



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
PROGRAMA DE DOCTORADO EN CIENCIAS BIOMÉDICAS
INSTITUTO DE ECOLOGÍA

DIVERSIDAD, RECURSOS GENÉTICOS Y GENÉTICA DE POBLACIONES EN
CALABAZAS (*Cucurbita*),
CON ÉNFASIS EN *Cucurbita moschata* Duchesne.

TESIS
QUE PARA OPTAR POR EL GRADO DE
DOCTORA EN CIENCIAS

PRESENTA
HELENA SOCORRO HERNÁNDEZ ROSALES

DIRECTOR DE TESIS:
DR. LUIS ENRIQUE EGUIARTE FRUNS
Instituto de Ecología, UNAM.

COMITÉ TUTOR
DR. RAFAEL LIRA SAADE
Facultad de Estudios Superiores Iztacala, UNAM.
DRA. ANA ELENA ESCALANTE HERNÁNDEZ
Instituto de Ecología, UNAM.

CIUDAD UNIVERSITARIA, CDMX.
JULIO 2021



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.

Agradecimientos institucionales.

Agradezco a la Universidad Nacional Autónoma de México (UNAM) por darme la oportunidad de realizar mi formación académica con la más alta calidad, partiendo desde el bachillerato (CCH Vallejo), la licenciatura (FES Iztacala) hasta el posgrado.

Gracias al Programa de Doctorado en Ciencias Biomédicas de la UNAM por permitirme llevar a cabo mis estudios de posgrado en el Instituto de Ecología de la UNAM en Ciudad Universitaria.

Al Consejo Nacional de Ciencia y Tecnología (CONACyT) por brindarme el apoyo económico durante la realización de mis estudios de doctorado con la Beca no. 385367. Al Programa de Apoyo a los Estudios de Posgrado (PAEP) de la UNAM por el financiamiento otorgado para poder asistir al Curso Internacional “Domesticación, manejo y conservación *in situ* de recursos genéticos”, que se llevó a cabo en la Universidad Nacional Agraria la Molina (UNALM) en Perú.

Esta tesis se pudo llevar a cabo gracias al financiamiento otorgado por la Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO) con los proyectos KE004 “Diversidad genética de las especies de *Cucurbita* en México e hibridación entre plantas genéticamente modificadas y especies silvestres de *Cucurbita*” y PE001 “Diversidad genética de las especies de *Cucurbita* en México. Fase II. Genómica evolutiva y de poblaciones, recursos genéticos y domesticación”. También se agradece al proyecto de Ciencia Básica CONACyT 2011.167826 “Genómica de poblaciones: estudios en el maíz silvestre, el teosinte (*Zea mays* ssp. *parviglumis* y *Zea mays* ssp. *mexicana*) y al proyecto de Problemas Nacionales (CONACyT) 247730 por el apoyo económico otorgado.

Agradezco el soporte técnico y ayuda brindada por parte del Biól. Marco Tulio Solano de la Cruz responsable del Laboratorio de Genética Molecular y a Rodrigo García Herrera, técnico académico a cargo de supercómputo en el LANCIS. Así mismo, agradezco a la Lic. Erika Rodríguez Reyes por su valiosa ayuda en todos los trámites del posgrado. Todos ellos del Instituto de Ecología.

Agradezco a los miembros del jurado de examen doctoral constituido por:

Presidenta: Dra. María Esperanza Martínez Romero

Secretario: Dr. Luis Enrique Eguiarte Fruns

Vocal: Dr. Antonio González Rodríguez

Vocal: Dr. Juan Servando Núñez

Vocal: Dra. Eria Alaide Rebollar Caudillo

quienes enriquecieron esta tesis con cada uno de sus comentarios y observaciones.

Finalmente agradezco a mi comité tutor integrado por el Dr. Luis E. Eguiarte Fruns del Instituto de Ecología UNAM, el Dr. Rafael Lira Saade de la UBIPRO de la Facultad de Estudios Superiores Iztacala UNAM, y la Dra. Ana Elena Escalante Hernández del Instituto de Ecología UNAM.

Agradecimientos a título personal.

Agradezco al Dr. Luis Eguiarte y al Dr. Rafael Lira por el voto de confianza y permitirme participar en el proyecto de las calabazas. Agradezco al Dr. Luis y a la Dra. Valeria Souza por permitirme realizar mis estudios de posgrado en su laboratorio y por el apoyo académico y personal durante estos años.

Agradezco el total apoyo técnico y logístico por parte de la Dra. Erika Aguirre Planter. Eri muchas gracias por recibirme y hacerme sentir parte del laboratorio desde el primer día. Gracias por tu valiosa amistad, eres una mujer admirable.

Agradezco a todos los miembros del Laboratorio de Evolución Molecular y Experimental quienes a lo largo de estos años me ayudaron desinteresadamente en las salidas de campo, trabajo de laboratorio, atendiendo mis dudas y enriqueciendo mi formación a partir de sus comentarios y sugerencias. Sobre todo, les agradezco la amistad, los momentos y experiencias que compartimos: Diego, Felicitas, Gabriel, Javier Luna, Jonás, Manuelín, Manuel Rosas, Margaux , Mariette, Mirna, Niza, Roberto, Santiago, Valerie, Yaxal y Yocelyn. Gracias Sra. Silvia Barrientos por la ayuda brindada en todo momento y sobre todo por ese ímpetu por mantener unido al laboratorio.

Mi formación doctoral no hubiera sido la misma sin la dirección, dedicación y paciencia de la Dra. Gabriela Castellanos Morales. Gaby, eres sin duda uno de los más importantes pilares en esta etapa de mi vida. Simplemente no tengo suficientes palabras para expresar toda la admiración, el agradecimiento y el cariño incondicional que siento por ti. **Gracias.**

Les agradezco a todos los miembros del *Team Cucurbita*. Al M. en C. Guillermo Sánchez quién me insistió y enseñó a trabajar en campo. Memo, agradezco enormemente tu sincera amistad y por compartir conmigo el desarrollo y dedicación a este proyecto. Gracias por las extensas charlas académicas y personales, gracias por ser un increíble ser humano conmigo. Agradezco a los técnicos y *posdocs* que enriquecieron el proyecto general y nos apoyaron en todo momento: Dr. Enrique Scheinvar, Dr. Jaime Gasca y Dra. Xitlali Aguirre. Muchas gracias a las calabacitas

tiernas: Biol. Leslie, Ing. Ag. Paulina, M. en C. Karen, Biol. Fernando, Brenda y M. en C. Carmina por el entusiasmo compartido durante el proyecto.

Fernando Tapia, gracias por tu confianza y crecer a mi lado, gracias por dejarme ser tu tutora y sobre todo tu amiga.

Gracias al Dr. Josué Barrera Redondo por llevar este proyecto a las grandes ligas y por siempre compartir tu conocimiento con nosotros. Josué, gracias por ser la persona quién eres y compartir ese brillo conmigo.

Gracias al Dr. Luis Eguiarte por brindarme la oportunidad y la confianza de pertenecer a uno de los más selectos grupos de investigación con el que cuenta la UNAM. Le agradezco enormemente haberme recibido y aceptar ser mi maestro, gracias por enseñarme evolución y genética de poblaciones. **Mil gracias** por las enseñanzas académicas y personales.

A la Biol. Itzel Cabral. Itzel, has sido la pieza clave que ha permitido transformar mi vida cuando más lo he necesitado.

Agradezco a mi familia y amigos por su amor y apoyo incondicional. Virginia y Domingo, mis padres, quienes con su ejemplo y amor me han impulsado toda la vida a luchar por mis metas respetando en todo momento mis reglas y convicciones. Héctor, Silvia y Valentina gracias por compartir este momento, los quiero. Dulce Carolina, mi *soulmate*. Hermana eres mi confidente y mi apoyo incondicional, te amo y siempre te amare.

Alexis gracias por regresarle los colores a mi vida y por las dosis de locura, se ha vivido. Gracias por acompañarme en esta nueva etapa y mil gracias más por compartir conmigo *al perro* (Iggy).

Resumen.

La domesticación es un proceso evolutivo dirigido por la selección artificial que resulta en la divergencia entre un taxón silvestre y un taxón domesticado. Las especies domesticadas son excelentes modelos para estudiar el proceso evolutivo desde distintos enfoques como la morfología, la fisiología y la genética. Los principales estudios sobre el tema se han realizado mediante marcadores moleculares que se analizan bajo el marco conceptual de la genética de poblaciones y la filogeografía. Con la domesticación de especies la humanidad ha generado una amplia diversidad de recursos que son el sostén de la economía mundial y la base para la seguridad alimentaria. Dentro de estos recursos se encuentran las especies domesticadas del género *Cucurbita* (*C. argyrosperma*, *C. moschata*, *C. pepo*, *C. ficifolia* y *C. maxima*) generalmente conocidas como calabazas. Esta tesis presenta un análisis de la genética de poblaciones, la filogeografía y un aporte etnobotánico de estas especies. En el caso particular de *C. moschata* (calabaza de castilla), una especie ampliamente apreciada por su fruto, se analizaron marcadores genéticos de núcleo y mitocondria de variedades tradicionales de México para evaluar la variación y estructura genética. También, se estudiaron sus relaciones filogeográficas con el objeto de comprender mejor su historia evolutiva. En los sistemas agrícolas tradicionales es común encontrar distintas especies de calabazas, con el objetivo de determinar los niveles de flujo genético, también con marcadores moleculares, se evaluó la diferenciación genética entre especies mediante análisis bayesianos. Finalmente, se documentó la diversidad de calabazas en una región de Andina del Perú. A lo largo de esta tesis se discuten con mayor detalle los patrones de diversidad genética y las posibles causas evolutivas que han modelado las variedades actuales de *C. moschata* en México. También se analiza la importancia de la hibridación en la conservación de los acervos genéticos que representan estas especies y sus variedades locales en México. Finalmente, la documentación de diversidad de variedades locales encontradas en Perú permitió aportar evidencia que apoya la hipótesis del origen andino de una de las especies.

Abstract.

Domestication is an evolutionary process driven by artificial selection that results in divergence between a wild taxon and a domesticated taxon. Domesticated species are excellent model systems to study evolution using morphological, physiological and genetic approaches. The main studies of domestication have been carried out using molecular markers that are analyzed under the conceptual framework of population genetics and phylogeography. With the domestication of species, mankind has generated a wide diversity of resources that are the mainstay of the world economy and the basis for food security. Among these resources are the domesticated species of the genus *Cucurbita* (*C. argyrosperma*, *C. moschata*, *C. pepo*, *C. ficifolia* and *C. maxima*) generally known as pumpkins, gourds and squashes. This thesis presents an extensive analysis of population genetics, phylogeography and an ethnobotanical contribution of these species. In the particular case of *C. moschata* (*calabaza de castilla*), a species widely appreciated for its fruit, I analyzed nuclear and mitochondrial genetic markers of traditional varieties from Mexico to evaluate its genetic variation and structure and phylogeographic relationships to understand its evolutionary history. In traditional agricultural systems it is common to find different pumpkin species; to determine the levels of gene flow among the different species, I evaluated the genetic differentiation between species by Bayesian analysis. Finally, I documented the diversity of pumpkins in an Andean region of Peru. Throughout this thesis, I discuss in detail the patterns of genetic diversity and possible evolutionary causes that have shaped the current varieties of *C. moschata* in Mexico. Also, I analyzed the importance of hybridization in the conservation of gene pools representing different species and their landraces in Mexico. Finally, the diversity of landraces found in Peru support the hypothesis of the Andean origin of one of the species.

Índice.

Introducción	1
<i>Evolución bajo domesticación.....</i>	<i>1</i>
Marco teórico para el estudio de la evolución bajo domesticación.....	4
<i>Origen y diversificación de las especies domesticadas.....</i>	<i>5</i>
<i>Diversidad y estructura genética de las especies domesticadas.....</i>	<i>6</i>
<i>Diversidad genética y el manejo de las especies domesticadas ¿qué papel tienen las fuerzas evolutivas en su mantenimiento?.....</i>	<i>9</i>
<i>Recursos fitogenéticos.....</i>	<i>13</i>
Proyecto general y objetivos.....	16
<i>Presentación del proyecto.....</i>	<i>16</i>
<i>Objetivos generales.....</i>	<i>17</i>
Justificación.....	19
Sistema de estudio: Especies domesticadas del género <i>Cucurbita</i>.....	20
<i>Historia natural del género <i>Cucurbita</i>.....</i>	<i>20</i>
<i>Las especies domesticadas del género <i>Cucurbita</i>.....</i>	<i>22</i>
<i><i>Cucurbita moschata</i> Duchesne.....</i>	<i>24</i>
<i><i>Cucurbita argyrosperma</i> Huber spp. <i>argyrosperma</i>.....</i>	<i>28</i>
<i><i>Cucurbita pepo</i> L.....</i>	<i>29</i>
<i><i>Cucurbita ficifolia</i> Bouché.....</i>	<i>30</i>
<i><i>Cucurbita maxima</i> spp. <i>maxima</i> Duchesne.....</i>	<i>30</i>
Capítulo 1: GENÉTICA DE POBLACIONES Y FILOGEOGRAFÍA DE <i>Cucurbita moschata</i>, Duchesne ex Lam. Duchesne ex Poir. (CUCURBITACEAE) EN MÉXICO.....	32
<i>Artículo de investigación: Hernández-Rosales, H. S., G. Castellanos-Morales, G. Sánchez-de la Vega, E. Aguirre-Planter, S. Montes-Hernández, R. Lira-Saade, y L. E. Eguiarte. 2020. Phylogeographic and population genetic analyses of <i>Cucurbita moschata</i> reveal divergence of two mitochondrial lineages linked to an elevational gradient. <i>American Journal of Botany</i>. 107: 510-525.....</i>	<i>32</i>
Capítulo 2: GENÉTICA DE POBLACIONES COMPARADA Y ANÁLISIS DE FLUJO GENÉTICO DE ESPECIES DOMESTICADAS DEL GÉNERO <i>Cucurbita</i>.....	58
<i>Artículo en preparación: Juntas, pero no revueltas: Diferenciación genética y tasas de migración de tres especies domesticadas del género <i>Cucurbita</i> (CUCURBITACEAE) en México.....</i>	<i>58</i>
Capítulo 3: GENÓMICA DE POBLACIONES E IDENTIFICACIÓN DE LOCI BAJO SELECCIÓN DE <i>Cucurbita moschata</i>.....	98
<i>Genómica de poblaciones e identificación de loci con señales de selección de <i>Cucurbita moschata</i>.....</i>	<i>98</i>
Capítulo 4: ESTUDIO ETNOBOTÁNICO DE ESPECIES DOMESTICADAS DEL GÉNERO <i>Cucurbita</i> EN PERÚ.....	108
<i>Artículo de investigación: Barrera-Redondo, J., H. S. Hernández-Rosales, V. Cañedo-Torres, K. Aréstegui-Alegría, J. Torres-Guevara, F. Parra, I. Torres-García, y A. Casas. 2020. Landrace diversity and local selection criteria of domesticated squashes and gourds (<i>Cucurbita</i>) in the Central Andean mountain range of Peru: Tomayquichua, Huánuco. <i>Botanical Sciences</i>. 98: 101-116.....</i>	<i>108</i>
Discusión general.....	125
Conclusiones generales.....	132
<i>PERSPECTIVAS.....</i>	<i>134</i>
Anexo I: Fechamiento de tiempos de divergencia entre los haplotipos de ADN mitocondrial de <i>Cucurbita moschata</i>.....	135
Literatura citada.....	142

Introducción.

