando se obtuvo el STRUCTURE únicamente para variedades cultivadas de *C. annuum* se observó el mismo patrón de asignación genética. Existen dos grupos genéticos de chiles cultivados.

Esta estructura nos refleja no solo el papel que juega la altitud en el cultivo de chiles, sino también, el grado de domesticación de cada uno y el manejo que se le da en cada región. La única excepción es el chile Tusta que, aunque es una variedad cultivada de amplio uso en la Costa y los Valles Centrales, no es agrupada con las demás

variedades cultivadas de tierras altas. Por otro lado, aunque el Costeño Rojo se cultiva principalmente en la Costa, este comparte cerca del 40% de similitud genética con las demás variedades cultivadas en tierras altas, además de presentar una forma generalmente alargada, triangular y de tamaño mediano (Muñoz-Zurita, 2013), muy similar al chile Taviche y Serrano de tierras altas. Probablemente, las características ambientales en las que son cultivados los chiles como la selección artificial están influyendo en la distribución de la variación genética de las poblaciones de chile *C. annuum*.

Se han definido dos principales formas de selección en plantas domesticadas: la selección intencional y la selección automática o mediada solamente por el ambiente de cultivo (Zohary, 2004). En el caso de los chiles la principal forma de selección se da mediante selección intencional, ya que las características del fruto como su tamaño, su textura, su color, y la tolerancia a plagas y enfermedades son determinantes en el proceso de selección del cultivo (Rodríguez, *et al.*, 2007; Latoumerie, *et al.*, 2001). Las decisiones que toman los agricultores sobre prácticas de manejo, localidad de cultivo, tamaño de la población y fuente de la semilla, también influyen en la diversidad genética de la planta cultivada (Brush, 2000).

La variedad de chile Guiña Danhi, el cual es cultivado en la localidad de Rancho Llano (se agrupó en tierras altas de Oaxaca en el STRUCTURE por población) y es representativo de la zona del Istmo de Tehuantepec (región sureste de la Costa), fue agrupada en zonas altas del estado de Oaxaca. La Información sobre este tipo de chile es escasa, sin embargo, la selección automática (con adaptaciones a ambientes creados por el ser humano) jugaría un rol más importante en la domesticación de este tipo de chile, al no verse afectado por la selección del fruto. Los campesinos no suelen

discriminar entre los frutos de este tipo de chile (comentario campesino). Tal vez, el chile Guiña Danhi derivó de zonas altas del estado de Oaxaca y su cultivo se adaptó a la región de la Costa, al observar en el árbol de Neighbor-Joining y en el análisis POPGRAPH por variedad un cierto grado de similitud con el chile tipo Chilcostle Rojo y Huacle de la región de la Cañada. Por otra parte, la palabra “*guiiña*” hace referencia al chile en lengua zapoteca, convirtiendo a Guiña Danhi en una variedad que aún conserva su nombre en Zapoteco (Long-Solís, *et al.*, 2012). De tal manera, cabe la posibilidad de que el parecido del chile Guiña Danhi con los chiles Chilcostle Rojo y Huacle se deba a una potencial migración de individuos Zapotecos de la Cañada hacia el Istmo de Tehuantepec, o bien, al posible intercambio de semillas entre estas dos regiones.

Una fuerte estructura genética puede ser asociada a la topografía como lo observó Aguilar-Meléndez, y colaboradores, en 2009. Ellos asociaron la estructura genética que encontraron, con la Sierra Madre Oriental, la Sierra Madre Occidental y la Península de Yucatán. En este trabajo se encontró con el análisis MONMONIER (sin la presencia de individuos de *C. frutescens* y *C. chinense*) una barrera entre las poblaciones localizadas al suroeste de la Costa de Oaxaca y al noroeste de los Valles Centrales que corresponde a la Sierra Madre del Sur. Esta barrera contribuiría al aislamiento entre regiones al interferir el flujo de la parte suroeste de la Costa hacia zonas altas. Los grupos obtenidos por SAMOVA con una $K=6$ corroboran a la Sierra Madre del Sur como barrera, mientras que ambos análisis Monmonier con y sin individuos de *C. frutescens* y *C. chinense* proponen otra barrera que correspondería al Cerro Nube (Quie Yelaag). Por otro lado, el análisis de POPGRAPH que también identifica barreras geográficas propone otra posible barrera en la zona suroeste de la Costa hacia la región de la Cañada y ésta correspondería a la Sierra de Tlaxiaco. Esto parece indicar que las principales barreras

geográficas para los chiles *C. annuum* son aquellas que se relacionan a sistemas montañosos.

En los análisis de estructura genética se observó que la población de Baltazar es agrupada con las poblaciones de la Costa, y que la barrera geográfica que conforma la Sierra Madre del Sur, la separa de San Pablo Nexila. Esto ocurre porque ambiental y económicamente esta población es considerada como parte de la Costa de Oaxaca (García-Mendoza, *et al.*, 2004), sin embargo, su altitud la ubica como parte de tierras altas (600 a 1, 700 msnm).

VIII.3. Manejo humano moldeador de la diversidad genética

Se ha establecido que pequeñas poblaciones de linajes endémicos con distancias geográficas más cortas (caso de los Valles Centrales y la Cañada, ver Fig. 18), pueden ser más susceptibles a efectos de deriva y a un limitado flujo de genes que incrementan la diferenciación entre poblaciones, pero la disminuye al interior de éstas (Loveless y Hamrick, 1984). El chile de Agua es de alta preferencia y consumo en los Valles Centrales siendo considerado endémico de la región mientras que los chiles tipo Paradito, Nanchita y Taviche son conocidos localmente por su consumo (Castellón-Martínez, *et al.*, 2012; Montaña-Lugo, *et al.*, 2014). Probablemente, las preferencias de consumo y el mercado local influyen en la diversidad genética de las poblaciones de chile al establecerse variedades endémicas en cada región, aumentando su diferenciación genética. Se trata además de chiles altamente diferenciados en morfología y propiedades de uso (color, aroma y sabor).

El chile Huacle es considerado como una variedad endémica de la región de la Cañada (García-Gaytan, *et al.*, 2017). Sus patrones de estructura y relaciones genéticas lo agrupan cercanamente con los chiles de Agua. Sin embargo, en el chile Huacle no fue posible encontrar loci polimórficos, tal vez debido a que la semilla empleada suele proceder de frutos cosechados en el ciclo inmediato anterior, y que existen alrededor de 15 productores y una superficie anual de aproximadamente 10 hectáreas para el cultivo de este chile en Cuicatlán, no efectuándose alguna estrategia para incrementar la variabilidad del chile Huacle (Muñoz-Zurita, 2013; López-López, *et al.*, 2016). En ocasiones, el chile Huacle que se vende en Cuicatlán, llega a proceder de un número reducido de familias, que al perder cosechas por efectos climáticos reservan semillas a partir de unas pocas proporcionadas entre estas mismas familias (comentario campesino a grupo de trabajo derivado de salidas al campo) conduciendo a fuertes cuellos de botella.

Los agricultores reportan que en el mismo chilar donde es posible encontrar chiles Huacles de pronto crecen, uno al lado del otro, chiles que ellos denominan Chilcostle, el cual en la cosecha y venta no es separado de los Huacles, ya que los comerciantes lo pagan por igual (Aguilar-Rincón, *et al.*, 2010; Muñoz-Zurita, 2013; comentario campesino). Este chile probablemente es una variante segregada del chile Huacle y se diferencian en la forma del fruto, el Chilcostle toma una forma triangular cónica, mientras que los Huacles son redondos y globosos (Muñoz-Zurita, 2013; ver Anexo 1). Aunque ambos chiles están sujetos a las mismas condiciones ambientales y de cultivo (ambos se cultivan en Cuicatlán), el chile Huacle no presentó variación genética, mientras que en el Chilcostle los niveles de variación genética ($H_E = 0.28$) fueron considerables con solo 4 individuos analizados, indicándonos que la fuerte selección a la que ha sido sometido el chile Huacle y la reducción en el número de sus productores han conducido a la

disminución de su variación genética. Estudios de diversidad genética que involucren una mayor cantidad de individuos son necesarios en esta variedad al encontrarse en peligro de extinción (García-Gaytán, *et al.*, 2017) y ser altamente susceptible a plagas (López-López, *et al.*, 2016).

Las reducciones de la diversidad de las variedades de chiles cultivados en relación al pariente silvestre van desde muy bajas (3% para Costeño Amarillo) hasta muy altas (66% para Mirasol). La diversidad que puede conservar el chile cultivado puede ser atribuida a una relación entre los habitantes y las plantas, derivadas de las prácticas de manejo agrícola que llevan a cabo los zapotecos y que se han registrado en regiones como los Valles Centrales (Martínez-Sánchez, *et al.*, 2010). Esta relación, también ha sido observada en Yucatán, donde la combinación de una influencia de la cultura maya y de los sistemas de manejo tradicional como milpas, solares y huertos familiares, permite tener una gran diversidad genética y morfológica de plantas como el chile y frijol, los cuales son conservados y aprovechados localmente por el campesino (Aguilar-Meléndez, *et al.*, 2009; Latoumerie, *et al.*, 2001; Martínez-Castillo, *et al.*, 2004). Por otra parte, un gran número de las variedades locales son cultivadas en parcelas pequeñas o en huertos y traspatios que se encuentran inmersos o cercanos a remanentes de bosque tropical caducifolio, conformando lo que clásicamente Perfecto y colaboradores, (2009), han denominado una matriz agroecológica, en la que los elementos de la vegetación natural y los elementos de los campos de cultivo interactúan de muchas formas; el flujo génico es una de ellas.

La selección direccional y artificial juega un papel importante en la diversidad de los chiles cultivados en zonas altas. En el caso particular del chile de Agua los sistemas agrícolas tradicionales en los Valles Centrales, han permitido la diversidad morfológica

de este tipo de chile (Martínez-Sánchez, *et al.*, 2010), y también, su diversidad genética permitiendo preservar a pesar de su amplio cultivo, niveles de diversidad genética considerables ($H_E= 0.27$).