Evolución bajo domesticación.

Charles Darwin en *El Origen de las Especies* (1859) propuso su teoría de evolución, la cual considera dos premisas principales: los organismos descienden con modificaciones a partir de un ancestro común y que la causa principal de modificación es la Selección Natural que actúa en la variación que se hereda. Para que su teoría fuera aceptada por la comunidad científica, Darwin proporcionó abundantes evidencias entre las que destaca su investigación acerca de la variación en las especies domesticadas, trabajo al cual dedicó al primer capítulo de su libro (*La variación en estado doméstico*) del *Origen*. En términos generales, Darwin propone que los animales y plantas domesticados difieren más entre sí que los individuos de cualquier especie en estado natural. La gran diversidad de las especies domesticadas se debe a la selección artificial que ejerce el ser humano al favorecer (o no) la frecuencia de ciertos individuos. A partir del trabajo de Charles Darwin y otros naturalistas de la época, como Alphonse de Candolle, con su obra *Origine des plantes cultivées* (1883), se fundaron las bases teóricas actuales de la evolución bajo domesticación.

La evolución bajo domesticación es el proceso de divergencia evolutiva de un organismo que resulta en el establecimiento de una nueva especie (la especie domesticada) y su pariente silvestre estrechamente relacionado (Meyer y Purugganan, 2013). Para que esto suceda es necesaria la interacción de interdependencia entre dos organismos: (1) **El domesticador**, que es el encargado de dirigir el proceso mediante la selección artificial, a partir de la cual cuida y controla la reproducción de los individuos de la otra especie. (2) **El domesticado**, que es el organismo (planta, animal, hongo o bacteria) que posee características de interés/utilidad para el domesticador (Zeder et al., 2006). Aunque se reconoce al humano como la especie que más ha domesticado organismos de distintas especies, por ejemplo también se han identificado algunas especies de insectos (hormigas, termitas y escarabajos) que llevan a cabo el cuidado y reproducción controlada de consorcios microbianos para su beneficio (Mueller et al., 2005).

Generalmente se considera que la domesticación es el punto final de un continuo que (1) comienza con la selección de especies silvestres para su explotación, (2) continúa a través del cultivo/crianza de los organismos seleccionados y (3) termina con la fijación/formación de una nueva población, linaje o especie que difiere morfológica, fisiológica y genéticamente del

progenitor silvestre (Pickersgill, 2007). Así, las especies domesticadas se pueden definir como aquellas poblaciones/linajes de organismos que presentan un conjunto de modificaciones resultado de la adaptación dirigida por selección artificial para su cultivo/crianza y destinadas para el uso o consumo (Gepts, 2010).

Dado que la domesticación comprende distintas etapas, tenemos un gradiente en el que se pueden reconocer por un lado a linajes/poblaciones que presentan características o rasgos de una domesticación parcial o incipiente y que son capaces de sobrevivir en un estado silvestre (Eguiarte et al., 2018). Y por el otro, tenemos a las especies totalmente domesticadas que son aquellas que han divergido suficiente de su pariente silvestre y requieren/dependen totalmente de los cuidados del domesticador para reproducirse y seguir sobreviviendo (Zeder, 2015).

El proceso de domesticación no toma el mismo tiempo para todos los organismos y no todas las especies fueron domesticadas a la vez; al contrario, es un proceso que incluso está sucediendo es este momento para muchas especies. Originalmente, se creía que la divergencia entre el organismo domesticado y el pariente silvestre ocurría rápidamente (Hillman y Davies, 1990), sin embargo, ahora se sabe que puede ser un largo proceso abarcando de cientos a miles de años (Purugganan y Fuller, 2011).

El registro arqueológico sugiere una primera etapa de pre-domesticación en la cual los humanos cuidaban conscientemente formas silvestres que mostraban atributos o rasgos favorables (Meyer y Purugganan, 2013). Una vez que concluye la etapa de pre-domesticación y la diferenciación es claramente marcada entre el organismo domesticado y su pariente silvestre más cercano, el proceso de domesticación continúa con una segunda etapa, que consiste en el incremento *in situ* de organismos con los rasgos favorables. La tercera etapa corresponde a una fase de diversificación y mayor divergencia genética, fisiológica y morfológica. A lo largo del periodo de diversificación los organismos domesticados se propagan y adaptan a nuevos ambientes agroecológicos y responden a diferentes preferencias culturales (Meyer y Purugganan, 2013). Finalmente, con la intención de maximizar los atributos de las especies domesticadas (p. ej., rendimiento, facilidad de cultivo/crianza, calidad, etc.), la última etapa implica la reproducción consciente y deliberada de estos organismos. Esta etapa continúa haciéndose desde hace un poco más de 11,000 años (Meyer y Purugganan, 2013).

Independientemente del grado de domesticación, todas las especies domésticas presentan características que establecen las diferencian entre ellas y sus progenitores/parientes silvestres,

estos rasgos se conocen como *síndromes de domesticación* y son los atributos seleccionados durante la transformación inicial (Hammer, 1984; Olsen y Wendel, 2013). Entre los síndromes de domesticación más comunes en las plantas se encuentran: el incremento en el tamaño de las semillas y los frutos, pérdida de la dormancia, cambios de sabor de frutos, semillas y raíces, frutos suaves y ciclos de vida más cortos. Las etapas asociadas a la difusión de las especies domesticadas traen consigo una mayor diversidad de rasgos y solo por citar algunos podemos señalar: el incremento en el número de semillas, reducción de la toxicidad de las raíces, mayor tolerancia al estrés biótico y abiótico, floración sincrónica y mayor rendimiento (para una lista más completa, ver Meyer y Purugganan, 2013).

Desde el punto de vista genético, el proceso de domesticación implica selección y fijación de alelos relacionados con rasgos favorables. Por ejemplo, la fijación del alelo *Sh1* evitó que las semillas del maíz, arroz y sorgo se desprendieran fácilmente y su cosecha fuera más sencilla, (Stetter et al., 2017). Algunos alelos que confieren fenotipos domesticados son el resultado de mutaciones que ocurrieron en el pariente silvestre y/o surgieron durante el proceso de domesticación (Jaenicke-Després et al., 2003). Algunos otros alelos representan variación preexistente que, bajo el proceso de domesticación, se vieron favorecidos. La velocidad de fijación de estos alelos depende en gran medida de la intensidad de la selección artificial, sin dejar de lado la acción continua y progresiva del resto de las fuerzas evolutivas [endogamia, flujo genético, deriva génica (Casas y Parra, 2016)].

Dado que el proceso de evolución bajo domesticación ocurre “rápidamente” respecto a lo que sucede de forma natural, las especies domesticadas representa excelentes modelos para el estudio de distintos procesos evolutivos y el papel de las distintas fuerzas evolutivas. Como inicialmente se señaló al inicio de esta tesis, Darwin (1859) tomó como ejemplo los taxa domésticos para estudiar la adaptación y la especiación. Por sus cualidades, las especies domesticas brindan un amplio espectro y grado de divergencia que permite a los científicos evaluar, analizar y entender los cambios morfológicos, fisiológicos y genéticos de las especies bajo un enfoque evolutivo.

Marco teórico para el estudio de la evolución bajo domesticación.

El estudio de la domesticación es multidisciplinario y lo que sabemos actualmente de este proceso proviene de áreas de estudio como la botánica, la arqueología, la biogeografía, la etnobotánica e incluso de la lingüística (Larson et al., 2014). No obstante, el descubrimiento y desarrollo de la genética ha proporcionado poderosas herramientas para su análisis.

La domesticación es un proceso que deja huellas en el material genético de las especies y esas señales pueden ser identificadas a través del uso de marcadores moleculares. Estos marcadores genéticos permiten hacer inferencias sobre la historia de la domesticación y las bases moleculares de este proceso. También es posible evaluar el estado actual de la diversidad genética de las poblaciones, así como reconocer los genes involucrados en los síndromes de domesticación para su aplicación en el mejoramiento y desarrollo de nuevas variedades (Gaut et al., 2015; Guerra-García y Piñero, 2017; Kantar et al., 2017; Eguiarte et al., 2018).

Por un lado, la filogeografía es una rama del conocimiento que permite estudiar los componentes históricos, filogenéticos, y la distribución espacial de los linajes de genes bajo un enfoque de dos dimensiones [tiempo–espacio] (Vázquez-Domínguez et al., 2009). A partir de la filogeografía es posible abordar preguntas acerca del origen geográfico y el número de eventos de domesticación que han sufrido las especies. Por ejemplo: ¿Las especies se domesticaron una sola vez o sucedió más de una vez? ¿Las especies se domesticaron en un solo lugar o fueron domesticadas de manera independiente en distintas regiones?

Por su parte, la genética de poblaciones proporciona el marco conceptual para analizar la distribución espacial de las frecuencias alélicas con el objetivo de reconocer los niveles de variación y los patrones de diversidad de las poblaciones. El equilibrio de Hardy-Weinberg es un modelo que describe cómo se comporta la variación en ausencia de cualquier fuerza evolutiva (Eguiarte et al., 2013). Teniendo como punto de partida esta hipótesis nula, con los estadísticos básicos de la genética de poblaciones (p. ej., heterocigosis esperada [H_E] y observada [H_O], diversidad haplotípica [H_d], riqueza alélica [A]) podemos hacer inferencias acerca del estado actual de la diversidad. También es posible inferir la intensidad y el balance de las fuerzas evolutivas sobre las diferencias en las proporciones de alelos entre las poblaciones. Asimismo, permite identificar poblaciones con variación y estructura particular o que las distingue del resto de las poblaciones que conforman una especie.

Origen y diversificación de las especies domesticadas.

A partir de filogenias moleculares se pueden evaluar las relaciones de parentesco entre las especies domesticadas y sus parientes silvestres. En términos filogenéticos existen dos grandes grupos de métodos de reconstrucción: los que utilizan distancias genéticas (p. ej. *neighbor-joining*) y los métodos que se basan en caracteres (p. ej. máxima parsimonia, máxima verosimilitud e inferencia Bayesiana) (Swofford et al., 1992; Lemey et al., 2009). Con estos estudios se ha podido reconocer que hay familias taxonómicas con mayor número de especies que se han logrado domesticar en comparación con otras. Por ejemplo, en la familia Solanaceae se han documentado más de 30 especies domesticadas; esta familia incluye géneros como *Solanum* [tomate, papa] y *Capsicum* [chile] (Meyer y Purugganan, 2013). En la familia Cucurbitaceae hay varias especies domesticadas, tal es el caso del género *Cucurbita*, reconocido por incluir las calabazas domesticadas, con al menos seis eventos de domesticación independientes (Kistler et al., 2015; Kates et al., 2017a; Castellanos-Morales et al., 2018).

La filogenética, la biogeografía y la filogeografía proveen el marco conceptual para estudiar el origen y las distribuciones geográficas de los linajes, especialmente dentro y entre especies estrechamente relacionadas, permitiendo identificar procesos históricos y contemporáneos ligados a los cuellos de botella, expansiones y migraciones de las especies domesticadas (Avice, 2000). Los análisis filogeográficos son particularmente útiles en el estudio del proceso de domesticación, porque en general estas especies incluyen más de un linaje evolutivo asociado a procesos de diferenciación poblacional en una ventana de tiempo (Aguirre, 2015).

Tradicionalmente, en los estudios filogeográficos se utilizan regiones del ADN de mitocondria y cloroplasto que son ideales para la reconstrucción filogeográfica de linajes de genes (Avice, 2009). Estos marcadores generalmente son regiones selectivamente neutras, se heredan vía materna (salvo ciertas excepciones) y no recombinan (también salvo ciertas excepciones). Además, tienen un cierto grado de polimorfismo intraespecífico y en algunas ocasiones cuentan con una alta tasa de mutación (Hickerson et al., 2010; Aguirre-Dugua y González-Rodríguez, 2016).

Las inferencias filogeográficas tienen su base en la teoría de coalescencia (Avice, 2000). Este enfoque se basa en la reconstrucción, desde el presente hacia el pasado, de las relaciones históricas entre los genes de una muestra haciéndolos converger en el último ancestro en común

(Rosenberg y Nordborg, 2002). En el caso de las especies domesticadas se analizan los linajes de una o muchas poblaciones, representantes de distintas variedades y los parientes silvestres más cercanos (Aguirre-Dugua y González-Rodríguez, 2016).

Un método ampliamente usado para entender estas relaciones intraespecíficas es la construcción de *redes de haplotipos*. Un haplotipo es un grupo de alelos ligados que corresponde a una secuencia de ADN, la cual se puede distinguir de sus secuencias homólogas (Hedrick, 2005). Estas redes permiten describir relaciones no jerárquicas y reticuladas entre haplotipos, medir la frecuencia gráfica de los haplotipos e incluso se permiten inferir los haplotipos no muestreados o faltantes (Vázquez-Domínguez et al., 2009). Entre los algoritmos ampliamente utilizados para la reconstrucción de redes de haplotipos se encuentran el *median-joining network* (Bandelt et al., 1999), implementado en programas como NETWORK (Fluxus-engineering, 2015). Una red de haplotipos muestra la relación entre los haplotipos con base en el número de mutaciones, las conexiones entre ellos representan eventos de coalescencia y el tamaño del símbolo que representa a cada haplotipo generalmente indica la frecuencia dentro de la muestra. De acuerdo con Clement et al. (2000) los patrones generales que se pueden observar son: (i) los haplotipos más frecuentes o con conexiones múltiples tienen mayor probabilidad de ser los haplotipos ancestrales y (ii) un haplotipo ancestral comúnmente se encuentra al centro de la red y los periféricos son los más recientes.

El estudio de la domesticación a partir de este tipo de análisis ha permitido conocer, por ejemplo, que el origen del maíz (*Zea mays*) es monofilético, deriva de la especie silvestre *Z. mays* spp. *parviglumis* y su domesticación ocurrió hacia el oeste de México (Matsuoka et al., 2002), mientras tanto que, la cebada (*Hordeum vulgare*) tuvo dos centros independientes de origen, la cebada de Europa y África del Norte y la cebada de origen asiático (Saisho y Purugganan, 2007).

Diversidad y estructura genética de las especies domesticadas.

La genética de poblaciones permite caracterizar en términos cuantitativos la diversidad genética de las poblaciones y determinar su estructura considerando el espacio geográfico (Eguiarte, 1999). Se sirve de los marcadores moleculares para detectar polimorfismos (variantes/mutaciones) entre alelos de un gen (Grover y Sharma, 2016). Las isoenzimas (formas múltiples de enzimas) fueron de los primeros marcadores moleculares en implementarse en los estudios para cuantificar variación genética en poblaciones naturales [principalmente humanos y moscas

del género *Drosophila* (Hubby y Lewontin, 1966)]. Posteriormente se implementaron en estudios de poblaciones domesticadas. Por ejemplo, a partir de datos de isoenzimas se obtuvo una mejor resolución para diferenciar a las especies del género *Psidium* donde se ubica la guayaba domesticada (Albany et al., 1997). No obstante, el rápido desarrollo de marcadores que provienen de las moléculas de ADN (cloroplasto, mitocondria y núcleo) desplazó a las isoenzimas y hoy su uso en genética de poblaciones es casi obsoleto.

Los microsatélites o secuencias simples repetidas (SSRs) han sido mayormente usados en la genética de poblaciones para estimar los niveles de diversidad genética (Litt y Luty, 1989). Marcaron la base de mucho del conocimiento de la diversidad genética de las poblaciones naturales y domesticadas. Son excelente marcadores para identificar patrones de estructura genética y eventos de flujo genético entre poblaciones e introgresión. Además, permiten determinar patrones evolutivos de los procesos histórico del origen de especies y razas o variedades locales en el caso de especies domesticadas (Formisano et al., 2012; Castellanos-Morales et al., 2019). Estos marcadores siguen siendo la principal herramienta en los estudios de diversidad genética en virtud de las ventajas que ofrecen: (i) alto grado de polimorfismo; (ii) segregan de manera mendeliana y son codominantes y; (iii) son selectivamente neutros (Goldstein y Pollock, 1994).

Uno de los estudios más completos de la historia evolutiva y la diversificación de una especie domesticada de importancia a nivel global se abordó utilizando datos de microsatélites nucleares. En 2002, Matsuoka et al. analizaron la diversidad genética del maíz (*Z. mays* spp. *mays*). Ellos analizaron poco más de 260 individuos, algunos pertenecían a poblaciones de teocintle (*Z. mays* spp. *parviglumis* y spp. *mexicana*) y otros a poblaciones domesticadas recolectadas desde Canadá hasta el norte de Chile. Con la información de 99 loci de microsatélites evaluaron el flujo y estructura genética mediante métodos de agrupamiento, estudiaron la covarianza de las frecuencias alélicas con un análisis de componente principales y con la distancia de alelos compartido obtuvieron una reconstrucción filogenética. Entre los resultados más relevantes de su investigación, delimitaron que el maíz se domesticó en un solo evento a partir de la spp. *parviglumis* en México hace ~9,000 años. También observaron que variedades domesticadas forman grupos definidos de acuerdo con la región o variedad local a la cual pertenecen. Finalmente, encontraron señales de introgresión entre los taxones (teocintle-

maíz), con lo cual demostraron la importancia del flujo genético con la spp. *mexicana* como un factor que contribuye en la diversidad del maíz domesticado.

El rápido desarrollo de las tecnologías de secuenciación masiva y el ensamble de genomas completos permitió identificar variantes a una resolución de un **solo sitio en el ADN**. Con las técnicas de secuenciación de siguiente generación NGS por sus siglas en inglés (*next generation sequencing*), también llamadas técnicas de secuenciación paralela masiva, la genética de poblaciones y otras ramas del conocimiento se beneficiaron con la identificación de los polimorfismos de nucleótido único o SNPs (*single nucleotide polymorphisms*) a lo largo de todo el genoma (Escalante et al., 2014). Pese a que no son altamente polimórficos como los microsatélites, la ventaja es que existen miles a lo largo del genoma, tanto en regiones neutrales como codificantes (Eguiarte et al., 2013). Los costos actualmente son mucho más accesibles y en los últimos años el desarrollo de paqueterías bioinformáticas exclusivas para su análisis hace los estudios de genómica de poblaciones muchos más sencillos que hace una década.

Los SNPs derivados de métodos NGS han sido ampliamente utilizados en estrategias de desarrollo de marcadores con aplicaciones en la agricultura, como el caso del tomate [*Solanum lycopersicum* (Lin et al., 2014)]. En términos evolutivos, permiten evaluar la diversidad genética de las especies y realizar estudios de adaptación local como el del pariente silvestre del maíz [*Zea mays ssp. parviglumis* (Aguirre-Liguori et al., 2017; 2019)]. En términos de conservación, han permitido la genotipificación de colecciones de germoplasma y la identificación de accesiones duplicadas (De Albuquerque et al., 2019).

Si la historia evolutiva de las especies se ha podido deducir a través de unas decenas de marcadores ¿por qué utilizar SNPs? Estos polimorfismos se encuentran relativamente homogéneos a lo largo del genoma (Escalante et al., 2014), identificar estas variantes permite analizar el efecto y patrones relativos de los procesos evolutivos y la intensidad de cada una de sus fuerzas a nivel genómico (Eguiarte et al., 2013). Además de conocer la historia evolutiva de las especies con mayor precisión, estos marcadores ofrecen la oportunidad de detectar los genes relacionados a las adaptaciones locales ya que presentarán frecuencias alélicas contrastantes entre poblaciones (Eguiarte et al., 2013).

Hay tres principales enfoques para la identificación de variantes a lo largo del genoma que incluyen, entre otras, la secuenciación del genoma completo WGS (*whole genome sequencing*), la genotipificación por secuenciación GBS (*genotyping-by-sequencing*) y variantes análogas de

representación reducida, como RAD-seq (*restriction site-associated DNA sequencing*) y la secuenciación del exoma completo WEC (*whole exome capture sequencing*) y cada uno de ellos puede tener distintas ventajas y aplicaciones sobre el otro (Yao et al., 2020).

En términos muy generales, la metodología basada bajo un enfoque de GBS y otros métodos de representación reducida consiste en: 1) extracción de ADN; 2) fragmentación del ADN mediante enzimas de restricción o sonicación; 3) selección de fragmentos (100-200 pb); 4) adición de adaptadores a los fragmentos; 5) la construcción de librerías (si la tecnología de secuenciación lo requiere), y 6) secuenciación (Escalante et al., 2014). Para conocer las ventajas y desventajas de las plataformas más comunes ver Eguiarte et al. (2013).

La metodología GBS ofrece la ventaja de trabajar con genomas de todo tipo de tamaños debido (Yao et al., 2020). Por el contrario, además de los costos elevados de genotipificación, el poder de cómputo y la dependencia de los conocimientos bioinformáticos son necesarios para alcanzar los objetivos y metas específicas de cada estudio (Eguiarte et al., 2013).

Un ejemplo que muestra las ventajas de la metodología GBS es la historia evolutiva del maíz domesticado. Recientemente se ha dado un giro gracias al análisis de poco más de 32,000 SNPs. Moreno-Letelier et al. (2020) analizaron poblaciones de teocintle y razas criollas colectadas en México. Contrario a lo que habían propuesto Matsuoka et al. (2002) y Van Heerwaarden et al. (2011) -quienes sugieren que el maíz se domesticó en la Cuenca del Balsas en el estado de Guerrero (Matsuoka et al., 2002; Van Heerwaarden et al., 2011)-, Moreno-Letelier et al. (2020) proponen que el proceso pudo haber iniciado en las costas del estado de Jalisco. Sin importar el tipo de marcador molecular utilizado (microsatélites o SNPs), ambos estudios (Matsuoka et al., 2002; Moreno-Letelier et al., 2020) ponen de manifiesto la importancia del flujo genético y los múltiples eventos de introgresión en la compleja historia de la domesticación del maíz.

Diversidad genética y el manejo de las especies domesticadas ¿qué papel tienen las fuerzas evolutivas en su mantenimiento?

La asociación de organismos domesticados con los humanos tiene ~12,000 años (Meyer y Purugganan, 2013). Darwin, como anteriormente se mencionó, encontró en las especies domesticadas un sistema útil para explicar el origen de la variación (Eguiarte et al., 2018). Al igual que los criadores de plantas y animales de aquella época, Darwin poseía conocimiento

acerca de la **variación preexistente** en los organismos y como a partir de ella se desarrollaban nuevas razas a través de la selección artificial (Casas y Parra, 2016). Pese a que desconocía los trabajos de Mendel (1866), él también reconocía que la herencia de las características era parte fundamental para la transmisión de la variación a lo largo del tiempo.

Desde el inicio y en cada generación del proceso de domesticación la selección y fijación de rasgos favorables suelen limitarse a un **número reducido de individuos** de la especie progenitora (Doebley et al., 2006), y durante este período, se aumenta la frecuencia *in situ* de los organismos seleccionados. La población domesticada que se establece a partir de esos pocos individuos difiere considerablemente en su diversidad genética respecto a la especie silvestre porque uno de los efectos más drásticos de los cuellos botella es la reducción o pérdida de variación genética (Doebley, 1989). Contrariamente, los estudios recientes de la diversidad genética de las especies domesticadas apuntan que muchas de las plantas domesticadas en realidad presentan poca pérdida de diversidad genética. En los siguientes párrafos se explicarán algunos procesos que contrarrestan la pérdida de variación causada por los cuellos de botella asociados a los eventos de domesticación.

El inicio de la domesticación se establece a partir de la selección de distintos individuos de una población progenitora que poseen atributos de interés o rasgos favorables. El resultado del proceso selectivo son poblaciones con caracteres que las distinguen, pero con diferencias en las frecuencias de los alelos (Casas et al., 1997). En otras palabras, el cuello de botella asociado a esta primera etapa del proceso se ve contrarrestado porque: (1) aunque se selecciona un atributo, los individuos seleccionados con ese rasgo difieren en las frecuencias de sus alelos para distintos genes; (2) no debe descartarse la posibilidad de que los individuos provengan de más de una población progenitora o linajes distintos y (3) es posible que se hayan domesticado más de una vez en el tiempo. Una vez que se han fundado poblaciones de individuos seleccionados por sus rasgos de interés y comienza la etapa de diversificación, las distintas presiones de selección entran al juego. Las presiones de selección han generado dentro de las especies domesticadas una amplia diversidad de variedades. Por ejemplo, para el arroz (*Oryza sativa*) se han reconocido más de 50,000 variedades (Sinha y Mishra, 2012).

Entonces, ¿cuáles son las presiones de selección que conllevan a la diversificación de especies domesticadas? La fuerza principal es la **selección artificial** que se puede dividir en selección consciente y selección inconsciente. La **selección consciente** o intencional es aquella

que se aplica deliberadamente para incrementar en frecuencia atributos de interés. En otras palabras, se preserva lo que se considera que tiene mayor valor de cada organismo y ese mismo criterio se aplica generación tras generación (Zohary, 2004).

Por otra parte, la **selección inconsciente** o automática es aquella que se lleva a cabo involuntariamente cuando los organismos domesticados son movidos de su hábitat de origen para ser colocados frecuentemente en lugares muy distintos. La difusión a nuevos ambientes automáticamente promueve cambios drásticos en las presiones de selección. Los organismos domesticados pierden y desarrollan nuevas adaptaciones para sobrevivir a las nuevas condiciones ecológicas (Zohary, 2004).

La selección artificial no anula las presiones de la **selección natural** sobre los organismos domesticados y su diversidad (Guillén et al., 2013). En ambos procesos de selección artificial hay consciencia e intención, solo que los efectos de la selección consciente son más rápidos, sistemáticos y explícitos que los de la selección inconsciente (Casas y Parra, 2016). Concisamente, la variación que heredan los organismos seleccionados de la población progenitora no se pierde drásticamente ni en el inicio del proceso de domesticación ni en las etapas posteriores. La variación de las especies domesticadas son expresiones de divergencias (diferencias morfológicas o genéticas) generadas intencionalmente a través de la selección artificial.