En este estudio se analizaron también diferentes asignaciones genéticas del análisis STRUCTURE independientemente del resultado de la prueba de Evanno. Uno de estos grupos fue $K= 3$ el cual indica que la región de la Costa presenta las 3 pozas génicas. Esta diversidad genética en cultivos locales se convierte en una defensa contra pestes, enfermedades y cambios ambientales (Galluzi, *et al.*, 2010), lo cual es importante ante el mayor riesgo en el cultivo del chile que consiste en su susceptibilidad a plagas y enfermedades (Rodríguez, *et al.*, 2007). Por otra parte, una de estas 3 pozas génicas coincide en la mayoría de los casos con la presencia de individuos silvestres. Lo mismo aplica para el caso de una $K= 4$.

Se ha indicado que los sistemas de manejo tradicional juegan un papel importante en la conservación de los recursos genéticos del chile *C. annuum*, especialmente cuando las poblaciones silvestres declinan (Aguilar-Meléndez, *et al.*, 2009, González-Jara, *et al.*, 2011; Pickersgill, 2016). Las poblaciones arvenses analizadas provienen de estos sistemas de manejo (traspacios, milpas y jardines) con niveles de diversidad genética similares a los silvestres. Estos chiles, así sea por efectos del azar, son transportados a estos sistemas por las aves (Hayano-Kanashiro, *et al.*, 2016), permitiéndoles su crecimiento como plantas arvenses, toleradas por las familias y utilizadas posteriormente para autoconsumo y en algunos casos como ingreso económico (Latoumerie, *et al.*, 2001).

Por otra parte, el hecho de categorizar un chile *C. annuum* con el nombre de una variedad según cada localidad parece otorgarle no solo una identidad cultural sino

también genética. El DAPC por grado de domesticación mostró que cuando se categoriza un chile como una forma cultivada, arvense o silvestre la variación genética para el caso de los cultivados parece no compartirse con las arvenses y silvestres, ya que el 97% de los individuos de las variedades cultivadas se agrupan en una nube que no se sobrelapa con las otras formas de manejo. Sin embargo, cuando un chile *C. annuum* lo categorizas como una variedad denominada según la localidad de cultivo y el agricultor que la siembra, el DAPC arroja que los chiles Tusta y Mirasol no comparten gran parte de su composición genética con las otras variedades cultivadas. Este patrón fue observado también en el árbol de Neighbor Joining con la distancia genética D_A , donde el chile Tusta y Mirasol son relacionados en un grupo diferente a las demás variedades cultivadas agrupándolos con los arvenses y silvestres. Estos resultados parecen ser un indicador de como el manejo o la identidad cultural que se le da a una variedad de chile *C. annuum* está asociado también a su composición genética.

VIII.4. *C. annuum* y las especies *C. frutescens* y *C. chinense* en la Costa de Oaxaca

La diversidad y estructura genética de la región de la Costa parece ser más compleja en comparación con los Valles Centrales y la Cañada. Los análisis de diversidad genética y los resultados STRUCTURE por variedad indican que el chile Costeño Rojo presenta dos pozas génicas: una similar a los chiles cultivados de tierras altas y otra muy similar a las dos pozas génicas de los chiles silvestres. La alta diversidad genética encontrada en esta variedad puede ser consecuencia del flujo génico entre los chiles silvestres y el Costeño Rojo. Sin embargo, el STRUCTURE con una $K= 3$ únicamente para variedades cultivadas

indicó que una de sus pozas génicas es similar a la de los chiles de Agua, Árbol, Chilcostle Rojo y Guiña Danhi, mientras que su otra poza génica es similar a la de Tusta y Mirasol.

Se ha reportado que el chile Costeño presenta una flor por axila con un estigma excerto (López y Montes, 2006), favoreciendo la polinización cruzada (Barrios, *et al.*, 2007). Tanto los chiles silvestres como los chiles Chilcostle, Tusta, Taviche, Paradito y Mirasol se han registrado con excerción del estigma, mientras que el chile de Agua y el Huacle presentan un estigma al mismo nivel que las anteras (López y Montes, 2006) conduciendo a la autopolinización (Barrios, *et al.*, 2007). Esto favorecería la hipótesis acerca del posible flujo génico entre el chile Costeño con los chiles silvestres, Tusta, Mirasol y Chilcostle Rojo, dando indicios acerca de que, probablemente el proceso de domesticación que han atravesado las variedades cultivadas de la región de la Costa les ha permitido presentar un estigma excerto igual que los silvestres (López y Montes, 2006).

El chile Costeño no ha sido ampliamente estudiado ni de interés como el chile de Agua o el chile Huacle. Por lo que se hace de especial importancia realizar estudios detallados sobre sus sistemas de apareamiento, su morfología y su diversidad genética. Este estudio revela que la variedad Costeño Rojo se encuentra afectada por fuertes cuellos de botella ($M= 0.28$) reduciendo su diversidad alélica en un 70% (conforme al tamaño mínimo de muestra), presenta niveles de endogamia considerables ($F_{IS}= 0.48$) y los mayores niveles de diversidad genética ($H_E= 0.36$).

El chile Tusta, variedad local del estado de Oaxaca cultivada en los Valles Centrales (Ejutla de Crespo) y en la Costa (en Pochutla) se ha reportado como una variedad con alta diversidad fenotípica en formas, tamaños, pesos y rendimiento de los frutos por planta y una alta preferencia de consumo debido a su sabor y picor en los

Valles Centrales (Castellón-Martínez, *et al.*, 2012; 2014; Santiago-Luna, *et al.*, 2016). En este trabajo el chile Tusta presentó niveles de F_{IS} negativos propios de la exogamia, una condición atípica entre las variedades cultivadas y que posiblemente se deba al factor de la polinización cruzada. La estructura genética con una $K= 4$ nos indica que probablemente esta variedad se ha diferenciado de todas las demás variedades cultivadas de *C. annuum*, como consecuencia de una domesticación paralela a las demás variedades, en la cual fuertes procesos de deriva génica ($M= 0.25$) y flujo génico ($F_{IS}= -0.04$) permiten su diferenciación. Esto es indicador de cómo la selección del agricultor y las condiciones ambientales propias de la Costa y las cañadas y barrancas de la Sierra Madre del Sur están influyendo en el cultivo de este chile. La implementación de trabajos que involucren su diversidad genética en todo el estado son importantes ante el panorama de que el acervo genético de esta variedad de chile se encuentra en riesgo, principalmente en los Valles Centrales (Castellón-Martínez., *et al.*, 2012), y posiblemente sea remplazada por variedades comerciales de mayor demanda en el mercado (Santiago-Luna, *et al.*, 2016).

Por otra parte, la relación Tusta-Silvestres-Arvenses-Mirasol mostrada en los análisis de Neighbor-Joining, DAPC, STRUCTURE y POPGRAH puede tener su origen en su interacción con las especies de *C. frutescens* y *C. chinense*. Por lo menos los chiles Tusta, Mirasol y arvenses comparten parte de sus pozas genéticas con las otras dos especies (ver Fig. 26) permitiendo su diferenciación de las demás variedades de *C. annuum*. El chile Mirasol y los individuos arvenses podemos localizarlos principalmente en sistemas de manejo tradicional como traspatios y jardines, donde es posible localizar también individuos de *C. frutescens* y *C. chinense* (observaciones de campo). Además,

el chile Tusta puede ser localizado en zonas cercanas a estos individuos arvenses, tal fue el caso de la población Juan Diegal.

En 2007, Barrios y colaboradores identificaron en sistemas de agricultura tradicional hibridaciones intra e interespecíficas dentro y entre las especies del complejo *C. annuum-frutescens-chinense* (grupo de flores blancas), excepto para el cruzamiento entre *C. chinense* x *C. frutescens*. Probablemente, estas hibridaciones se efectúen también entre las variedades de *C. annuum* (Tusta-Mirasol-Arvenses) analizadas en este estudio. Más aún las agrupaciones generadas en el análisis de DAPC al incluir a *C. frutescens* parecen reflejar la posibilidad de que esta especie intercambie material genético con las poblaciones arvenses y de chile Mirasol dentro de *C. annuum*. La prueba de Mantel para la región de la Costa, la cual presenta la mayor cantidad de individuos de *C. frutescens* y *C. chinense*, indicó que no existe aislamiento por distancia a este nivel ($p= 0.052$), por lo tanto, probablemente se deba al flujo génico entre estas especies.

En general la inclusión en los análisis de individuos de *C. frutescens* y *C. chinense*, no afectó considerablemente la dispersión de los datos obtenidos comparándola como cuando se analizó únicamente a *C. annuum*. Para el caso de *C. chinense* solo se contaba con tres individuos en dos poblaciones, mientras que los 16 individuos de *C. frutescens* analizados por medio del DAPC mostraron que su variación genética es muy similar a *C. annuum* al presentar un valor de asignación de individuos de 0. Estas especies son bastante similares entre sí, especialmente a nivel ancestral (Long-Soliís, 2012). Por lo tanto, se sugiere que futuros trabajos de diversidad genética en la especie *C. annuum* incluyan, si se encuentran en las cercanías de muestreo, a estas otras dos especies ya que podrían aportar información importante sobre los patrones y distribución de la diversidad genética de la especie *C. annuum*.

Hoy en día, es necesario generar conocimiento sobre los sistemas de agricultura tradicional, sus patrones de diversidad genética, y su capacidad de responder a cambios ambientales (Galluzi, *et al.*, 2010). Este trabajo, da un primer acercamiento sobre la diversidad genética de poblaciones cultivadas, arvenses y silvestres en tres regiones del estado de Oaxaca y cómo los sistemas de agricultura tradicional juegan un rol importante en la conservación de los recursos genéticos del chile *C. annuum*.

IX. CONCLUSIONES

- La altitud en la que crecen las poblaciones de chile *Capsicum annuum* influye en los patrones de diversidad y estructura genética de las poblaciones al encontrarse mayores niveles de diversidad genética y menor diferenciación debido a eventos de dispersión en tierras bajas (región de la Costa de Oaxaca).
- El manejo de los agricultores o campesinos define grandes grupos de diversidad genética de chile *C. annuum* al localizarse por lo menos 3 pozas génicas para variedades cultivadas: la poza de las variedades de tierras altas (incluyendo a Guiña Danhi), la de Tusta y Mirasol, y la de los Costeños (fundamentalmente tierras bajas).
- La región de la Costa presenta la mayor diversidad genética y de sistemas de manejo, tanto de poblaciones cultivadas como de arvenses y silvestres, por lo que puede considerarse como un área prioritaria de conservación de los recursos genéticos del chile *C. annuum* en México.
- Las formas de manejo y uso (sabor, aroma y grado de picor) que se les da a las variedades de chile *C. annuum* como el chile de Agua en los Valles Centrales y el chile Huacle en la Cañada aunado al sistema de apareamiento en cada variedad permite la diferenciación genética entre las poblaciones.
- Las poblaciones arvenses presentaron niveles de diversidad genética similares a las silvestres. Estas son localizadas principalmente en sistemas de manejo tradicional como jardines, traspatios y huertos familiares, recalcando la importancia de mantener y conservar estos sistemas de manejo, ante el declive de poblaciones silvestres debido a la fragmentación del hábitat.

- La variedad Costeño Rojo presentó dos pozas génicas probablemente producto de la polinización cruzada con los individuos silvestres de *C. annuum* produciendo altos niveles de diversidad genética a comparación de las otras variedades analizadas.
- Los datos obtenidos sugieren que la presencia de las especies *C. frutescens* y *C. chinense* alrededor de los campos de cultivo y en los sistemas de manejo como los traspatios puede conducir a hibridación con la especie *C. annuum* contribuyendo a la diversidad y diferenciación genética de las variedades de Chile *C. annuum*
- La variedad de Chile Tusta se diferenció de las demás variedades cultivadas de *C. annuum*, pero es claramente diferente morfológicamente respecto a las silvestres y arvenses, lo cual podría hablarnos de una domesticación paralela en relación a otras variedades cultivadas de *C. annuum* y/o de la polinización cruzada con los individuos arvenses o de flujo génico con las especies *C. frutescens* y *C. chinense*.

X. REFERENCIAS

- Agatón, G., Santiago, A., Vallejo, J. M. S., y Montaña, A. (2016). Estudio de impacto ambiental, económico y social en la región del istmo de Tehuantepec, Oaxaca, México, debido a la instalación de parques eólicos. *Tlamati*, 7(1):14-21
- Aguilar-Meléndez, A., Morrell, P. L., Roose, M. L., y Kim, S. C. (2009). Genetic diversity and structure in semiwild and domesticated chiles (*Capsicum annum*; Solanaceae) from Mexico. *American Journal of Botany*, 96(6), 1190-1202.
- Aguilar-Rincón, V. H., Corona-Torres, T., López-López, P., Latournerie-Moreno, L., Ramírez-Meraz, M., Villalón-Mendoza, H., y Aguilar-Castillo, J. A. (2010). *Los chiles de México y su distribución* (p. 114). SINAREFI, Colegio de Postgraduados, INIFAP, ITConkal, UANL, UAN. Montecillo, Texcoco, Estado de México.
- Allendorf, F. W., y Luikart, G. (2009). *Conservation and the genetics of populations* (p. 664). John Wiley y Sons, USA.
- Azofeita-Delgado, A. (2006). Uso de marcadores moleculares en plantas; aplicaciones en frutales del trópico. *Agronomía mesoamericana*, 17(2): 221-242.
- Balloux, F., y Lugon-Moulin, N. (2002). The estimation of population differentiation with microsatellite markers. *Molecular ecology*, 11(2): 155-165.
- Barrios, O., Fuentes, V., Shagarosdky, T., Cristóbal, R., Castiñeiras, L., Fundora, Z.,... y Fernández, F. (2007). Nuevas combinaciones híbridas de *Capsicum* spp. en sistemas de agricultura tradicional de occidente y oriente de Cuba. *Agrotecnia de Cuba*, 31(2): 327-335.
- Bellón, M. R., Barrientos-Priego, A. F., Colunga-GarcíaMarín, P., Perales, H., Reyes-Agüero, J. A., Rosales-Serna, R., y Zizumbo-Villareal, D. (2009). Diversidad y

- conservación de recursos genéticos en plantas cultivadas. *Capital natural de México*, 2: 355-382.
- Bradburd, G. S., Ralph, P. L., y Coop, G. M. (2013). Disentangling the effects of geographic and ecological isolation on genetic differentiation. *Evolution*, 67(11): 3258-3273.
- Bran, R. A. A., Moya, C., Ponce, P., Álvarez, M., y Varela, M. (2007). Diagnóstico participativo de las condiciones socioculturales asociadas a la conservación de los chiles silvestres (*Capsicum* spp.) en la Depresión Central de Chiapas, México. *Cultivos Tropicales*, 28(1): 69-73.
- Brown, A. H. D. (1978). Isozymes, plant population genetic structure and genetic conservation. *Theoretical and Applied Genetics*, 52(4): 145-157.
- Brush, S. B. (2000). *Genes in the field: on-farm conservation of crop diversity* (p. 228). IDRC.
- Brush, S. B. (2004). *Farmers Bounty: Locating Crop Diversity in the Contemporary World* (p. 327). Yale University Press.
- Burger, J. C., Chapman, M. A., y Burke, J. M. (2008). Molecular insights into the evolution of crop plants. *American Journal of Botany*, 95(2): 113-122.
- Caballero, J., Cortés, L., Martínez-Alfaro, M. A., y Lira-Saade, R. (2004). *Uso y manejo tradicional de la diversidad vegetal. Biodiversidad de Oaxaca*. Instituto de Biología, Universidad Nacional Autónoma de México-Fondo Oaxaqueño para la Conservación de la Naturaleza, México-World Wildlife Fund, México, DF, 541-564.
- Casas, A., Otero-Arnaiz, A., Pérez-Negrón, E., y Valiente-Banuet, A. (2007). *In situ* management and domestication of plants in Mesoamerica. *Annals of Botany*, 100 (5): 1101-1115.

- Castellón-Martínez, E., Chávez-Servia, J. L., Carrillo-Rodríguez, J. C., y Vera-Guzmán, A. M. (2012). Preferencias de consumo de chiles (*Capsicum annuum* L.) nativos en los valles centrales de Oaxaca, México. *Revista fitotecnia mexicana*, 35(5): 27-35.
- Castellón Martínez, E., Carrillo-Rodríguez, J. C., Chávez-Servia, J. L., y Vera-Guzmán, A. M. (2014). Variación fenotípica de morfotipos de chile (*Capsicum annuum* L.) nativo de Oaxaca, México. *Phyton*, 83(2): 225-236.
- Chapuis, M.-P. y Estoup, A. (2007) Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution*, 24(3): 621-631
- Charlesworth, B., y Charlesworth, D. (2017). Population genetics from 1966 to 2016. *Heredity*, 118(1): 2-9.
- Clement, C. R. (1999). 1492 and the loss of Amazonian crop genetic resources. I. The relation between domestication and human population decline. *Economic Botany*, 53(2): 188-202.
- Clement, C. R., de Cristo-Araújo, M., Coppens D'Eeckenbrugge, G., Alves Pereira, A., y Picanço-Rodríguez, D. (2010). Origin and domestication of native *Amazonian crops*. *Diversity*, 2(1): 72-106.
- Conner, J. K., y Hartl, D. L. (2004). *A primer of ecological genetics* (p 304). Sinauer Associates Incorporated.
- Contreras Toledo, A. R., López Sánchez, H., Santacruz Varela, A., Valadez Moctezuma, E., Aguilar Rincón, V. H., Corona Torres, T., y López, P. A. (2011). Diversidad Genética en México de variedades nativas de chile poblano mediante microsatélites. *Revista fitotecnia mexicana*, 34(4): 225-232.

- Cooper, G., Amos, W., Bellamy, R., Siddiqui, M. R., Frodsham, A., Hill, A. V., y Rubinsztein, D. C. (1999). An empirical exploration of the $(\delta\mu)^2$ genetic distance for 213 human microsatellite markers. *The American Journal of Human Genetics*, 65(4): 1125-1133.
- Cornuet J.M. y Luikart G. (1996). Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, 144: 2001-2014.
- Darwin, C. (1868). *The variation of animals and plants under domestication* (Vol. 2). O. Judd.
- De Candolle, A. (1883). *Origine des plantes cultivées* (Vol. 43). G. Baillièrre et cie.
- Di Rienzo, A., Peterson, A.C., Garza, J.C., Valdés, A.M., Slatkin M., y Freimer, N.B. (1994) Mutational processes of simple-sequence repeat loci in human populations. *Proceedings of the National Academy of Sciences of the USA*, 91: 3166–3170.
- Djian-Caporalino, C., Lefebvre, V., Sage-Daubèze, A. M., y Palloix, A. (2006). *Capsicum*. Genetic Resources, Chromosome Engineering, and Crop Improvement. *Vegetable Crops* (p. 186-245 Vol. 3).
- Do, C., Waples, R. S., Peel, D., Macbeth, G. M., Tillett, B. J., y Ovenden, J. R. (2014). NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (N_e) from genetic data. *Molecular Ecology Resources*, 14(1): 209-214.
- Doebley, J. F., Gaut, B. S., y Smith, B. D. (2006). The molecular genetics of crop domestication. *Cell*, 127(7): 1309-1321.
- Doyle, J.J. y J. L. Doyle. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf material. *Phytochemical Bulletin*, 19: 11–15.

- Dupanloup, I., Schneider, S., y Excoffier, L. (2002). A simulated annealing approach to define the genetic structure of populations. *Molecular Ecology*, 11(12): 2571-2581.
- Dyer, R. J. (2015). Population graphs and landscape genetics. *Annual Review of Ecology, Evolution and Systematics*, 46: 327-342.
- Dyer, R. J., y Nason, J. D. (2004). Population graphs: the graph theoretic shape of genetic structure. *Molecular ecology*, 13(7): 1713-1727.
- Dyer, R. J., Nason, J. D., y Garrick, R. C. (2010). Landscape modelling of gene flow: improved power using conditional genetic distance derived from the topology of population networks. *Molecular Ecology*, 19(17): 3746-3759.
- Earl, Dent A. y vonHoldt, Bridgett, M. (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, 4(2): 359-361.
- Eguiarte, L. (2009). Nueva guía para principiantes a la genética de poblaciones. En J. J. Morrone y P. Magaña. (2009). *Evolución biológica: Una visión actualizada desde la revista Ciencias* (1° ed., pp. 83-102). México: Facultad de Ciencias, Universidad Nacional Autónoma de México.
- Ellegren, H., y Galtier, N. (2016). Determinants of genetic diversity. *Nature Reviews Genetics*, 17(7): 422-433.
- Ellstrand, N. C., y Elam, D. R. (1993). Population genetic consequences of small population size: implications for plant conservation. *Annual review of Ecology and Systematics*, 24(1): 217-242.
- Ellstrand, N. C., Prentice, H. C., y Hancock, J. F. (1999). Gene flow and introgression from domesticated plants into their wild relatives. *Annual review of Ecology and Systematics*, 30(1): 539-563.