Además de las presiones de selección, la deriva génica, la endogamia, la mutación y el flujo genético son importantes en el establecimiento de nuevas variedades porque regulan los niveles de variación y moldean los patrones de diversidad genética. La **deriva genética** es una fuerza evolutiva que afecta a todos los loci, pero en el proceso de domesticación influye fuertemente en la fijación o eliminación de alelos. En las poblaciones naturales es un proceso aleatorio y el tiempo que tardan en fijarse y/o perderse un alelo dependerá del tamaño efectivo de la población. (Hedrick, 2005). Sin embargo, en el proceso de domesticación la deriva es particularmente importante, porque el manejo selectivo de porciones de diversidad genética determina los efectos fundadores para establecer las divergencias evolutivas en pocas generaciones (Casas y Parra, 2016). Se fijan y pierden alelos en cada generación, pero la diversidad genética, al menos en plantas, se mantiene debido a que los tamaños efectivos son grandes, debido a la difusión y establecimiento de nuevas poblaciones (Hernández-Rosales et al., 2020).

La **endogamia** es una fuerza que opera de manera natural, sin embargo, en la domesticación de especies es generalmente un proceso guiado y favorecido intencionalmente. Mediante la reproducción controlada de individuos, se asegura que la progenie herede los atributos seleccionados de tal manera que las cruza endógamas permiten el establecimiento en pocas generaciones de nuevas variedades. Sin embargo, esta fuerza evolutiva disminuye la diversidad genética de las poblaciones, porque la fijación de los rasgos favorables requiere un control de la reproducción y generalmente los individuos manejados resultan estar estrechamente emparentados. El caso más extremo de una cruce endógama se da en las plantas con la autofecundación o con clonalidad. Así como la deriva genética, la endogamia tiene efectos sobre toda la variación del genoma y por ello no se descarta la posibilidad de resultados no deseables, como la expresión de genotipos no favorables y la disminución de la variabilidad genética. No obstante, la endogamia es una fuerza evolutiva considerablemente importante en el proceso de evolución bajo domesticación, es la expresión del manejo ancestral y actual que continúa hoy en día para el desarrollo y mejoramiento de plantas y animales (Hawkes, 1983; Casas y Parra, 2016).

El **flujo genético** es la principal fuerza evolutiva que compensa la pérdida de variación causada por la deriva genética y la endogamia. La introgresión con los parientes silvestres y/o especies hermanas permite intercambio y aumento de la diversidad genética. Entre las prácticas de manejo más comunes en plantas se encuentra el movimiento e intercambio de germoplasma, que promueve que nueva diversidad se introduzca entre las poblaciones domesticadas. Además, se ha propuesto que el flujo entre especies se ve facilitado e intensificado por el manejo humano (Aguirre-Dugua et al., 2012; Sánchez-de La Vega et al., 2018). Estudios recientes han confirmado que el flujo genético entre la especie progenitora silvestre y el taxón domesticado ha sucedido desde el inicio de la domesticación y continúa actualmente, tal es el caso del maíz y su pariente silvestre el teocinte (Moreno-Letelier et al., 2020).

Finalmente, ninguna especie ni silvestre ni domesticada está exenta de la **mutación**. Esta es la fuerza evolutiva que genera diversidad genética *de novo* (Futuyma, 2013). Las mutaciones son las señales en el material genético que permiten reconocer en las especies domesticadas los procesos asociados a adaptaciones geográficas y reconstruir el origen. Las mutaciones pueden cambiar las funciones de los genes en la especie progenitora, teniendo como resultado fenotipos o atributos favorables que después por selección artificial pueden incrementar sus frecuencias hasta fijarse en las poblaciones domesticadas (Meyer et al., 2012). El estudio de la historia evolutiva de

las especies y su diversificación depende en gran medida de la identificación de estas mutaciones.

En el proceso de evolución bajo domesticación, las presiones de selección y su interacción con el resto de las fuerzas evolutivas da como resultado la amplia diversidad de **variedades criollas o razas locales** (generalmente llamadas *landraces* en inglés) que actualmente podemos observar (Meyer et al., 2012).

Recursos fitogenéticos.

Los recursos genéticos son parte de la base biológica de la seguridad alimentaria y, directa o indirectamente, sostienen los medios de subsistencia de todos los habitantes de la Tierra (FAO, 2020). Poseen las unidades funcionales de herencia (genes) con valor actual y potencial para la economía, la ciencia y las necesidades sociales de la humanidad (Allem, 2000).

En particular, los organismos de origen vegetal que tienen potencial o valor para la alimentación, agricultura y silvicultura son reconocidos como **recursos fitogenéticos** o por sus siglas en inglés PGR [*plant genetic resources* (Allem, 2000)]. Estos incluyen a las poblaciones y acervos genéticos de variedades criollas o razas locales, parientes y especies silvestres que están disponibles para el mejoramiento de especies de plantas domesticadas (o en proceso). Son el mayor **recurso natural**, del cual el humano ha dependido y dependerá para su sobrevivencia (Salgotra y Zargar, 2020).

El valor, potencial e importancia de los recursos fitogenéticos radica en su **diversidad genética**, la cual se aprovecha para múltiples propósitos en la atención de necesidades sociales, tecnológicas y económicas (Casas y Parra, 2007). Bajo el panorama de cambio climático que estamos viviendo como humanidad, la variabilidad de estos recursos es y podrá ser usada para el incremento en la resiliencia y productividad de los sistemas de producción agrícola para la seguridad alimentaria futura (Salgotra y Zargar, 2020). Su variabilidad genética se resguarda en gran medida en las **variedades locales o razas criollas**, así como en los **parientes silvestres** de los cultivos (Casas y Parra, 2007).

Los parientes silvestres tienen una historia evolutiva más antigua y por ello poseen reservorios de diversidad genética más extensos; esta variación les confiere mayor potencial evolutivo (Casas y Parra, 2007; Ochoa-Zavala et al., 2016) y posiblemente posean diferentes caracteres deseables en las plantas cultivadas, por ejemplo, genes de resistencia a estrés biótico y

abiótico (Salgotra y Zargar, 2020). Por su parte, las variedades criollas son aquellas seleccionadas por agricultores locales, quienes las cultivan y manejan de manera tradicional de acuerdo a su cultura campesina local (Drozdowska y Hermanowski, 2016). Estas variedades tradicionales son importantes porque poseen variación genética que puede ser utilizada como materia prima, incluso algunas pueden tener niveles de diversidad genética similar a las de sus parientes silvestres (Bunting y Pickersgill, 1966; Castellanos-Morales et al., 2019). Estas variedades locales generalmente no son consideradas para su cultivo tecnificado debido a su bajo rendimiento. El valor de las variedades locales reside en que muchas veces presentan mayor tolerancia a factores de estrés hídrico, sequías, heladas, suelos salinos y altas temperaturas y son estos rasgos los que se pueden utilizar para el mejoramiento de cultivos de élite (Salgotra y Zargar, 2020).

Por otra parte, la **pérdida de variación genética** de los recursos fitogenéticos es o debería ser una de las mayores preocupaciones de la humanidad. Por un lado, esta pérdida puede darse por factores naturales, principalmente por los cuellos de botella que se producen generación tras generación debido a la selección de unos pocos individuos. Cuando la disminución de la diversidad genética es debido a la pérdida de los parientes silvestres, o de las especies silvestres relacionadas o de variedades locales, se le denomina **erosión genética** (Velásquez-Milla et al., 2011). La erosión genética es un proceso que involucra el remplazo de la agrobiodiversidad tradicional por cultivares modernos de agricultura intensiva (Van De Wouw et al., 2010). Una de las consecuencias de la Revolución Verde (1960–1990) fue el desarrollo de cultivares mejorados con mayor rendimiento y adaptabilidad a la cosecha tecnificada (Mastretta et al., 2018). Como resultado de este proceso, muchas variedades locales fueron y seguramente seguirán siendo desplazadas (Ceccarelli, 2009). Entender las consecuencias de la erosión genética es difícil debido a la limitada o inexistente documentación de los niveles de variación genética de la agrobiodiversidad. Entonces, ¿cómo saber cuánto variación se está perdiendo si no sabemos cuánta variación existe?

La información respecto a la diversidad genética de razas criollas es poca, inclusive en regiones consideradas como centros de origen y diversificación (como es el caso de México; Piñero et al., 2009). Nuestro país es considerado agrodiverso, ya que cuenta con alrededor de 130 plantas nativas domesticadas y muchas de ellas tiene una gran importancia económica en todo el mundo (Acevedo et al., 2009; Mastretta-Yanes et al., 2018). Destacan aquellas asociadas a los

sistemas de agricultura tradicional denominados *milpa*. Estos agroecosistemas generalmente incluye maíz, diferentes especies de frijol, calabazas y chiles (Acevedo et al., 2009). Asimismo, cabe destacar que las poblaciones de parientes silvestres de estos cultivos hoy en día coexisten muy cercanamente. Es preciso señalar que, si bien existen varios trabajos relacionados a la identificación de variedades nativas del maíz y su diversidad genética, así como estudios de su domesticación, genómica de poblaciones y hasta estudios de transcriptomas, la información más básica para el resto de las especies de origen Mesoamericano es limitado (p. ej., el registro y distribución de variedades nativas o razas criollas) (Mastretta-Yanes et al., 2019).

Por estos motivos considero que tanto *i)* el registro de las razas criollas o variedades locales de las especies domesticadas y parientes silvestres; *ii)* la documentación de su variación genética y distribución espacial; *iii)* la caracterización de los patrones de estructura genética poblacional y *iv)* la identificación de regiones de alta diversidad genética, son críticos para el entendimiento de la historia evolutiva de estos recursos genéticos bajo un marco conceptual de genética de poblaciones y filogeografía. A su vez, esta información contribuirá a desarrollar planes de conservación y manejo sostenible adecuados para salvaguardar la variabilidad genética a mediano y largo plazo.

Proyecto general y objetivos.

Presentación del proyecto.

Esta tesis forma parte del proyecto “Diversidad genética de las especies de *Cucurbita* en México. Fase I” y “Fase II. Genómica evolutiva y de poblaciones, recursos genéticos y domesticación” financiado por la Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (núm. proyectos CONABIO KE004 y PE001, respectivamente).

México es centro de origen y diversificación del género *Cucurbita* que incluye especies cultivadas de importancia económica y cultural, entre las que destacan *C. argyrosperma* spp. *argyrosperma* (calabaza pipiana), *C. pepo* spp. *pepo* (calabacín, zucchini), y *C. moschata* (calabaza de castilla). Una revisión de estudios en 2016 reveló que para México no existía un verdadero registro de la diversidad genética de estas especies (incluyendo variedades locales y razas criollas, así como las especies silvestres nativas). Por lo anterior, el objetivo general del proyecto buscaba realizar estudios para analizar la diversidad genética de las distintas especies de calabazas domesticadas y silvestres (*Cucurbita* spp.) que se cultivan y crecen en México. Es así que este proyecto ha generado la documentación más completa de la variación genética y distribución espacial de estas especies en México (Sánchez-de la Vega et al., 2018; Castellanos-Morales et al., 2019; Hernández-Rosales et al., 2020).

Esta tesis se enfocó en analizar la variación y patrones de estructura genética de *Cucurbita moschata* bajo el marco conceptual de la genética de poblaciones y la filogeografía. Originalmente, este trabajo estaba planteado para analizar solo accesiones resguardadas en el banco de germoplasma del Campo Experimental Bajío, Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias (INIFAP). Sin embargo, tras una serie de pruebas de germinación y debido a que muchas de las accesiones ya habían perdido variabilidad, se tomó la decisión de recorrer su distribución potencial en México con base en los puntos de colectas de distintas bases de datos públicas y así recolectar germoplasma viable.

Es así que se plantearon los siguientes objetivos generales de esta tesis y que corresponden a los capítulos que la componen.

Objetivos generales.

1. Utilizar la información de la base de datos del banco de germoplasma del Campo Experimental Bajío, Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias (INIFAP) y la base de datos de CONABIO para mapear la distribución potencial de las colecciones y detectar sitios en los que se requiera hacer trabajo de recolección. **Capítulo 1.**
2. Determinar y analizar la variación genética de las variedades locales y razas criollas de *Cucurbita moschata* (Duchesne ex Lam. Duchesne ex Poir) distribuidas en México bajo un enfoque de genética de poblaciones. **Capítulo 1.**
3. Mapear la distribución de la diversidad genética y analizar los recursos genéticos de la especie para inferir su historia evolutiva en México y las fuerzas que han modelado su filogeografía. **Capítulo 1.**
4. Determinar los tiempos de divergencia de esta especie y analizarlos bajo un marco conceptual de la evolución bajo domesticación. **Anexo 1.**
5. Comparar la variación genética y determinar los niveles de flujo genético de las tres especies de calabaza con mayor importancia económica y cultural en México (*C. moschata*, *C. argyrosperma* spp. *argyrosperma*, el pariente silvestre *C. argyrosperma* spp. *sororia* y *C. pepo* spp. *pepo*). **Capítulo 2.**
6. Determinar y analizar los niveles de variación genómica de las variedades locales y razas criollas de *C. moschata* mediante marcadores de polimorfismos de nucleótido único o SNPs e identificar genes candidatos con potencial para el mejoramiento de cultivos de élite. **Capítulo 3.**

Adicionalmente, como parte de una estancia de investigación en Perú y ante la preocupación de la falta de información de los recursos fitogenéticos del género *Cucurbita* a

nivel global, se planteó un proyecto de investigación para seguir contribuyendo al conocimiento de las calabazas domesticadas. En Perú y el resto de los países de Sudamérica, *C. maxima* es generalmente conocida como zapallo y *C. moschata* es comúnmente conocida en algunas regiones de Perú como shupe o loche. Ambas especies son cultivos de gran importancia económica, mientras que, como en otras regiones del mundo, *C. ficifolia*, conocida en ese país como calabaza, presenta un impacto menor (Lira et al., 2016; Paris, 2016). Perú al igual que México, carece de información detallada de la variación genética y la distribución espacial de sus recursos fitogenéticos. Fue así que en colaboración entre la Universidad Nacional Agraria La Molina, Lima, Perú y la Universidad Nacional Autónoma de México se llevó a cabo un estudio etnobotánico con el objetivo de enriquecer el conocimiento sobre la diversidad biológica y el manejo de las calabazas y zapallos en la región de Tomayquichua, Huánuco. Para este proyecto se plantearon las siguientes preguntas de investigación:

1. ¿Cuáles son las variedades locales de *Cucurbita* que reconocen, promueven y conservan los agricultores en el Distrito de Tomayquichua? **Capítulo 4.**
2. ¿Qué mecanismos promueven la diversidad de estas variedades? **Capítulo 4.**

Justificación

Esta tesis de investigación se justifica partiendo de las siguientes premisas:

- I. La importancia de los recursos fitogenéticos que representa *Cucurbita moschata* por su versatilidad como alimento y su extensa distribución en distintas condiciones ecológicas.
- II. La poca información que existe para esta especie del acervo genético que representan sus numerosas variedades locales y razas criollas que se cultivan en México.
- III. Considerando el posible origen Mesoamericano de esta especie, es indispensable el conocimiento de la historia evolutiva de las variedades mexicanas bajo el marco conceptual de la genética de poblaciones y filogeografía.
- IV. Debido a la capacidad de hibridación entre taxa domesticados del género *Cucurbita* es necesario evaluar el flujo genético entre ellas (*C. moschata*, *C. argyrosperma* spp. *argyrosperma* y *C. pepo* spp. *pepo*) para así determinar el impacto dentro de los acervos genéticos.
- V. Hay poca información acerca del efecto que tienen los criterios de selección por parte de los agricultores sobre la domesticación y mantenimiento de las variedades de las especies cultivadas del género *Cucurbita* en Latinoamérica.