- Emshwiller, E. (2006). Genetic data and plant domestication. En Zedder M.A., *et al.*, eds. (2006). *Documenting Domestication: New Genetic and Archaeological Paradigms*. (pp. 99-122).
- Engels, J. M. M., Ebert, A. W., Thormann, I., y De Vicente, M. C. (2006). Centres of crop diversity and/or origin, genetically modified crops and implications for plant genetic resources conservation. *Genetic Resources and Crop Evolution*, 53(8): 1675-1688.
- Eshbaugh, W. H. (2012). The Taxonomy of the Genus *Capsicum* (pp. 14-28). En Ruso, V. M. (2012). *Peppers: Botany, Production and Uses*. CABI.
- Evanno, G., Regnaut, S., y Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular ecology*, 14(8): 2611-2620.
- Excoffier, L., Laval, G., y Schneider, S. (2005). Arlequin (version 3.5): an integrated software package for population genetics data analysis. *Evolutionary bioinformatics*, 1.
- Franceschetti, U. (1971). Natural cross pollination in pepper (*Capsicum annum* L.). En *Eucarpia Meeting on Genetics and Breeding of Capsicum*, (pp 16-18).
- Frankel, O. H. (1974). Genetic conservation: our evolutionary responsibility. *Genetics*, 78(1): 53-65.
- Galluzzi, G., Eyzaguirre, P., y Negri, V. (2010). Home gardens: neglected hotspots of agro-biodiversity and cultural diversity. *Biodiversity and conservation*, 19(13): 3635-3654.
- García-Gaytán, V., Gómez-Merino, F. C., Trejo-Téllez, L. I., Baca-Castillo, G. A., y García-Morales, S. (2017). The Chilhuacle Chili (*Capsicum annum* L.) in Mexico:

- Description of the Variety, Its Cultivation, and Uses. *International Journal of Agronomy*, 2017: 1-13.
- García-Mendoza, A. J., Díaz, M. D. J. O., y Briones-Salas, M. (2004). *Biodiversidad de Oaxaca* (p. 605). UNAM.
- Garrido-Garduño, T., y Vázquez-Domínguez, E. (2013). Métodos de análisis genéticos, espaciales y de conectividad en genética del paisaje. *Revista mexicana de biodiversidad*: 84(3): 1031-1054.
- Garza, J. C., y Williamson, E. G. (2001). Detection of reduction in population size using data from microsatellite loci. *Molecular ecology*, 10(2): 305-318.
- Gepts, P. (2004). Plant Crop Domestication as a Long-Term Selection Experiment. *Breeding Reviews*, 24: 1-44.
- Gepts, P., y Papa, R. (2002). *Evolution during domestication* (pp. 1-7). Encyclopedia of Life Sciences.
- Gilbert, K. J. (2016). Identifying the number of population clusters with structure: problems and solutions. *Molecular ecology resources*, 16(3): 601-603.
- Goldstein, D. B., Linares, A. R., Cavalli-Sforza, L. L., y Feldman, M. W. (1995). Genetic absolute dating based on microsatellites and the origin of modern humans. *Proceedings of the National Academy of Sciences*, 92(15): 6723-6727.
- Goldstein, D. B., y Pollock, D. D. (1997). Launching microsatellites: a review of mutation processes and methods of phylogenetic inference. *Journal of heredity*, 88(5): 335-342.
- González-Jara, P., Moreno-Letelier, A., Fraile, A., Piñero, D., y García-Arenal, F. (2011). Impact of human management on the genetic variation of wild pepper, *Capsicum annum* var. *glabriusculum*. *PLoS One*, 6(12): 1-11.

- Goudet, J. (2002). *FSTAT version 2.9. 3.2, a program to estimate and test gene diversities and fixation indices*. Institute of Ecology, Lausanne, Switzerland.
- Gross, B. L., y Olsen, K. M. (2010). Genetic perspectives on crop domestication. *Trends in plant science*, 15(9): 529-537.
- Harlan, J. R. (1971). Origins of agriculture: centers and no centers. *Science*, 174: 468-474.
- Harlan, J. R. (1975). Geographic patterns of variation in some cultivated plants. *Journal of Heredity*, 66(4): 182-191.
- Harlan, J. R., Wet, J. D., y Price, E. G. (1973). Comparative evolution of cereals. *Evolution*, 27(2): 311-325.
- Hartl, D. L., y Clark, A. G. (2007). *Principles of Population Genetics* (4ta ed., p. 653). Sunderland, Massachusetts: Sinauer Associates, Inc. Publishers.
- Hayano-Kanashiro, C., Gámez-Meza, N., y Medina-Juárez, L. A. (2016). Wild pepper *Capsicum annuum* L. var. *glabriusculum*: Taxonomy, plant morphology, distribution, genetic diversity, genome sequencing, and phytochemical compounds. *Crop Sci*, 56: 1-11.
- Hedrick, P. W. (2011). *Genetics of populations* (4ta ed., p. 675). Sudbury, Massachusetts: Jones and Bartlett publishers.
- Hernández-Verdugo, S., Dávila, A. P. y Oyama, K. (1999). Síntesis del conocimiento taxonómico, origen y domesticación del género *Capsicum*. *Boletín de la Sociedad Botánica de México*, 64: 65-84.
- Hernández-Verdugo, S., Luna-Reyes, R., y Oyama, K. (2001). Genetic structure and differentiation of wild and domesticated populations of *Capsicum annuum* (Solanaceae) from Mexico. *Plant Systematics and Evolution*, 226 (3-4): 129-142.

- Heslop-Harrison, J. S., Schwarzacher, T., Altman, A., y Hasegawa, P. M. (2011). Genetics and genomics of crop domestication. *Plant biotechnology and agriculture: prospects for the 21st century*, 1-16.
- Jara, E. C. (2011). Problemática en torno a la construcción de parques eólicos en el istmo de Tehuantepec. *Revista Desarrollo local sostenible*, 4(12):1-14
- Jardón-Barbolla, L. (2015). Orígenes y diversidad a la mitad de las montañas: Nikolai Vavílov, México y las plantas domesticadas. *Oikos*, 14: 6-10.
- Jardón-Barbolla, L. (2017). De Sonora a Yucatán. Chiles en México: diversidad y domesticación. *Oikos*, 17: 25-29.
- Jombart, T., S. Devillard y F. Balloux. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics*, 11: 94 108.
- Kalia, R. K., Rai, M. K., Kalia, S., Singh, R., y A. K. Dhawan. (2011). Microsatellite markers: an overview of the recent progress in plants. *Euphytica*, 177(3): 309-334.
- Kimura, M., y Crow, J.F. 1964. The number of alleles that can be maintained in a finite population. *Genetics*, 49: 725-738.
- Kraft, K. H., Brown, C. H., Nabhan, G. P., Luedeling, E., Ruiz, J. D. J. L., d'Eeckenbrugge, G. C., ... y Gepts, P. (2014). Multiple lines of evidence for the origin of domesticated chili pepper, *Capsicum annum*, in Mexico. *Proceedings of the National Academy of Sciences*, 111(17): 6165-6170.
- Ladizinsky, G. (2012). *Plant evolution under domestication* (p. 254). Springer Science y Business Media.
- Langella, O. (1999). Populations 1.2.32: a population genetic software. Disponible en: http://www.bioinformatics.org/project/?group_id=84

- Latoumerie, L., Chávez, J. L., Pérez, M., Hernández, C. F., Martínez, R., Arias, L. M., y Castañón, G. (2001). Exploración de la diversidad morfológica de chiles regionales en Yaxcaba, Yucatán, México. *Agronomía mesoamericana*, 12(1): 41-47.
- Loaiza-Figueroa, F., Ritland, K., Cancino, J. A. L., y S. D. Tanksley. (1989). Patterns of genetic variation of the genus *Capsicum* (Solanaceae) in Mexico. *Plant Systematics and Evolution*, 165(3-4): 159-188.
- Long-Solís, J. (2012). *Capsicum* y cultura: La historia del chilli (p. 203). México: Fondo de Cultura Económica.
- López L. P y S. Montes H. 2006. Avances de investigación de la red de hortalizas del SINAREFI (p. 466 vol. I). Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias, Campo Experimental Bajío. Celaya, Gto. México.
- López-López, P., Rodríguez, H. R., y Bravo, E. M. (2016). Impacto económico del chile huacle (*Capsicum annuum* L) en el estado de Oaxaca. *Revista Mexicana de Agronegocios*, 38: 317-328.
- Loveless, M. D., y Hamrick, J. L. (1984). Ecological determinants of genetic structure in plant populations. *Annual review of ecology and systematics*, 15(1): 65-95.
- Luikart G., Allendorf, F. W., Cornuet, J. M. y William, B., S. (1997). Distortion of allele frequency distributions provides a test for recent population bottlenecks. *Journal of Heredity*, 89(3): 238-247.
- Mapes, C., y Basurto. F. (2016). Biodiversity and edible plants of Mexico. En *Ethnobotany of Mexico* (pp. 83-131). Springer New York.
- Martínez-Castillo, J., Zizumbo-Villarreal, D., Perales-Rivera, H., y Colunga-Garcíaamarin, P. (2004). Intraspecific diversity and morpho-phenological variation in *Phaseolus lunatus* L. from the Yucatan Peninsula, Mexico. *Economic Botany*, 58(3): 354-380.

- Martínez-Sánchez, D., Pérez-Grajales, M., Rodríguez-Pérez, J. E., Pérez, M., y del Carmen, E. (2010). Colecta y caracterización morfológica de chile de agua (*Capsicum annuum* L.) en Oaxaca, México. *Revista Chapingo. Serie horticultura*, 16(3): 169-176.
- Maruyama, T., y Fuerst, P. A. (1985). Population bottlenecks and non equilibrium models in population genetics. II. Number of alleles in a small population that was formed by a recent bottleneck. *Genetics*, 111: 675-689.
- Matton, D. P., Nass, N., Clarke, A. E., y Newbigin, E. (1994). Self-incompatibility: how plants avoid illegitimate offspring. *Proceedings of the National Academy of Sciences*, 91(6): 1992-1997.
- McLeod, M. J., Guttman, S. I., Eshbaugh, W. H., y Rayle, R. E. (1983). An electrophoretic study of evolution in *Capsicum* (Solanaceae). *Evolution*, 37(3): 562-574.
- Meave, J. A., Romero-Romero, M. A., Salas-Morales, S. H., Pérez-García, E. A., y Gallardo-Cruz, J. A. (2012). Diversidad, amenazas y oportunidades para la conservación del bosque tropical caducifolio en el estado de Oaxaca, México. *Revista Ecosistemas*, 21(1-2): 85-100.
- Meirmans, P. G. (2015). Seven common mistakes in population genetics and how to avoid them. *Molecular Ecology*, 24(13): 3223-3231.
- Meyer, R. S., y Purugganan, M. D. (2013). Evolution of crop species: genetics of domestication and diversification. *Nature Reviews. Genetics*, 14(12): 840.
- Milla, R., Osborne, C. P., Turcotte, M. M., y Violle, C. (2015). Plant domestication through an ecological lens. *Trends in ecology and evolution*, 30(8): 463-469.
- Mohammadi, S. A., y Prasanna, B. M. (2003). Analysis of genetic diversity in crop plants—salient statistical tools and considerations. *Crop science*, 43(4): 1235-1248.

- Monmonier, M. S. (1973). Maximum-difference barriers: an alternative numerical regionalization method. *Geographical Analysis*, 5: 245-261
- Montaño-Lugo, M. L., Velasco Velasco, V. A., Ruíz Luna, J., Ángeles, C., Virginia, G., Rodríguez Ortiz, G., y Martínez, M. L. (2014). Contribución al conocimiento etnobotánico del chile de agua (*Capsicum annuum* L.) en los Valles Centrales de Oaxaca, México. *Revista mexicana de ciencias agrícolas*, 5(3): 503-511.
- Montes, H. S., y Eguiarte, L. E. (2002). Genetic structure and indirect estimates of gene flow in three taxa of *Cucurbita* (Cucurbitaceae) in western Mexico. *American Journal of Botany*, 89(7): 1156-1163.
- Muñoz-Zurita, R. (2013). Los chiles nativos de México (p. 256). DGE, El equilibrista-Conservas la Costeña, México.
- Nagy, I., Stagel, A., Sasvari, Z., Röder, M., y Ganal, M. (2007). Development, characterization, and transferability to other Solanaceae of microsatellite markers in pepper (*Capsicum annuum* L.). *Genome*, 50(7): 668-688.
- Narez-Jiménez, C. A., de-la-Cruz-Lázaro, E., Gómez-Vázquez, A., Castañón-Nájera, G., Cruz-Hernández, A., y Márquez-Quiroz, C. (2014). La diversidad morfológica in situ de chiles silvestres (*Capsicum* spp.) de Tabasco, México. *Revista fitotecnia mexicana*, 37(3): 209-215.
- Nei, M. (1987). *Molecular Evolutionary Genetics*. Columbia University Press, New York, NY, USA.
- Nei, M., Tajima, F., y Tatenó, Y. (1983). Accuracy of estimated phylogenetic trees from molecular data. *Journal of molecular evolution*: 19(2): 153-170.
- Nicolai, M., Cantet, M., Lefebvre, V., Sage-Palloix, A. M., y Palloix, A. (2013). Genotyping a large collection of pepper (*Capsicum* spp.) with SSR loci brings new evidence for

the wild origin of cultivated *C. annuum* and the structuring of genetic diversity by human selection of cultivar types. *Genetic resources and crop evolution*, 60(8): 2375-2390.

Nuez, F. V., Gil-Ortega, R., y Costa, J. G. (2003). *El cultivo de pimientos, chiles y ajíes* (p. 586). Ediciones Mundi-Prensa.

Nybom, H. (2004). Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Molecular ecology*, 13(5): 1143-1155.

Olsen, K. M., y Wendel, J. F. (2013). A bountiful harvest: genomic insights into crop domestication phenotypes. *Annual Review of Plant Biology*, 64: 47-70.

Onus, A. N., y Pickersgill, B. (2004). Unilateral incompatibility in *Capsicum* (Solanaceae): occurrence and taxonomic distribution. *Annals of Botany*, 94(2): 289-295.

Oyama, K., Hernández-Verdugo, S., Sánchez, C., González-Rodríguez, A., Sánchez-Peña, P., Garzón-Tiznado, J. A., y Casas, A. (2006). Genetic structure of wild and domesticated populations of *Capsicum annuum* (Solanaceae) from northwestern Mexico analyzed by RAPDs. *Genetic Resources and Crop Evolution*, 53(3): 553-562.

Pacheco-Olvera, A., Hernández-Verdugo, S., Rocha-Ramírez, V., González-Rodríguez, A., y Oyama, K. (2012). Genetic Diversity and Structure of Pepper (*L.*) from Northwestern Mexico Analyzed by Microsatellite Markers. *Crop Science*, 52(1): 231-241.

Page, R. D. M. (1996). TREEVIEW: An application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences*, 12: 357-358.

- Peery, M. Z., Kirby, R., Reid, B. N., Stoelting, R., DOUCET-BËER, E. L. E. N. A., Robinson, S., ... y PALSBOÏLL, P. J. (2012). Reliability of genetic bottleneck tests for detecting recent population declines. *Molecular Ecology*, 21(14): 3403-3418.
- Perfecto, I., Vandermeer, J. H., y Wright, A. L. (2009). *Nature's matrix: linking agriculture, conservation and food sovereignty* (p. 242). Routledge.
- Perry, L., y Flannery, K. V. (2007). Precolumbian use of chili peppers in the Valley of Oaxaca, Mexico. *Proceedings of the National Academy of Sciences*, 104(29): 11905-11909.
- Petit, R. J., El Mousadik, A., y Pons, O. (1998). Identifying populations for conservation on the basis of genetic markers. *Conservation biology*, 12(4): 844-855.
- Pickersgill, B. (1971). Relationships between weedy and cultivated forms in some species of chili peppers (genus *Capsicum*). *Evolution*, 25(4): 683-691.
- Pickersgill, B. (1997). Genetic resources and breeding of *Capsicum* spp. *Euphytica*, 96(1): 129-133.
- Pickersgill, B. (2007). Domestication of plants in the Americas: insights from Mendelian and molecular genetics. *Annals of botany*, 100(5): 925-940.
- Pickersgill, B. (2016). Chile Peppers (*Capsicum* spp.). En *Ethnobotany of Mexico* (pp. 207-231). Springer New York.
- Piñero, D., et al. (2008a). La diversidad genética como instrumento para la conservación y el aprovechamiento de la biodiversidad: estudios en especies mexicanas (pp. 437-494). En *Capital natural de México*, vol. I: Conocimiento actual de la biodiversidad. Conabio, México.

- Piñero, D., Barahona, A., Eguiarte, L., Rocha, A., y Salas, R. (2008b). La variabilidad genética de las especies: aspectos conceptuales y sus aplicaciones y perspectivas en México. *Conabio*, 1(1): 415-435.
- Piry, S., Luikart, G., y Cornuet, J. M. (1999). Computer note. BOTTLENECK: a computer program for detecting recent reductions in the effective size using allele frequency data. *Journal of heredity*, 90(4), 502-503.
- Pritchard, J. K., M. Stephens y Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155: 945-959.
- Putman, A. I., y Carbone, I. (2014). Challenges in analysis and interpretation of microsatellite data for population genetic studies. *Ecology and evolution*, 4(22): 4399-4428.
- Rendón-Aguilar, B., Bernal-Ramírez, L., y Sánchez-Reyes, G. (2017). Las plantas arvenses: más que hierbas del campo. *Oikos*, 17: 30-34.
- Rentería-Alcántara, M. (2007). Breve revisión de los marcadores moleculares (pp. 541-566). En: Eguiarte, L. E., Souza, V. y E. Aguirre. *Ecología molecular*. Instituto de Ecología, UNAM.
- Rodríguez, J., Peña Olvera, B. V., Gil Muñoz, A., Martínez Corona, B., Manzo, F., y Salazar Liendo, L. (2007). Rescate in situ del chilepoblano en Puebla, México. *Revista Fitotecnia Mexicana*, 30(1): 25-32.
- Romero, G., Adeva, C., y Z. Battad. (2009). Genetic fingerprinting: advancing the frontiers of crop biology research. *Philipp. Sci. Lett*, 2: 8-13.
- Rousset, F. (2008). genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Molecular ecology resources*, 8(1): 103-106.

- SAGARPA, Secretaria de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación (2016). 4to Informe de Labores 2015-2016. *Secretaria de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación* (SAGARPA). (154p). Disponible en: http://www.sagarpa.gob.mx/Transparencia/POT_2016/Informe/CuartoInformeDeLabores_SAGARPA.pdf
- Santiago-Luna, E. G., Carrillo-Rodríguez, J. C., Chávez-Servia, J. L., Enríquez del Valle, R., y Villegas-Aparicio, Y. (2016). Variación fenotípica en poblaciones de chile tusta. *Agronomía Mesoamericana*, 27(1): 139-149.
- Sato, S., Isobe, S., Asamizu, E., Ohmido, N., Kataoka, R., Nakamura, Y., ... y Sasamoto, S. (2005). Comprehensive structural analysis of the genome of red clover (*Trifolium pratense* L.). *DNA Research*, 12(5): 301-364.
- Sexton, J. P., Hangartner, S. B., y Hoffmann, A. A. (2014). Genetic isolation by environment or distance: which pattern of gene flow is most common?. *Evolution*, 68(1): 1-15.
- Shirasawa, K., Ishii, K., Kim, C., Ban, T., Suzuki, M., Ito, T., Muranaka, T., Kobayashi, M., Nagata, N., Isobe, S., y Tabata, S. (2013). Development of *Capsicum* EST-SSR markers for species identification and *in silico* mapping onto the tomato genome sequence. *Mol Breed*, 31: 101-110.
- SIAP, Sistema de Información Agroalimentaria y Pesquera (2015) Anuario Estadístico de la Producción Agrícola 2015. Sistema de Información Agroalimentaria y Pesquera de la Secretaria de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación. Disponible en: http://infosiap.siap.gob.mx/aagricola_siap_gb/ientidad/index.jsp (Agosto 2017).