Sistema de estudio: Especies domesticadas del género *Cucurbita*.

Historia natural del género *Cucurbita*.

El género *Cucurbita* (CUCURBITACEAE) es reconocido generalmente por sus especies cultivadas, comúnmente llamadas en español calabazas o zapallos. El género es nativo de América y se considera de origen relativamente reciente, el tiempo de divergencia con su género hermano *Peponopsis* se ha calculado hace 16 ± 7 millones de años (Castellanos-Morales et al., 2018). Se reconocen entre 12–15 especies de calabazas y alrededor de 20 taxa de los cuales cinco son domesticados (Lira et al., 1995; Paris 2016). Antes del Holoceno (hace unos 11,000 años) *Cucurbita* ya presentaba un amplia distribución, así que su dispersión y adaptación a nuevos ambientes parece haber sido facilitada por la megafauna actualmente extinta (Kistler et al., 2015).

Tradicionalmente las especies del género han sido divididas en dos grandes grupos siguiendo sus características ecológicas, fisiológicas y su distribución. El primer grupo corresponde al de las especies **xerofíticas** (climas áridos), que se caracterizan por poseer raíces tuberiformes y forma de vida perenne. El segundo grupo es el de las especies **mesofíticas** (climas húmedos), que poseen raíces fibrosas, y con la excepción de *C. ficifolia* (que es perenne), conservan una forma de vida anual (Paredes, 2016). Dentro del grupo de taxa **mesofíticos** se incluyen las cinco especies cultivadas del género: *C. argyrosperma* spp. *argyrosperma*, *C. ficifolia*, *C. moschata*, *C. maxima* spp. *maxima* y *C. pepo* spp. *pepo* (Figura 1).

A pesar de contar con pocas especies, *Cucurbita* es un género evolutivamente exitoso y un sistema de estudio interesante. Cuenta con al menos seis eventos independientes de domesticación (dos eventos para *C. pepo* spp. *pepo*) (Kates et al., 2017; Castellanos-Morales et al., 2018; Castellanos-Morales et al., 2019). En el Nuevo Mundo, los únicos géneros que pueden ser comparables con *Cucurbita* por el número de especies domesticadas son *Capsicum* (con 3-5 especies) y *Solanum* (con 7 especies) (Nee, 1990). Además, dentro de las angiospermas *Cucurbita* es un género con gran variedad morfológica y esto se debe mayormente a que las especies cultivadas contienen una extensa diversidad a escala local, regional y mundial, además de poseer una extensa riqueza de variedades locales y razas criollas (Bisognin, 2002; Sanjur et al., 2002).

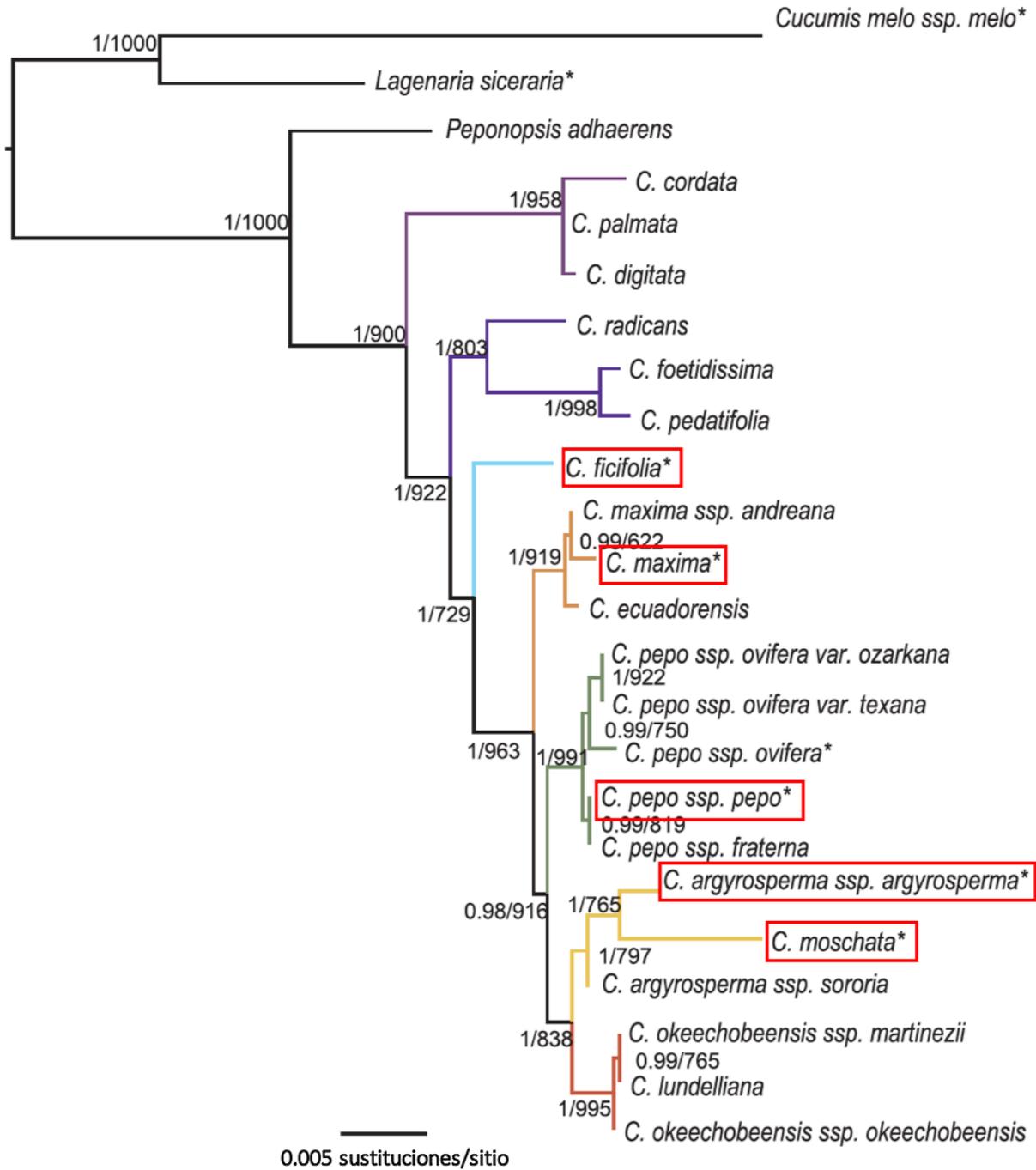


Figura 1. Filogenia del género *Cucurbita* (Castellanos-Morales et al., 2018). Las cinco especies domesticadas de importancia alimenticia se resaltan en rojo.

Casi todas las especies del género se distribuyen y se cultivan en México y por esta razón se le considera centro de diversidad. De igual forma, México es considerado un importante centro de origen, porque de acuerdo con la evidencia molecular y arqueobotánica, *C. moschata*, *C. argyrosperma* spp. *argyrosperma* y *C. pepo* spp. *pepo* parecen haber sido domesticadas en esta región de América (Smith, 1997b; Bisognin, 2002; Sanjur et al., 2002).

En cuanto a la biología general, todos los taxa del género son plantas monoicas, rastreras o trepadoras. Poseen flores gamopétalas con corolas tubular-campanuladas. Las flores masculinas (estaminadas) y femeninas (pistiladas) pueden presentar una coloración que varía entre amarillo-pálido a tonos más amarillos y naranjas brillantes (Lira, et al., 1995). Las flores normalmente son polinizadas por abejas solitarias del género *Peponapis* y *Xenoglossa* (Hurd et al., 1971), lo que propicia el flujo genético y la hibridación entre las especies silvestres y cultivadas (Montes-Hernández y Eguiarte, 2002). Los frutos son del tipo pepo, en las especies silvestres son relativamente uniformes en forma, superficie, tamaño y coloración. Por el contrario, los frutos de las especies cultivadas presentan gran variedad en dichos caracteres (Lira et al., 1995; Eguiarte et al., 2018). Los ensambles de genomas más recientes confirman que todas las especies de *Cucurbita* tienen 20 pares de cromosomas (Sun et al., 2017; Montero-Pau et al., 2018; Barrera-Redondo et al., 2019).

Las especies domesticadas del género *Cucurbita*.

Entre las especies domesticadas de la familia *Cucurbitaceae*, las calabazas domesticadas son los cultivos más destacados e importantes, y se pueden encontrar como frutas y verduras en casi todas partes del mundo (Paris, 2016; Chomicki et al., 2019). De acuerdo con su relevancia económica y producción, de las cinco especies domesticadas, *C. pepo* spp. *pepo* (calabacín) es la especie más cultivada a nivel global (FAO, 2020). El segundo lugar le corresponde a *C. maxima* spp. *maxima* (kabocha) y en tercer lugar se encuentra *C. moschata* (calabaza de castilla). Finalmente, la producción de las especies *C. argyrosperma* spp. *argyrosperma* (calabaza pipiana) y *C. ficifolia* (chilacayote) más bien responden a las demandas del consumo local.

Las calabazas tienen gran importancia alimenticia debido su valioso aporte nutricional; por ejemplo, las semillas poseen alto contenido de proteínas y aceites (Jarret et al., 2013). Son cultivadas en todo el mundo, principalmente para el consumo de sus frutos (maduros e inmaduros), aunque particularmente en América Latina el consumo de las flores, semillas y las

partes tiernas de los tallos es común (OECD, 2012). En la industria son consideradas como materia prima para la elaboración de detergentes, debido a su contenido de saponinas (Nee, 1990) y gracias a sus compuestos, se han registrado múltiples usos en la medicina tradicional (Lira, et al., 1995).

Las calabazas domesticadas no solo son fuentes de alimentos relevantes y versátiles, también aportan beneficios ecológicos dentro de los agroecosistemas tradicionales en los que se cultivan. En México, principalmente en la región que corresponde a Mesoamérica, las variedades locales y razas criollas de calabazas se cultivan en las milpas. En estos sistemas, es habitual encontrar en asociación al maíz, distintas variedades o especies de chile (*Capsicum* sp.), frijol (*Phaseolus* sp.) y de una o más especies de calabazas (*Cucurbita* spp.) (Hernández Xolocotzi, 1988). Dentro del agroecosistema milpa, las calabazas previenen el desarrollo de malezas, ayudan a mitigar la erosión del suelo y gracias a sus grandes hojas forman una densa capa sobre la superficie, que evita la pérdida de humedad del suelo (Altieri, 1995; Gliessman, 1990; Lozada-Aranda et al. 2017).

Los cultivos de élite de calabazas que se producen bajo un esquema de agricultura intensiva tienen gran importancia económica a nivel global. Tan solo en 2018 se registró una producción cercana a los 28 millones de toneladas en el mundo. La región con mayor producción de calabazas fue Asia (62%) y entre los países, México es el quinto mayor productor mundial de este cultivo con más de 200,000 toneladas (FAO, 2020). Cabe mencionar que estos informes no consideran al resto de las especies cultivadas, ni las variedades locales o razas criollas. Estas últimas representan un importante recurso destinado en gran medida para el autoconsumo familiar, mientras que el excedente puede ser utilizado para alimentar animales de granja o para la venta local (Barrera-Redondo et al., 2020).

Asimismo, los estudios de los recursos genéticos de las calabazas se han enfocados en la especie de mayor producción (*C. pepo* spp. *pepo*) y sus variedades comerciales y mejoradas. Hay mucha información de la diversidad de esta especie que salvaguardan algunos bancos de germoplasma, especialmente de accesiones variedades comerciales y mejoradas [p. ej. Paris et al., 2003; Formisano et al., 2012; Gong et al., 2012)]. Incluso, hasta hace pocos años el único genoma disponible para el género *Cucurbita* pertenecía a esta especie (Montero-Pau et al., 2018).

El valor de los recursos genéticos reside en las variedades locales, razas criollas y parientes silvestres, y las calabazas domesticadas que son importantes por su versatilidad como

alimento, su rol ecológico en los agroecosistemas y sobre todo por su innumerable riqueza de formas, variedades criollas y razas locales [p. ej. Lira et al., 1995; Sanjur et al., 2002; Sánchez-de La Vega et al., 2018; Castellanos-Morales et al., 2019; Barrera-Redondo et al., 2020; Hernández-Rosales et al., 2020)].