- Simpson, J. (1997). Molecular markers. *Boletín de la Sociedad Botánica de México*, 60: 73-76.
- Slatkin, M. (1987). Gene flow and the geographic structure of natural populations. *Science*, 236(4803): 787-792.
- Slatkin, M. (1995). A measure of population subdivision based on microsatellite allele frequencies. *Genetics*, 139(1), 457-462.
- Szpiech, Z. A., Jakobsson, M., y Rosenberg, N. A. (2008). ADZE: a rarefaction approach for counting alleles private to combinations of populations. *Bioinformatics*, 24(21): 2498-2504.
- Takezaki, N., y Nei, M. (1996). Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics*, 144(1): 389-399.
- Takezaki, N., y Nei, M. (2008). Empirical tests of the reliability of phylogenetic trees constructed with microsatellite DNA. *Genetics*, 178(1): 385-392.
- Tang, H., Sezen, U., y A. H. Paterson. (2010). Domestication and plant genomes. *Current opinion in plant biology*, 13(2): 160-166.
- Toledo-Aguilar, R., López-Sánchez, H., Santacruz-Varela, A., Valadez-Moctezuma, E., López, P. A., Aguilar-Rincón, V. H.e,... y Vaquera-Huerta, H. (2016). Characterization of genetic diversity of native Ancho chili populations of Mexico using microsatellite markers. *Chilean journal of agricultural research*, 76(1): 18-26.
- Van de Wouw, M., Kik, C., van Hintum, T., van Treuren, R., y Visser, B. (2009). Genetic erosion in crops: concept, research results and challenges. *Plant Genetic Resources*, 8(1): 1-15.
- Vavilov, N. I. (1926). The origin of the cultivation of primary crops, in particular cultivated hemp. *Studies on the origin of cultivated plants*, 221-233.

- Wang, I. J., y Bradburd, G. S. (2014). Isolation by environment. *Molecular Ecology*, 23(23): 5649-5662.
- Weir, B. S., y Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, 1358-1370.
- Young, A., Boyle, T., y Brown, T. (1996). The population genetic consequences of habitat fragmentation for plants. *Trends in Ecology and Evolution*, 11(10): 413-418.
- Yumnam, J. S., Tyagi, W., Pandey, A., Meetei, N. T., y Rai, M. (2012). Evaluation of genetic diversity of chilli landraces from North Eastern India based on morphology, SSR markers and the Pun1 locus. *Plant Molecular Biology Reporter*, 30(6): 1470-1479.
- Zeder, M. A., Emshwiller, E., Smith, B. D., y Bradley, D. G. (2006). Documenting domestication: the intersection of genetics and archaeology. *TRENDS in Genetics*, 22(3): 139-155.
- Zhang, X. M., Zhang, Z. H., Gu, X. Z., Mao, S. L., Li, X. X., Chadœuf, J., ... y Zhang, B. X. (2016). Genetic diversity of pepper (*Capsicum* spp.) germplasm resources in China reflects selection for cultivar types and spatial distribution. *Journal of Integrative Agriculture*, 15(9): 1991-2001.
- Zohary, D. (2004). Unconscious selection and the evolution of domesticated plants. *Economic botany*, 58(1): 5-10.

ANEXO 1

Fotos de las variedades de chiles *C. annuum* analizados. El chile Serrano Criollo no tuvo registro fotográfico. Fotos tomadas por Cristina Alonso Fernández



ANEXO 2

Extracción de ADN, protocolo CTAB-miniprep (Doyle y Doyle, 1987). Modificado

- 1) Colocar alrededor de 1g de tejido fresco en un tubo ependorf de 2µl, poner dos balines entre el tejido (balín-tejido-balín).
- 2) Preparar un stock de CTAB 2X con Mercaptoetanol (81.6ml de CTAB 2x y 244.8µl de Mercaptoetanol, solución para 48 muestras). Agregar 5µl de RNasa (por muestra).
- 3) Agregar 1ml de buffer CTAB2x-Mercaptoetanol-RNasa a cada muestra.
- 4) Colocar tubos en el Tissue Lyser y moler el tejido. Rotar tubos 2 veces cada tres minutos.
- 5) Recuperar los balines trasladando la mayor cantidad del tejido molido a un tubo nuevo de 2ml y centrifugar a 10, 000 rpm durante 8 minutos a 4°C.
- 6) Eliminar el sobrenadante y resuspender con 600µl de buffer CTAB2x-Mercaptoetanol-RNasa. Incubar en baño María a 60°C durante 10 minutos.
- 7) Agregar a cada tubo 600µl de cloroformo: octanol 24:1 a temperatura ambiente y agitar hasta homogenizar. Centrifugar a 12,000 rpm durante 10 minutos a 4°C
- 8) Trasladar el sobrenadante a un tubo nuevo de 1.5ml y precipitar el ADN con 2/3 (400µl) del volumen final (600µl) de isopropanol frío. Dejar reposar 12 horas a -20°C.
- 9) Centrifugar a 10, 000 rpm durante 10 minutos a 4°C. Eliminar el sobrenadante.
- 10) El pellet debe limpiarse agregando 1ml de etanol al 70% frío.
- 11) Centrifugarse a 10,000 rpm durante 9 minutos. Eliminar el sobrenadante.
- 12) Repetir el paso 10 y 11 dos veces más, centrifugando a 10, 000 rpm durante 7 minutos.
- 13) Eliminar el sobrenadante y dejar secar la pastilla en la campana.
- 14) Resuspender el pellet con 200µl de DH₂O.

ANEXO 3

SCRIPTS para análisis genéticos en R

INSTALAR LIBRERÍAS

```
install.packages("adegenet")
install.packages("ade4")
install.packages("pegas")
install.packages("ape")
install.packages("hierfstat")
install.packages("fossil")
githubinstall
### Para instalar gstudio y popgraph ###
library(githubinstall)
githubinstall("dyerlab/gstudio")
githubinstall("dyerlab/popgraph")
githubinstall("dyerlab/igraph")
```

CARGAR LIBRERÍAS

```
### Para cargar gstudio y popgraph ###
library(githubinstall)
require(gstudio)
require(popgraph)
require(igraph)
```

CARGAR ARCHIVOS

```
### Formato genepop ###
chiles<-read.genepop("Nombre.gen")
### Formato genepop para gstudio ###
chilesgs<-read.population("Nombre.gen", type="genepop")
### Formato csv para popgraph ###
chilespg<-read_population("Nombre.csv", type= "separated",
locus.columns=c(7,8,9,10,11,12,13,14,15,16), header= TRUE, sep= ",")
### Archivo csv con coordenadas geográficas ###
xy<-read.csv("Nombre.csv", header=TRUE, sep=",")
```

ANÁLISIS GENÉTICOS

Alelos efectivos con GSTUDIO

```
library(githubinstall)
require(gstudio)
chilesgs<-read.population("ChilesGenepop.gen", type= "genepop")
chilespart<-partition(chilesgs, stratum= "Population")
chilesAe<-genetic_diversity(chilesgs, mode= "Ae")
chilesAePop<-lapply(chilespart, genetic_diversity)
```

Distancias geográficas con GGMAP en GSTUDIO (base 2 columnas longitud y latitud)

```
Geo<- read.csv("Coor.csv", header= FALSE, sep= ",")
pDistGeo<- Geo$Lon, Geo$Lat), method = "euclidean", diag = TRUE, upper= FALSE)
```

Barreras geográficas con ADEGENET

```
library(adegenet)
xy<-read.csv("Coordenadas.csv", header= TRUE)
cn1<-chooseCN(xy, type= 1, ask= FALSE) ## type puede ser de 1 a 4 ##
chiles<-read.genepop("ChilesGenepop.gen")
### Obtener distancias genéticas ###
Matriz con distancias genéticas  $R_{ST}$  en EXCEL ###Leer matriz con la librería Rcmdr###
### Elegir threshold "d" por default ###
plot(m1, col= "springgreen3")
om1<-optimize.monmonier(xy, distanciasfst, cn1)
plot(om1, method= "greylevel", add.arr= FALSE, bwd= 6, col= "blue", csize= .5)
```

Análisis multivariado con ADEGENET

```
library(adegenet)
chiles<-read.genepop("ChilesGenepop.gen")
MultChiles<-dapc(chiles)
### Elegir el número de discriminantes y pcas ###
summary(MultChiles)
### Figura con leyenda ###
scatter(MultChiles, col=c("darkred", "coral", "darkviolet", ... ), cex= 2, legend= TRUE,
clabel= FALSE, posi.leg= "bottomleft", scree.pca= TRUE, posi.pca= "topleft", cleg= 0.75,
xax= 1, yax= 2, inset.solid= 1)
### Figura sin leyenda ###
scatter(MultChiles, clabel=.8, col=c("darkred", "coral", "darkviolet", ...), scree.pca=
TRUE, posi.pca= "topleft", cleg= 0.75, xax= 1, yax= 2, inset.solid= 1)
```

Análisis gráfico poblacional con GITHUBINSTALL

```
### Cargar archivo con datos genéticos ###
chilespg<-read_population("ChilesGraph.csv", type= "separated",
locus.columns=c(7,8,9,10,11,12,13,14,15,16), header= TRUE, sep= ",")
### Obtener el gráfico poblacional más básico ###
ChilesPob<-to_mv(chilespg)
ChilesPoblas<-chilespg$Population
popgraph<-popgraph(ChilesPob, ChilesPoblas, alpha= 0.05, tol= 1e-04)
plot(popgraph, edge.color= "yellow3", vertex.label.color= "turquoise4", vertex.color=
"springgreen3", vertex.label.dist= 2)
E(popgraph)$weight
```