De las especies domesticadas de *Cucurbita*, poco estudiadas, pese a ser la segunda más diversa morfológicamente del género, es *C. moschata* (Goldman, 2004; Ferriol y Pico, 2008; OECD, 2012; Paris, 2016). Esta especie fue ampliamente apreciada en las culturas mesoamericanas, donde hasta la fecha todas sus partes son consumidas con propósitos alimenticios (Lira et al., 1995). Había sido motivo de interés en algunos trabajos debido a la estrecha relación con el Grupo *Argyrosperma* (Montes-Hernandez S.; Luis E. Eguiarte, 2002) pero su estudio en México era limitado.

***Cucurbita moschata* Duchesne.**

Cucurbita moschata (Duchesne ex Lam.) Duchesne ex Poir, de acuerdo con su ficha taxonómica, fue descrita por primera vez en 1818. Whitaker en 1947 la reconoce como una especie de calabaza que se domesticó en América Latina. Para 1966, Filov ya había reconocido más de 20 variedades de esta especie. En México esta especie es comúnmente conocida como calabaza de Castilla o calabaza criolla. Aunque en distintas regiones (también de México) recibe innumerables nombres, por mencionar algunos: calabaza de pellejo (Michoacán), tamalayota o támara que son nombres de origen Náhuatl (Oaxaca), *k'uum* y numerosas variaciones de origen Maya (Yucatán). En el resto de América Latina, se le conoce como zapallo (Colombia), auyama (Venezuela), loche o yuwi (Perú), en Estados Unidos y países de habla inglesa es llamada *seminole pumpkin* o *butternut squash* (Lira et al., 1995; Eguiarte et al., 2018; Barrera-Redondo et al., 2020).

Aunque su posible área de origen y domesticación aun es incierta, existen dos importantes evidencias que permiten teorizar acerca de su origen Mesoamericano, principalmente en la zona que corresponde a México:

- I. En primer lugar, tenemos que si bien se han encontrado distintos restos arqueobotánicos para esta especie a lo largo de América Latina (Whitaker y Bird, 1949; Dillehay, 2007; OECD, 2012), los más antiguos conocidos hasta la fecha datan de hace ~7,000 años (y

otros de 4,000 años) que se encontraron en México dentro de unas cuevas conocidas como “Cuevas de Ocampo” en Tamaulipas (Whitaker et al., 1957).

II. Actualmente no se conoce al pariente silvestre de *C. moschata* y es posible que aún no se haya colectado o que esté extinto (Nee, 1990; Sanjur et al., 2002; OECD, 2012; Kates et al., 2017; Castellanos-Morales et al., 2018; Eguiarte et al., 2018), pero esta especie tiene una estrecha relación taxonómica y genética con *C. argyrosperma* (Lira et al., 1995, Kates et al., 2017; Castellanos-Morales et al., 2018). Adicionalmente, se sabe que diversificó de *C. argyrosperma* hace 1.06 (0.29–1.9) millones de años (Castellanos-Morales et al., 2018). y que es posible a una eficiente producción de híbridos fértiles entre *C. moschata* y el taxa silvestre *C. argyrosperma* spp. *sororia* (Merrick, 1990). Esta última en la actualidad se distribuye de manera natural desde México hasta Centroamérica (Balvino-Olvera et al., 2017).

III. La versatilidad como alimento y su amplia adaptación a diversas condiciones ecológicas hace de *C. moschata* la tercera principal especie de cultivo dentro de las calabazas domesticadas y la segunda más diversa del género (Ferriol and Pico, 2008; Gong et al., 2013; Paris, 2016; Chomicki et al., 2019).

El uso más importante que se le da a *C. moschata* es el alimenticio, ya que los frutos maduros e inmaduros, las semillas, las flores y las puntas de las guías (parte tiernas de los tallos) son empleadas con este fin. Los frutos inmaduros son consumidos como verdura, mientras que los frutos maduros de *C. moschata* son particularmente apreciados por su pulpa, que se caracteriza por ser espesa, no fibrosa, de aroma dulce y sabores suaves, lo que permite sea fácilmente consumida en platillos dulces y salados (Lira et al., 1995; Paris, 2016). La intensidad del color es un importante componente de los frutos y la pulpa, por ello el alto contenido de carotenoides (fuente de antioxidantes) que es un elemento de gran valor que ofrece este cultivo (Goodwin, 1980; Azevedo-Meleiro y Rodríguez-Amaya, 2007). Otra particularidad de *C. moschata* es que las cáscaras de los frutos maduros no son lignificadas, lo que facilita su cosecha. Estas características son evidencias de que esta especie se ha cultivado durante mucho tiempo, principalmente para el consumo de la pulpa del fruto maduro (Schaffer et al., 1986; Merrick, 1995). Adicionalmente, los frutos maduros son un cultivo nutritivo y de bajo costo, lo que lo convierte en un alimento básico en países económicamente poco desarrollados (Paris, 2016). Aunque parece existir mayor preferencia por otras especies en distintas regiones de México, los

tallos tiernos y las flores masculinas (estaminadas) de *C. moschata* son consumidas como verdura cocida, frita, entre otros, comúnmente formando parte de diversos platillos. Por su tamaño y cantidad, las semillas enteras de esta especie son consumidas como botanas asadas y simplemente sazonadas con sal; hoy en día es común encontrarlas en pequeños puestos ambulantes en zonas urbanas de México (Lira et al., 1995).

Cucurbita moschata es una especie primordialmente adaptada a condiciones tropicales, por lo cual es preferentemente cultivada en regiones de clima cálido y alta humedad, de manera que la distribución tradicional del cultivo de *C. moschata* abarca desde las tierras bajas de México hasta Sudamérica. Gracias a que tolera un relativamente amplio intervalo geográfico-altitudinal, también se siembra al norte de México y suroeste de Estados Unidos (Whitaker y Cutler, 1956; Lira et al., 1995; OECD, 2012). Si bien es un cultivo que se maneja principalmente en tierras bajas, los límites altitudinales de esta especie parecen no ser tan estrictos, siendo cultivadas en México en un intervalo superior entre los 2,200–2,3000 msnm (Bukasov, 1981; OECD, 2012). De las especies de calabaza tropicales, esta especie presenta gran tolerancia a altas temperaturas y entre las especies domesticadas del género es la que menos soporta bajas temperaturas, pero es relativamente tolerante a la sequía (Wien, et al., 2002; Paris, 2016). *Cucurbita moschata* es generalmente cultivada en condiciones de temporada larga, aunque en distintas regiones de la Península de Yucatán (México) se han identificado algunas variantes con dos ciclos de vida de diferente duración (Lira, 1985, 1988; Paris, 2016).

Inicialmente, *C. moschata* solo había sido motivo de interés en algunos estudios de relaciones genéticas y de poblaciones debido a la estrecha relación con el Grupo *Argyrosperma* (que incluye el taxón domesticado *C. argyrosperma* spp. *argyrosperma* y su pariente silvestre *C. argyrosperma* spp. *sororia*). Las primeras reconstrucciones filogenéticas utilizando marcadores genéticos (RFLPs de cloroplasto) no resolvieron completamente las relaciones entre estas especies (Wilson et al., 1992). Por ello, los primeros estudios buscaban describir los niveles de diversidad y diferenciación genética de estos tres taxa (Wilson, 1989; Wilson et al., 1992; Merrick, 1991; Decker-Walters et al., 1990). Sin embargo, para este tipo de estudios solo se utilizaban unas cuantas accesiones de la especie y no se consideró el enorme acervo genético depositado en las variedades locales.

La literatura es relativamente vasta respecto a la descripción general de la especie y las variedades que se cultivan en todo el mundo, pero los esfuerzos para detallar los niveles de

variación genética, estructura genética y patrones filogeográficos siguen siendo limitados. Entre los que destacan, tenemos Montes-Hernández y Eguiarte (2002), quienes determinaron los niveles de variación genética de 11 poblaciones *C. argyrosperma* y cuatro poblaciones *C. moschata* a nivel regional en Jalisco (México). Mediante el análisis de aloenzimas reportaron para las muestras de *C. moschata* un porcentaje alto de loci polimórficos ($P = 0.97$) y una diversidad relativamente alta ($H_E = 0.416$). En este estudio, las poblaciones de *C. moschata* mostraban un exceso de heterocigos (con respecto a lo que se esperaría si los apareamientos fueran al azar). La diferenciación genética para el conjunto de taxa fue baja ($F_{ST} = 0.087$). Este último resultado sugiere constante flujo genético entre todo el conjunto de taxa y/o alta cercanía genética entre ellos.

El primer estudio en caracterizar genéticamente accesiones de *C. moschata* de origen Mesoamericano fue el de Barboza et al. (2012). En este trabajo, analizaron la diversidad genética de 218 accesiones de germoplasma de *C. moschata* que conforman parte de la Colección Internacional del Centro Agronómico Tropical de Investigación y Enseñanza (CATIE) en Turrialba, Costa Rica. La mayoría de las semillas analizadas fueron colectadas en México y Centroamérica, y solo algunas cuantas provenientes de Perú y Colombia. Los niveles de variación genética se evaluaron a través de marcadores de ADN ribosomal (ITS_1-ITS_2) y ADN de cloroplasto ($trnL-trnF$). Entre sus resultados, destaca que la diversidad genética de las muestras provenientes de México posee los valores más altos ($H_E = 0.874$ y 24 haplotipos; Tabla 1). Mientras que su análisis de agrupamiento generado mediante la metodología UPGMA no permitió obtener una clara separación entre las accesiones de México, Centroamérica y Perú. Este estudio mostró que *C. moschata* es una especie con una extensa diversidad de variedades y con una alta diversidad genética. Proponen que la distribución de la diversidad genética de esta especie se debe al aislamiento geográfico y las diferencias en las condiciones ambientales entre las regiones donde se cultiva. Finalmente, el acervo en cada región se mantiene gracias a la preservación de las variedades generación tras generación mantenida por los agricultores locales.

El trabajo de Barboza et al. (2012) es pionero en el estudio de la diversidad genética de esta especie, y permitió tener una visión general respecto a los niveles de variación genética de *C. moschata* a una escala regional. Sin embargo, debido a que se estudiaron accesiones de germoplasma y no poblaciones, es posible que dicho trabajo no refleje adecuadamente los acervos genéticos de la especie a una escala local.

***Cucurbita argyrosperma* Huber spp. *argyrosperma*.**

En México *C. argyrosperma* spp. *argyrosperma* es conocida como calabaza pipiana o calabaza mixta. Se trata de la subespecie domesticada del grupo *Argyrosperma* y su pariente silvestre directo es *C. argyrosperma* spp. *sororia* (Lira et al., 1995). La distribución de esta especie parte desde el suroeste de Estados Unidos hasta Centroamérica desde nivel del mar hasta los 1,900 m. Dado que no tolera las temperaturas bajas, su cultivo es principalmente en regiones con climas cálidos y algo secos (Lira et al., 2009). Los restos arqueobotánicos más antiguos fueron encontrados en el estado de Guerrero (México) con una antigüedad de ~8,600 años (Ranere et al., 2009). A partir de datos genómicos, recientemente se propuso que esta especie tiene un origen monofilético y que su domesticación pudo haber sucedido en las tierras bajas de Jalisco en México (Barrera-Redondo et al., 2021).

La calabaza pipiana, como su nombre lo indica, es utilizada en México principalmente para la elaboración del *pipian* a partir del aprovechamiento de sus semillas. Los frutos de esta especie son utilizados como alimentos y aprovechados por sus propiedades medicinales, aunque en menor medida (Villanueva, 2007). Aunque no es la especie morfológicamente más diversa, si despliega una amplia variedad en la forma de sus frutos y sobre todo en el tamaño, forma y coloración de sus semillas (Lira et al., 1995).

El estudio más reciente que analizó ampliamente la diversidad genética de las variedades locales que se distribuyen a lo largo de todo México es el de Sánchez-de la Vega et al. (2018). En este trabajo se analizaron 19 localidades, al mismo tiempo se evaluó y comparó esta especie con su pariente silvestre directo (spp. *sororia*) con datos de seis poblaciones. La variación genética, estimada con nueve loci de microstélites nucleares, fue similar entre ambas subespecies ($H_E=0.428$ en *sororia* y $H_E=0.410$ en *argyrosperma*). Los análisis de estructura genética, implementados en Structure determinaron que existen tres grupos genéticos bien diferenciados que distinguen a las variedades del norte de México, las variedades de la Península de Yucatán y las variedades de la costa del Pacífico. Con base en los modelos de distribución de especies, este estudio fue el primero en proponer que el centro de domesticación del taxa cultivado pudo haber sido en la región del Balsas en Jalisco México.

***Cucurbita pepo* L.**

El taxón domesticado y comestible del Grupo Pepo corresponde a la subespecie *C. pepo* spp. *pepo* (Decker, 1988). Esta calabaza es globalmente conocida como *zucchini* y es la especie de la familia Cucurbitaceae de mayor importancia económica en todo el mundo (Lust y Paris, 2016; Chomicki et al., 2019). Los restos arqueológicos más antiguos para un taxa domesticado del género *Cucurbita* también corresponden a esta especie y datan de hace 10,000 años (Smith, 1997a); en conjunto con la evidencia lingüística y molecular se propone que su domesticación ocurrió en México (Paris et al., 2003; Brown et al., 2012; Castellanos-Morales et al., 2019). *Cucurbita pepo* se cultiva a gran escala en todo el mundo, está adaptada principalmente a climas templados y fríos, aunque algunas pocas de sus variedades se pueden cultivar en regiones cálidas de baja altitud (Villanueva, 2007).

El uso más importante de esta especie es el alimenticio, especialmente los frutos inmaduros e maduros son las estructuras vegetales que más se consumen (Lira et al., 1995; Paris, 2016). Debido al desarrollo de innumerables cultivares comerciales y variedades locales en todo el mundo es la especie que despliega la mayor variación morfológica de frutos (Lira et al., 1995; OECD, 2012).