Diferenciar los grupos en el gráfico

```
V(popgraph)$name <- c("ABA", "AL", "CUI", "CUE" ...)
V(popgraph)$Region <- c("Valles Centrales", "Costa", "Cañada", "Cañada" ...)
V(popgraph)$color <-c("springgreen2", "royalblue2", "yellow3", "yellow3", ...)
```

```

list.vertex.attributes(popgraph)
plot(popgraph, edge.color= "purple", vertex.label.color= "black", vertex.label.dist= 1.8)

### Cargar archivo con datos geográficos y ambientales ###
coordenadas<-read.csv("Coordenadas.csv", header=TRUE, sep=",")

### Sobrelapar el gráfico en un mapa ###
require(ggmap)
geo <- decorate_graph(popgraph, coordenadas, stratum="Population")
ubicacion <- c(mean(V(geo)$Lon), mean(V(geo)$Lat))
mapa<- get_map(location, matype="toner-lite", zoom=7)
p <- ggmap(mapa)
p <- p + geom_edgeset(aes(x=Lon,y=Lat), geo, color="black")
p <- p + geom_nodeset(aes(x=Lon, y=Lat, size=size), geo)
p + xlab("Longitud") + ylab("Latitud")

### Análisis de aislamiento por distancia ###
cGD <- to_matrix(geo, mode="shortest path")
pDistGeo<- dist(cbind(V(geo)$Lon, V(geo)$Lat), method = "euclidean", diag = TRUE,
upper= FALSE)
pDistAlt<- dist(cbind(V(geo)$Alt), diag=FALSE, upper=FALSE)

### Prueba de Mantel con ADE4 ###
library(ade4)
cGDdist<-as.dist(cGD)
pDistGeodist<-as.dist(pDist)
pDistAltdist<-as.dist(pDistAltdist)
mantel1 <-mantel.randtest(cGDdist, pDistGeodist, nrepet = 1000)
mantel2 <-mantel.randtest(cGDdist, pDistAltdist, nrepet = 1000)
plot (pDistAltdist, cGDdist, xlab="Distancia Altitudinal (m)", ylab="Distancia genética
condicional (cGD)", pch=20, col="darkmagenta")
Regresión<-lm(as.vector(cGDdist) ~ as.vector(pDistAltdist))
abline (Regresión, col="black", lwd=1)

```

ANEXO 4

Frecuencias de alelos nulos por locus en las poblaciones

Población	CaES4787	CaES1003	CaES3862	CaES2332	CaES0425	CaES4584	CaES4192	CaES5392	EPMS386	EPMS310	Promedio
ROS	0.00	0.00	0.15	0.06	0.17	0.00	0.00	0.00	0.17	0.04	0.06
COA	0.00	0.00	0.00	0.09	0.33*	0.00	0.29*	0.15	0.33*	0.41*	0.16
JDI	0.10	0.00	0.00	0.24*	0.20*	0.00	0.12	0.00	0.24*	0.19	0.11
AL	0.10	0.16	0.25*	0.00	0.24*	0.00	0.25*	0.11	0.02	0.00	0.11
CUE	0.00	0.00	0.00	0.22*	0.18	0.00	0.26*	0.00	0.08	0.14	0.09
CUI	0.00	0.00	0.00	0.09	0.26*	0.00	0.21*	0.00	0.29*	0.16	0.10
RAN	0.00	0.00	0.04	0.00	0.26*	0.00	0.31*	0.00	0.23*	0.34*	0.12
EJU	0.00	0.00	0.00	0.00	0.27*	0.18	0.22*	0.00	0.35*	0.31*	0.13
LOB	0.17	0.15	0.00	0.02	0.27*	0.00	0.26*	0.00	0.15	0.28*	0.13
TOS	0.00	0.00	0.00	0.25*	0.32*	0.00	0.32*	0.00	0.19	0.42*	0.15
SDO	0.18	0.00	0.00	0.18	0.24*	0.00	0.11	0.00	0.28*	0.31*	0.13
ETL	0.00	0.00	0.00	0.00	0.00	0.00	0.26*	0.00	0.26*	0.15	0.07
REY	0.00	0.00	0.00	0.33*	0.20*	0.00	0.30*	0.20*	0.26*	0.00	0.13
MIX	0.21*	0.27*	0.29*	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.09
SCR	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.29*	0.00	0.03
ABA	0.21*	0.21*	0.16	0.00	0.32*	0.00	0.00	0.00	0.26*	0.25*	0.14
POL	0.00	0.00	0.00	0.00	0.00	0.00	0.24*	0.21*	0.28*	0.42*	0.12
TOM	0.00	0.00	0.30*	0.28*	0.00	0.39*	0.18	0.00	0.22*	0.18	0.15
LAB	0.00	0.00	0.00	0.22*	0.00	0.00	0.22*	0.00	0.38*	0.33*	0.12
ZAC	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.22*	0.15	0.04
HUA	0.00	0.00	0.17	0.00	0.25*	0.00	0.29*	0.00	0.37*	0.00	0.11
JGB	0.00	0.00	0.00	0.00	0.00	0.00	0.25*	0.00	0.35*	0.29*	0.09
BAL	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.20*	0.00	0.02
COY	0.00	0.00	0.00	0.29*	0.29*	0.00	0.32*	0.00	0.24*	0.39*	0.15
SPA	0.00	0.00	0.00	0.15	0.00	0.00	0.00	0.00	0.29*	0.24*	0.07
TEZ	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.29*	0.00	0.03
COP	0.00	0.00	0.00	0.32*	0.00	0.00	0.32*	0.00	0.00	0.40*	0.10
PIN	0.00	0.00	0.00	0.33*	0.00	0.00	0.00	0.00	0.18	0.33*	0.08
Promedio	0.04	0.03	0.05	0.11	0.14	0.02	0.17	0.02	0.23*	0.20*	

* Frecuencias altas (≥ 0.20).

ANEXO 5

Desequilibrio de ligamiento

Locus 1	Locus 2	X²	g.l	p
CaES4787	CaES1003	37.26	14	0.001*
CaES4787	CaES3862	33.93	18	0.013*
CaES1003	CaES3862	41.55	18	0.001*
CaES4787	CaES2332	15.004	18	0.661
CaES1003	CaES2332	11.93	10	0.290
CaES3862	CaES2332	20.89	28	0.830
CaES4787	CaES0425	6.21	16	0.986
CaES1003	CaES0425	1.67	10	0.998
CaES3862	CaES0425	7.87	22	0.997
CaES2332	CaES0425	21.75	26	0.702
CaES4787	CaES4584	1.19	6	0.977
CaES1003	CaES4584	1.81	4	0.771
CaES3862	CaES4584	5.20	10	0.877
CaES2332	CaES4584	0.00	8	1.000
CaES0425	CaES4584	0.48	8	0.999
CaES4787	CaES4192	16.72	18	0.542
CaES1003	CaES4192	13.40	14	0.495
CaES3862	CaES4192	42.71	30	0.062
CaES2332	CaES4192	45.14	36	0.141
CaES0425	CaES4192	9.80	30	0.999
CaES4584	CaES4192	14.30	8	0.074
CaES4787	CaES5392	4.00	10	0.950
CaES1003	CaES5392	1.43	4	0.840
CaES3862	CaES5392	2.71	12	0.997
CaES2332	CaES5392	7.14	12	0.848
CaES0425	CaES5392	4.76	12	0.966
CaES4584	CaES5392	19.04	6	0.004*
CaES4192	CaES5392	17.39	16	0.360
CaES4787	EPMS386	24.04	22	0.345
CaES1003	EPMS386	6.98	18	0.990
CaES3862	EPMS386	6.97	32	0.999
CaES2332	EPMS386	48.81	40	0.160
CaES0425	EPMS386	26.42	30	0.653
CaES4584	EPMS386	3.44	8	0.903
CaES4192	EPMS386	13.89	40	0.999
CaES5392	EPMS386	11.53	18	0.870
CaES4787	EPMS310	20.68	20	0.416
CaES1003	EPMS310	1.69	10	0.998
CaES3862	EPMS310	10.80	24	0.990
CaES2332	EPMS310	38.90	36	0.341
CaES0425	EPMS310	20.42	26	0.771
CaES4584	EPMS310	4.86	12	0.962
CaES4192	EPMS310	21.74	34	0.949
CaES5392	EPMS310	6.95	14	0.936
EPMS386	EPMS310	76.21	44	0.002*

*significativos ($p < 0.05$)

ANEXO 6

Resultados de la prueba ANOVA

ANOVA para valores de H_E por región geográfica					
Fuente de variación	Suma de cuadrados	Grados de libertad	Varianza	F	p
Entre grupos	0.029	2	0.015	0.214	0.809
Dentro de los grupos	1.573	23	0.068		
Total	1.602	25			

ANOVA para valores de H_o por región geográfica					
Fuente de variación	Suma de cuadrados	Grados de libertad	Varianza	F	p
Entre grupos	0.005	2	0.003	0.118	0.889
Dentro de los grupos	0.526	23	0.023		
Total	0.531	25			

ANOVA para valores de H_E por grado de domesticación					
Fuente de variación	Suma de cuadrados	Grados de libertad	Varianza	F	p
Entre grupos	0.025	2	0.013	0.186	0.832
Dentro de los grupos	0.811	12	0.068		
Total	0.836	14			

ANOVA para valores de H_o por grado de domesticación					
Fuente de variación	Suma de cuadrados	Grados de libertad	Varianza	F	p
Entre grupos	0.005	2	0.002	0.089	0.915
Dentro de los grupos	0.307	12	0.026		
Total	0.312	14			

ANOVA para valores de F_{IS} por población					
Fuente de variación	Suma de cuadrados	Grados de libertad	Varianza	F	p
Modelo	0.017	2	0.008	0.033	0.968
Error	5.912	23	0.257		
Total	5.929	25			

ANOVA para valores de F_{IS} por variedad					
Fuente de variación	Suma de cuadrados	Grados de libertad	Varianza	F	p
Modelo	0.248	2	0.124	0.813	0.467
Error	1.835	12	0.153		
Total	2.083	14			

ANEXO 7

Valores F_{CT} obtenidos de las 15 corridas en el análisis SAMOVA

SAMOVA con grupos externos		SAMOVA sin grupos externos	
K	F_{CT}	K	F_{CT}
2	-0.012	2	0.050
3	-0.021	3	0.013
4	-0.025	4	-0.021
5	-0.019	5	-0.019
6	0.067	6	0.033
7	-0.024	7	0.009
8	-0.035	8	-0.004
9	-0.005	9	-0.001
10	0.002	10	0.048
11	-0.007	11	0.029
12	0.018	12	0.025
13	-0.070	13	0.038
14	-0.040	14	0.020
15	0.019	15	-0.014

ANEXO 8

Valores obtenidos del DAPC

Probabilidad de asignación y número de individuos asignados a cada grupo por DAPC.

Por población con grupos externos			
Población	Probabilidad de asignación	Tamaño del grupo inicial	Tamaño del grupo posterior
CUE	0.37	19	21
CUI	0.74	19	83
COA	0.58	24	72
EJU	0.00	17	0
LOB	0.00	16	0
ETL	0.00	11	0
MIX	0.00	10	0
SCR	0.00	10	0
ABA	0.00	9	0
LAB	0.00	8	0
ZAC	0.25	8	4
JGB	0.00	6	0
BAL	0.00	5	1
SPA	0.00	4	0
ROS	0.58	33	75
JDI	0.87	24	31
AL	0.00	22	2
RAN	0.00	19	0
TOS	0.53	15	24
SDO	0.85	13	11
REY	0.00	11	0
POL	0.33	9	14
TOM	0.11	9	1
HUA	0.00	6	0
COY	0.00	4	0
TEZ	0.00	4	0
COP	0.00	3	2
PIN	0.00	3	0

Por región geográfica con grupos externos			
Región geográfica	Probabilidad de asignación	Tamaño del grupo inicial	Tamaño del grupo posterior
Cañada	0.32	38	21
Valles Centrales	0.86	128	151
Costa	0.86	175	169

Probabilidad de asignación y número de individuos asignados a cada grupo por DAPC.

Por población sin grupos externos			
Población	Probabilidad de asignación	Tamaño del grupo inicial	Tamaño del grupo posterior
CUE	0.00	19	0
CUI	0.57	19	56
COA	0.50	24	89
EJU	0.00	17	4
LOB	0.19	16	20
ETL	0.00	11	0
MIX	0.00	10	0
SCR	0.00	10	0
ABA	0.00	9	0
LAB	0.00	8	0
ZAC	0.17	6	1
JGB	0.00	5	0
BAL	0.00	5	0
SPA	0.00	4	0
ROS	0.67	33	71
JDI	0.88	24	31
AL	0.00	22	0
RAN	0.00	17	0
TOS	0.14	14	15
SDO	0.82	11	9
REY	0.00	5	0
POL	0.11	9	4
TOM	0.56	9	16
HUA	0.00	4	0
COY	0.00	4	0
TEZ	0.50	4	4
COP	0.00	3	2

Por región geográfica sin grupos externos			
Región geográfica	Probabilidad de asignación	Tamaño del grupo inicial	Tamaño del grupo posterior
Cañada	0.32	38	21
Valles Centrales	0.87	125	152
Costa	0.84	159	149

Probabilidad de asignación y número de individuos asignados a cada grupo por DAPC.

Por variedad			
Variedad	Probabilidad de asignación	Tamaño del grupo inicial	Tamaño del grupo posterior
Agua	0.91	79	168
Taviche	0.00	21	0
Árbol	0.42	12	6
Huacles	0.00	12	0
Paradito	0.00	10	0
Serrano Criollo	0.00	9	0
Nanchita	0.00	8	0
Chilcostle Rojo	0.00	4	0
Achilito	0.00	3	0
Guiña Dahni	0.00	17	0
Costeño Rojo	0.58	45	62
Costeño Amarillo	0.00	2	0
Tusta	0.80	35	37
Mirasol	0.00	4	0
Arvenses	0.61	33	36
Silvestres	0.70	23	24
<i>C. frutescens</i>	0.00	16	0
<i>C. chinense</i>	0.67	3	3
Por grado de domesticación			
Grado de domesticación	Probabilidad de asignación	Tamaño del grupo inicial	Tamaño del grupo posterior
Cultivados	0.97	261	274
Arvenses	0.70	33	43
Silvestres	0.57	23	17
<i>C. frutescens</i>	0.00	16	0
<i>C. chinense</i>	0.67	3	2

Eig-valores obtenidos del DAPC por población y región geográfica.

DAPC poblacional con grupos externos		DAPC poblacional sin grupos externos		DAPC por variedad	
Eig-valores		Eig-valores		Eig-valores	
1	38.72	1	56.22	1	56.22
2	26.73	2	32.21	2	32.21
3	14.04	3	21.27	3	21.27
4	12.47	4	19.57	4	19.57
5	8.58	5	11.59	5	11.59
6	7.83	6	10.37	6	10.37
7	6.27	7	8.95	7	8.95
8	5.11	8	4.81	8	4.81
9	4.25	9	3.91	9	3.91
10	3.75	10	3.08	10	3.08
11	3.40	11	2.12	11	2.12
12	2.60	12	1.89	12	1.89
13	1.92	13	1.83	13	1.83
14	1.71	14	0.98	14	0.98
15	1.42	15	0.60	15	0.60
16	1.33	16	0.15	16	0.15
17	1.07	17	0.11	17	0.11
18	0.94	18	0.69		
19	0.75	19	0.39		
20	0.37	20	0.27		
21	0.16	21	0.12		
22	0.10	22	0.08		
23	0.04	23	0.03		
24	0.02	24	0.02		
25	0.01	25	0.004		
DAPC regional con grupos externos		DAPC regional sin grupos externos		DAPC por grado de domesticación	
Eig-valores		Eig-valores		Eig-valores	
1	212.262	1	198.24	1	161.97
2	45.77	2	44.36	2	57.01
				3	38.76
				4	11.97

ANEXO 9

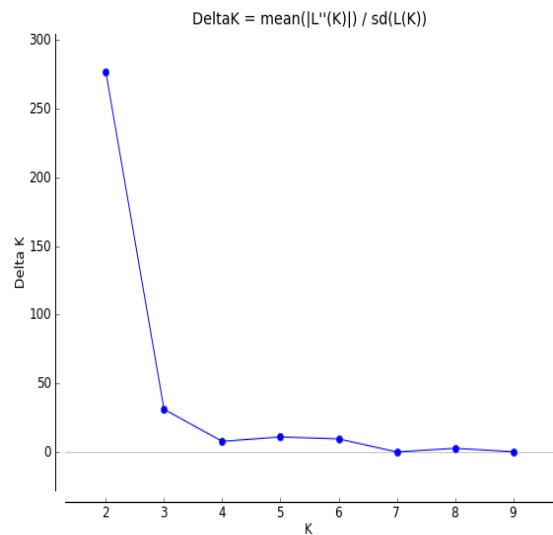
Valores obtenidos de STRUCTURE

Resultados de la prueba de Evanno, *et al.*, 2005.

K	Para poblaciones					Para variedades				
	L(K)	SD	[L'(K)]	L''(K)	ΔK	L(K)	SD	L'(K)	[L''(K)]	ΔK
1	-4614.62	0.04	-	-	-	-4540.71	0.03	-	-	-
2	-3983.69	0.85	630.93	350.16	413.66	-3916.05	0.96	624.66	353.42	369.25
3	-3702.92	0.63	280.77	147.05	231.99	-3644.81	0.63	271.24	135.84	215.71
4	-3569.2	6.44	133.72	8.3	1.29	-3509.41	1.84	135.40	29.44	16.03
5	-3427.18	22.40	142.02	46.99	2.10	-3403.45	43.57	105.96	47.51	1.09
6	-3332.15	43.58	95.03	0.8	0.02	-3345.00	67.95	58.45	104.31	1.54
7	-3236.32	138.58	95.83	44.28	0.32	-3182.24	71.25	162.76	65.46	0.92
8	-3096.21	2.76	140.11	82.06	29.74	-3084.94	99.19	97.30	43.42	0.44
9	-3038.16	9.63	58.05	30.14	3.13	-3031.06	47.16	53.88	6.95	0.15
10	-3010.25	21.56	27.91	-	-	-2970.23	23.15	60.83	-	-

Resultados de la prueba de Evanno, *et al.*, 2005 para las variedades cultivadas.

K	L(K)	SD	Ln'(K)	[Ln''(K)]	ΔK
1	-3170.35	0.05	-	-	-
2	-2789.28	0.79	381.07	217.73	276.62
3	-2625.94	4.20	163.34	131.35	31.25
4	-2593.95	19.64	31.99	156.07	7.95
5	-2405.89	7.37	188.06	82.16	11.15
6	-2299.99	8.48	105.90	82.26	9.70
7	-2276.35	10.51	23.64	2.38	0.23
8	-2250.33	12.21	26.02	35.38	2.90
9	-2188.93	20.15	61.40	7.41	0.37
10	-2134.94	7.82	53.99	-	-



Probabilidad de asignación de los individuos en las poblaciones definidas para cada grupo ($K=2$) obtenidos con el programa STRUCTURE.

Población	Grupo 1	Grupo 2
Valles Centrales		
COA	0.82	0.18
EJU	0.83	0.17
LOB	0.86	0.14
ETL	0.90	0.10
MIX	0.83	0.17
SCR	0.96	0.04
ABA	0.79	0.21
LAB	0.90	0.10
ZAC	0.98	0.03
JGB	0.93	0.07
BAL	0.13	0.87
SPA	0.93	0.07
Cañada		
CUE	0.83	0.17
CUI	0.93	0.07
Costa		
ROS	0.44	0.56
JDI	0.05	0.95
AL	0.29	0.71
RAN	0.70	0.30
TOS	0.19	0.81
SDO	0.03	0.98
REY	0.08	0.93
POL	0.09	0.91
TOM	0.08	0.92
HUA	0.07	0.93
COY	0.26	0.74
TEZ	0.06	0.94
COP	0.19	0.82
PIN	0.36	0.64

Probabilidad de asignación de los individuos en las variedades y variedades cultivadas definidas para cada grupo ($K=2$) obtenidos con el programa STRUCTURE.

Variedades			Variedades cultivadas		
Población	Grupo 1	Grupo 2	Población	Grupo 1	Grupo 2
Agua	0.14	0.86	Agua	0.18	0.82
Taviche	0.16	0.84	Taviche	0.20	0.80
Árbol	0.26	0.74	Árbol	0.28	0.72
Huacles	0.03	0.97	Huacles	0.04	0.96
Paradito	0.20	0.80	Paradito	0.20	0.80
Serrano Criollo	0.13	0.87	Serrano Criollo	0.13	0.87
Nanchita	0.07	0.93	Nanchita	0.08	0.92
Chilcostle Rojo	0.13	0.87	Chilcostle Rojo	0.15	0.85
Achilito	0.17	0.83	Achilito	0.20	0.80
Guiña Dahni	0.23	0.77	Guiña Dahni	0.30	0.70
Costeño Rojo	0.60	0.40	Costeño Rojo	0.69	0.31
Costeño Amarillo	0.89	0.11	Costeño Amarillo	0.95	0.05
Tusta	0.93	0.07	Tusta	0.94	0.06
Mirasol	0.95	0.05	Mirasol	0.91	0.09
Arvenses	0.88	0.12			
Silvestres	0.84	0.16			
Grupos externos					
<i>C. frutescens</i>	0.72	0.28			
<i>C. chinense</i>	0.97	0.03			