En vista de su importancia comercial a nivel mundial es la especie de calabaza cultivada más estudiada. La documentación de la diversidad genética y su historia evolutiva comprende un amplio número de trabajos que analizan innumerables accesiones y variedades mejoradas en todo el mundo (p. ej., Wilson, 1992; Wilson et al., 1992; Decker-Walters, 1993; Decker, 1988; Katzir et al., 2000; Sanjur et al., 2002; Ferriol et al., 2003; Paris et al., 2003; Gong et al., 2012). El trabajo más reciente es el de Castellanos-Morales et al. (2019) y sobresale por incluir variedades locales de México en donde se analizaron dos regiones de ADN de cloroplasto y nueve loci de microsátelites nucleares de 13 variedades locales de *C. pepo* spp. *pepo*, cinco variedades comerciales, una población del taxa silvestre *C. pepo* spp. *fraterna* y una accesión de *C. pepo* spp. *ovifera* (taxa cultivado con fines ornamentales). Se estimó la diversidad genética y pusieron a prueba el origen del taxón domesticado con aproximaciones Bayesianas. Las variedades locales del taxón domesticado presentaron alta variación genética y baja diferenciación genética. Debido a la hibridación entre taxa del grupo Pepo definir el pariente silvestre directo de la especie domesticada *C. pepo* spp. *pepo* ha sido complicado, sin embargo, en este estudio se propone a *C. pepo* spp. *fraterna* como el más probable. Los niveles de diversidad genética de las variedades

locales de *C. pepo* spp. *pepo* de Castellanos-Morales et al. (2019) son congruentes con lo reportado por Martínez-González et al. (2021) quien analizó miles de SNPs.

***Cucurbita ficifolia* Bouché**

En México, *C. ficifolia* es mejor conocida como chilacayote, mientras que en regiones de Sudamérica es simplemente llamada calabaza (Lira et al., 2016). Así como *C. moschata*, el pariente silvestre de *C. ficifolia* parece corresponder a una especie aun no descrita o no colectada (Lira et al., 1995; Hernández-Rosales et al., 2020). Sin embargo, debido a que los registros arqueológicos más antiguos (~5,000 años) han sido encontrados en Perú (Whitaker, 1981), así como la documentación de una importante diversidad morfológica de sus frutos en esa región, se ha propuesto que su origen y domesticación pudiera ser Andino (Barrera-Redondo et al., 2020).

Comparada con el resto de las especies cultivadas del género, *C. ficifolia* es la menos diversa y no se conoce la existencia de cultivares comerciales. Su diversidad predomina en los patrones de coloración y dimensiones de sus frutos (Lira et al., 1995; Barrera-Redondo et al., 2020). Generalmente se cultiva en traspatios y huertos familiares (Lira y Montes-Hernández, 1992; Lira et al., 1995). Esta especie se caracteriza por su hábitat restringido, ya que su distribución bajo cultivo abarca las zonas medias altas (1,000–3,000 msnm) en climas preferentemente templados (Lira et al., 1995). Los frutos inmaduros se consumen como verduras en distintos platillos, aunque predomina la preferencia de consumir sus frutos en platillos dulces (Lira et al., 1995).

Poco se sabe de la diversidad genética de esta especie, solo el estudio de Moya-Hernández et al. (2018) analizó algunas poblaciones mexicanas de *C. ficifolia* mediante polimorfismos de longitud de fragmentos amplificados (AFLPs por sus siglas en inglés) e identificó agrupaciones relacionadas con el color del fruto y el color de la semilla. Con estos resultados sugiere que los fenotipos observados en las accesiones están ligados a sus genotipos, formando linajes evolutivos discretos.

***Cucurbita maxima* spp. *maxima* Duchesne.**

Aunque es una especie cultivada en muchas partes del mundo, la evidencia arqueológica muestra que el origen y domesticación de este taxón sucedió en América del Sur (Lira et al., 1995). Su distribución comprende desde Chile hasta Bolivia, sin embargo, al igual que otras especies

cultivadas del género, su adaptación a un amplio intervalo de condiciones ecológicas se ha logrado a través del desarrollo de múltiples cultivares comerciales y razas criollas (Lira et al., 1995; OECD, 2012). Esta calabaza es comúnmente reconocida en todo el sur de América como zapallo, mientras que en México se le conoce como calabaza kabosha o calabaza de pepita rusa.

Esta especie despliega una importante diversidad de razas locales y abundantes cultivares comerciales tanto para su consumo como alimento, así como para su uso ornamental (Lira et al., 1995). Los frutos maduros que se destinan al consumo alimenticio son principalmente apreciados por su pulpa. Los estudios de su diversidad genética y morfológica en México son pocos, debido a que en esta región del continente no se conocen variedades locales o razas criollas. Sin embargo, se ha identificado que si se llegan a producir algunas de las variedades mejoradas como la kabosha (Hernández, 2017). Por su parte, en regiones del América del Sur se ha reportado un aumento en el desplazamiento de razas criollas por variedades mejoradas (Barrera-Redondo et al., 2020). Aun con la importancia económica que representa esta especie, poco se sabe de la diversidad genética y de los recursos genéticos que representa en regiones donde se cultiva ampliamente.

Capítulo 1: GENÉTICA DE POBLACIONES Y FILOGEOGRAFÍA DE *Cucurbita moschata*, Duchesne ex Lam. Duchesne ex Poir. (CUCURBITACEAE) EN MÉXICO.

Artículo de investigación: Hernández-Rosales, H. S., G. Castellanos-Morales, G. Sánchez-de la Vega, E. Aguirre-Planter, S. Montes-Hernández, R. Lira-Saade, y L. E. Eguiarte. 2020. Phylogeographic and population genetic analyses of *Cucurbita moschata* reveal divergence of two mitochondrial lineages linked to an elevational gradient. *American Journal of Botany*. 107: 510-525.

Cucurbita moschata es una de las cinco especies domesticadas de calabazas y es un cultivo importante que se distribuye y utiliza ampliamente desde Mesoamérica hasta Sudamérica. Esta especie se caracteriza por su gran diversidad morfológica dentro y entre variedades locales. En este primer capítulo se presenta el primer estudio formal de las relaciones filogeográficas de las variedades locales que se cultivan en México de esta especie y se determinaron los niveles de diversidad genética y su estructura genética. Se realizaron coletas de frutos maduros de *C. moschata* a lo largo de la República Mexicana durante los veranos de 2014–2016. Aunque la colección de germoplasma obtenida para esta especie es de alrededor de 50 localidades, inicialmente se eligieron 24 localidades para su análisis genético (siete de ellas corresponden a colecciones de banco de germoplasma). Se tuvo el cuidado de seleccionar muestras que representaran la distribución natural de esta especie, partiendo de los estados más al norte de México, la vertiente del Pacífico, el centro del país, la costa del Golfo de México y la Península de Yucatán. Cada localidad estaba representada con 4–5 frutos, de los cuales se sembraron, en condiciones controladas de invernadero, entre 5–6 semillas. A partir del tejido fresco se realizaron las extracciones de ADN. Para caracterizar la diversidad genética y evaluar los patrones de estructura genética y filogeografía se utilizaron los siguientes marcadores moleculares: tres regiones no codificantes de ADN de cloroplasto, una región no codificante de ADN mitocondrial y once loci de microsatélites de ADN nuclear. Se obtuvieron los estadísticos clásicos de diversidad genética (S , H_d , π , P , A , H_o , H_e , F_{IS}) y una red de haplotipos. La caracterización de los patrones de estructura se realizó con distintas metodologías de asignación bayesiana y análisis de conglomerados. La diferenciación genética se estimó con las medidas básicas de F_{ST} . Durante el desarrollo de los primeros análisis se identificó una estructura asociada

a un gradiente altitudinal, el cual no había sido previamente reportado. Así, que para poner a prueba esta hipótesis se dividieron las localidades en dos grupos: variedades locales de tierras altas (1,269–2,975 m) y variedades locales de tierras bajas (8–1,193 m). El análisis de varianza molecular del marcador de ADNmt y una prueba de regresión múltiple en matrices de distancia validaron nuestra propuesta. Aún cuando se eligieron marcadores de cloroplasto reportados como altamente variables, los niveles de variación fueron nulos para esta especie y por ello se descartaron de los análisis. Por el contrario, y para nuestra sorpresa, nuestros cebadores para los PCR de la región *trnL-trnF* de origen de cloroplasto, amplificaron un locus de ADNmt. Con esta región se encontraron 11 haplotipos y logramos identificar dos linajes asociados a nuestra clasificación por gradiente altitudinal. La estructura genética identificada mediante los marcadores nucleares nos reveló una distribución de la variación en dos grupos: las variedades de la Península de Yucatán y las variedades del resto de México a las que llamamos “variedades Continentales”. Nuevamente, un análisis de varianza molecular soportó nuestra propuesta. Los niveles intermedios de variación para esta especie en México fueron congruentes con lo reportado en la literatura, mientras que los niveles de endogamia resultaron más altos respecto a otras especies domesticadas del género. Sin embargo, la endogamia no es un problema crucial en esta especie, aparentemente gracias a las prácticas de los agricultores quienes fomentan el flujo genético mediante intercambio de semillas entre localidades.

Phylogeographic and population genetic analyses of *Cucurbita moschata* reveal divergence of two mitochondrial lineages linked to an elevational gradient

Helena S. Hernández-Rosales^{1,2}, Gabriela Castellanos-Morales³, Guillermo Sánchez-de la Vega¹, Erika Aguirre-Planter¹, Salvador Montes-Hernández⁴, Rafael Lira-Saade^{5,6}, and Luis E. Eguiarte^{1,6}

Manuscript received 13 September 2019; revision accepted 26 November 2019.

¹ Departamento de Ecología Evolutiva, Instituto de Ecología, Universidad Nacional Autónoma de México, Circuito Exterior s/n Anexo al Jardín Botánico, 04510, Ciudad de México, Mexico

² Programa de Doctorado en Ciencias Biomédicas, Instituto de Ecología, Universidad Nacional Autónoma de México, Circuito Exterior s/n Anexo al Jardín Botánico, 04510, Ciudad de México, Mexico

³ Departamento de Conservación de la Biodiversidad, El Colegio de la Frontera Sur, Villahermosa, Carretera Villahermosa-Reforma km 15.5 Ranchería El Guineo 2ª sección, 86280, Villahermosa, Tabasco, Mexico

⁴ Campo Experimental Bajío, Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias, Km 6.5 carretera Celaya-San Miguel de Allende, C.P. 38110, Celaya, Guanajuato, México, Mexico

⁵ UBIPRO, Facultad de Estudios Superiores Iztacala, Universidad Nacional Autónoma de México, Av. de los Barrios #1, Col. Los Reyes Iztacala, 54090, Tlanepantla, Edo. de Mex, Mexico

⁶ Authors for correspondence (e-mail: fruns@unam.mx; rlira@unam.mx)

Citation: Hernández-Rosales, H. S., G. Castellanos-Morales, G. Sánchez-de la Vega, E. Aguirre-Planter, S. Montes-Hernández, R. Lira-Saade, and L. E. Eguiarte. 2020. Phylogeographic and population genetic analyses of *Cucurbita moschata* reveal divergence of two mitochondrial lineages linked to an elevational gradient. *American Journal of Botany* 107(3): 1–16.

doi:10.1002/ajb2.1424

PREMISE: Domestication usually involves local adaptation to environmental conditions. *Cucurbita* species are a promising model for studying these processes. *Cucurbita moschata* is the third major crop in the genus because of its economic value and because it displays high landrace diversity, but research about its genetic diversity, population structure, and phylogeography is limited. We aimed at understanding how geography and elevation shape the distribution of genetic diversity in *C. moschata* landraces in Mexico.

METHODS: We sampled fruits from 24 localities throughout Mexico. We assessed 11 nuclear microsatellite loci, one mtDNA region, and three cpDNA regions but found no variation in cpDNA. We explored genetic structure with cluster analysis, and phylogeographic relationships with haplotype network analysis.

RESULTS: Mitochondrial genetic diversity was high, and nuclear genetic differentiation among localities was intermediate compared to other domesticated *Cucurbita*. We found high levels of inbreeding. We recovered two mitochondrial lineages: highland (associated with the Trans-Mexican Volcanic Belt) and lowland. Nuclear microsatellites show that localities from the Yucatan Peninsula constitute a well-differentiated group.

CONCLUSIONS: Mexico is an area of high diversity for *C. moschata*, and these landraces represent important plant genetic resources. In Mexico this species is characterized by divergence processes linked to an elevational gradient, which could be related to adaptation and may be of value for applications in agriculture. The Isthmus of Tehuantepec may be a partial barrier to gene flow. Morphological variation, agricultural management, and cultural differences may be related to this pattern of genetic structure, but further studies are needed.

KEY WORDS Cucurbitaceae; evolutionary divergence; genetic diversity; genetic resources; Mexico; squash.

Domestication usually involves local adaptation, as human populations locally select attributes related to morphological and physiological differentiation, enabling ecological adaptation to local climatic and soil conditions (Casas et al., 2007; Paris, 2016). These processes result in genetically structured subgroups of the domesticated species. In this sense, the genetic structure and morphological differentiation of certain domesticated species—such as maize (*Zea mays*), beans (*Phaseolus* spp.), and squashes (*Cucurbita* spp.)—have been associated with differences among human cultural management (Casas et al., 1996; Blancas et al., 2016; Sánchez-de la Vega

et al., 2018) and to fruit or seed selection, with favored traits resulting in a particular morphology (Aguirre-Dugua et al., 2013; Blancas et al., 2016). In addition, patterns of genetic structure have been attributed to ecological factors such as elevation gradients in the case of barley and maize (Tanto et al., 2010; Van Heerwaarden et al., 2011).

The genus *Cucurbita* (Cucurbitaceae) is native to the Americas and includes 15 to 20 taxa (species and subspecies), five of which are domesticated (Lira, 1995; Organization for Economic Cooperation and Development [OECD], 2012; Paris, 2016; Kates et al., 2017;

Castellanos-Morales et al., 2018; Eguiarte et al., 2018). Cultivated species comprise a large diversity of fruit shapes, including numerous landraces or varieties, as a result of their wide adaptability and popularity all over the world (OECD, 2012; Castellanos-Morales et al., 2018; Eguiarte et al., 2018). *Cucurbita pepo*, *C. maxima*, and *C. moschata* are considered major crops of worldwide economic importance, as they are cultivated globally (Paris, 2016; Chomicki et al., 2019). For 2017 alone, the world production of pumpkins, squashes, and gourds was 28 million tons on 2 million ha. Asia was the primary region of production (60%), and among countries, Mexico was the sixth-largest producer (FAO, 2019). In Mesoamerica, these plants are an important element of the *milpa* system, not only as a relevant and versatile food source, but also ecologically, because they help prevent weed development, mitigate soil erosion, and maintain soil moisture (Gliessman, 1983; Altieri, 1995; Lozada-Aranda et al., 2017). These species are a promising model for studying domestication, genetic resources, and the process of artificial selection.

Cucurbita moschata is cultivated in many parts of the world for its ripe and unripe fruits, seeds, young stems, tendrils, and flowers, all of which are consumed (OECD, 2012; Paris, 2016). This species is commonly known as *calabaza de castilla* in Mexico and as *zapallo* in Colombia, Ecuador, and Peru (Whitaker, 1947; Lira, 1995). *Cucurbita moschata* and *C. argyrosperma* are sister species (Kates et al., 2017). The precise area of their origin and domestication is unclear, the available evidence being inconclusive because (1) the wild ancestor of *C. moschata* is still unknown—it has been proposed that it may be related to wild *C. argyrosperma* ssp. *sororia* (Nee, 1990; Sanjur et al., 2002; OECD, 2012; Castellanos et al., 2018; Eguiarte et al., 2018), but it may have not been collected yet or may even be extinct (Kates et al., 2017); and (2) high landrace diversity is observed in Mesoamerica and South America (Whitaker and Davis, 1962; Nee, 1990; Sanjur et al., 2002).

The oldest archaeological remains of cultivated species of *Cucurbita* have been found in Mexico, and the oldest *C. moschata* found in archaeological seed records is dated to ~7000 yr BP and was retrieved from Tehuacán, Puebla, Mexico (Whitaker et al., 1957). Additionally, Whitaker and Bird (1949) in Peru (Huaca Prieta) identified seeds and rinds of *C. moschata* with a date of ~5000 yr BP, and Dillehay et al. (2007) recovered macrobotanical specimens (7000 to 4500 yr BP) from the Peruvian Andes. More recent archaeological evidence of *C. moschata* came from the Ocampo caves in Tamaulipas (Mexico), dated ~4000 yr BP (Whitaker et al., 1957).

Landraces of *C. moschata* are mostly cultivated in humid tropical climates in Mexico, Central America, and the Caribbean Islands, even though the species is relatively drought tolerant (Whitaker and Cutler, 1965; Lira, 1995; Lira et al., 2016) and can grow on rainfall alone (OECD, 2012). *Cucurbita moschata* is mainly grown at low altitudes, but some landraces are able to grow above 2200 m a.s.l. (Whitaker, 1968; Bukasov, 1981). In Mesoamerica, cultivation of *C. moschata* is closely associated with the *milpa* traditional management system, in which it is grown alongside maize, beans, and one or two additional *Cucurbita* species (Aguilar et al., 2003; OECD, 2012).

Cucurbita moschata is considered the second most diverse species of the genus after *C. pepo* (Ferriol and Pico, 2008; Gong et al., 2013; Paris, 2016). It exhibits remarkable morphological diversity in its fruit and seed types (e.g., shape, color, size; Figs. 1 and 5B), and life-cycle variants have been reported in Yucatan Mexico (e.g., *xmejen k'uum* [short cycle] and *xnuk k'uum* [long cycle];

Zizumbo-Villareal and Terán, 1985; Zizumbo-Villareal et al., 2010; Paris, 2016). *Cucurbita moschata* has been categorized in four main cultivar groups (Goldman, 2004) and has been classified in more than 20 varieties (OECD, 2012). After European contact with America, *C. moschata* spread rapidly to other continents and adapted to different ecological conditions, promoting further diversification (Lira et al., 1995; Decker-Walters and Walters, 2000).

Many studies have assessed the levels of genetic diversity in germplasm collections of domesticated *Cucurbita* species, mainly in *C. pepo* (Baranek et al., 2000; Ferriol et al., 2003a, b, 2004a, b; Paris et al., 2003; Cerón et al., 2010; Wu et al., 2011; Formisano et al., 2012; Gong et al., 2012, 2013; Inan et al., 2012; Kong et al., 2014; Balvino-Olvera et al., 2017; Kaźmińska et al., 2017; Enríquez et al., 2018; Sánchez-de la Vega et al., 2018; Castellanos-Morales et al., 2019). Few studies have analyzed genetic variation in *C. moschata*, and these included only a few landraces from Mexico (Montes-Hernández and Eguiarte, 2002; Barboza et al., 2012). Therefore, the previous studies lack a broad geographic sampling, particularly from Mexico, where some of the oldest archaeological records have been found, and appropriate phylogeographic and population genetics analyses are lacking.

The present study aims at understanding how geography and elevation shape the distribution of genetic diversity of *C. moschata* landraces in Mexico using organellar DNA regions and nuclear microsatellite loci. It represents a first wide-range analysis of the levels of genetic diversity and the genetic structure within and among landraces in Mexico. Additionally, based on our results, we suggest possible areas for the conservation of these important plant genetic resources.

MATERIALS AND METHODS

Sampling and DNA extraction

We analyzed 24 localities that represent landraces of *Cucurbita moschata* from Mexico. Sampling sites are distributed in altitudes ranging from 8 to 2975 m a.s.l. (Fig. 2A). We recognized, a priori, two levels of elevation: highland (>1200 m a.s.l.) and lowland (<1200 m a.s.l.) (Fig. 2A). To test for independence among highland and lowland localities, we performed a Student's *t*-test using R package "MASS" (Venables and Ripley, 2002).

We collected four or five mature fruits of different maternal plants during field trips (autumns of 2014–2016). We took samples to the Instituto de Ecología, Universidad Nacional Autónoma de México (UNAM), where we opened the fruits and washed and dried the seeds. Samples from six localities came from the germplasm collection of the Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias INIFAP-Campo Experimental Bajío, and another from the germplasm collection of the Facultad de Estudios Superiores Iztacala (FESI, UNAM). The localities, geographic locations, and altitudes are listed in Table 1.

In *Cucurbita* self-fertilization is usually low (given that the flowers are either male or female) and each seed may be the result of cross pollination with a different father (Hurd et al., 1971). To ensure that we obtained a good representation of genetic diversity per locality, we planted four or five seeds per collected fruit from each sampled locality. For the six localities granted by germplasm collections, we selected samples that came from the originally collected fruits and that have not been bred, to minimize both the risk of

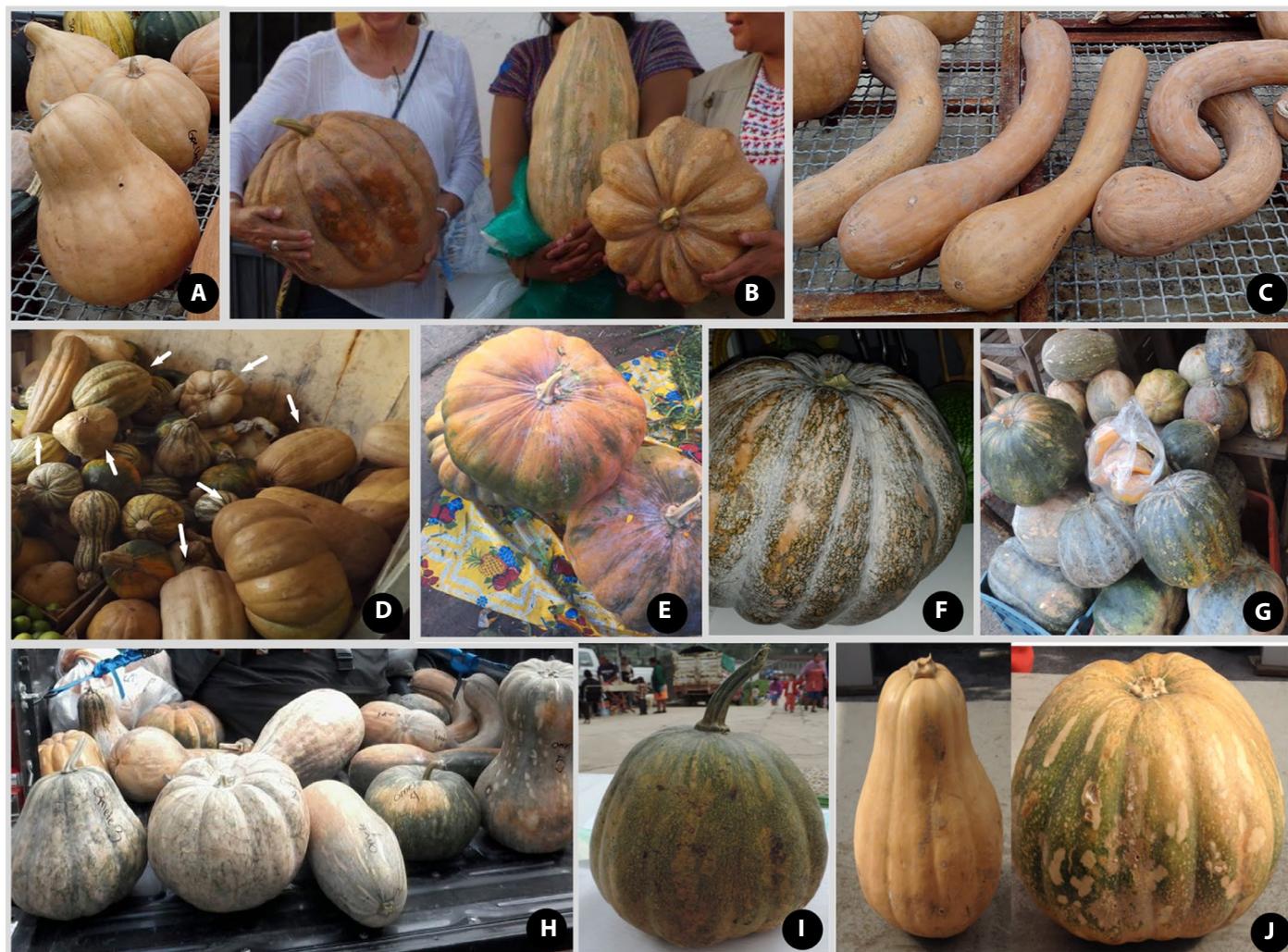


FIGURE 1. Morphological diversity of landraces of *Cucurbita moschata* from north to south in Mexico. (A) Sonora, (B) Coahuila, (C) Sinaloa-MA, (D) Nayarit, (E) Hidalgo, (F) Edomex, (G) Guerrero-CO, (H) Guerrero-OM, (I) Puebla, and (J) Oaxaca-TE. Locality information is shown in Table 1.

genetic contamination and the loss of genetic variation. Of these localities we planted between 14 and 20 random seeds to sample genetic variation.

We grew each plant in a pot with a mix of fertile soil and Miracle-Gro (Marysville, Ohio, USA) medium, under controlled temperature conditions (at 35°C average). Plants were maintained in a greenhouse for 30 d. We extracted total genomic DNA using a modified CTAB method from fresh foliar tissue (Doyle and Doyle, 1987).

Molecular data collection

We amplified four noncoding chloroplast regions: *rpL120-rpS22* (Hamilton, 1999); *psbJ-petA* and *atpI-atpH* (Shaw et al., 2007); and *trnL* (UAA)5'-*trnF* (GAA; Taberlet et al., 1991). We performed amplification by polymerase chain reaction (PCR) in a 25 μ L reaction volume using 10 ng/ μ L of target DNA, 1X buffer, 1.2 mM of MgCl₂, 0.3 mM of dNTPs, 0.32 μ M of each primer, and 2 U/ μ L of GoTaq DNA polymerase (Promega, Madison, Wisconsin, USA). We used a 2720 Thermal Cycler (Applied Biosystems, Foster City, California, USA) with the following program: 5 min at 94°C, then 35 cycles of

40 s at 94°C, 40 s at 52°C, and 120 s at 72°C, followed by 5 min at 72°C and a final step of 4°C. For *psbJ-petA* and *atpI-atpH* regions, we used a primer annealing temperature at 50°C for 1 min, followed by a ramp of 0.3°C/s to 65°C, and primer extension at 65°C for 4 min, followed by a final extension step of 5 min at 65°C. We confirmed PCR products by electrophoresis in 1.5% agarose gels. We sequenced amplicons with forward and reverse primers at Macrogen USA's Maryland headquarters (<http://www.macrogenusa.com>).

We filtered base quality and assembled the amplified sequences with Consed version 6 (Ewing et al., 1998; Gordon et al., 1998). To corroborate the identity of sequences, we performed a BLAST analysis in GenBank, considering 100% average of query cover and >90% identity. We used MEGA version 5 (Tamura et al., 2011) to align DNA sequences and to identify polymorphic sites. We found no variation among our samples in three of the cpDNA regions (*rpL120-rpS22*, *psbJ-petA*, and *atpI-atpH*). Our *trnL-trnF* PCR primers amplified a mitochondrial locus (previously reported by Alverson et al., 2010; GenBank ID GQ856148; 100% query cover and 93% identity) that was variable across our samples and was thus used for downstream analysis. Therefore, we will refer to this

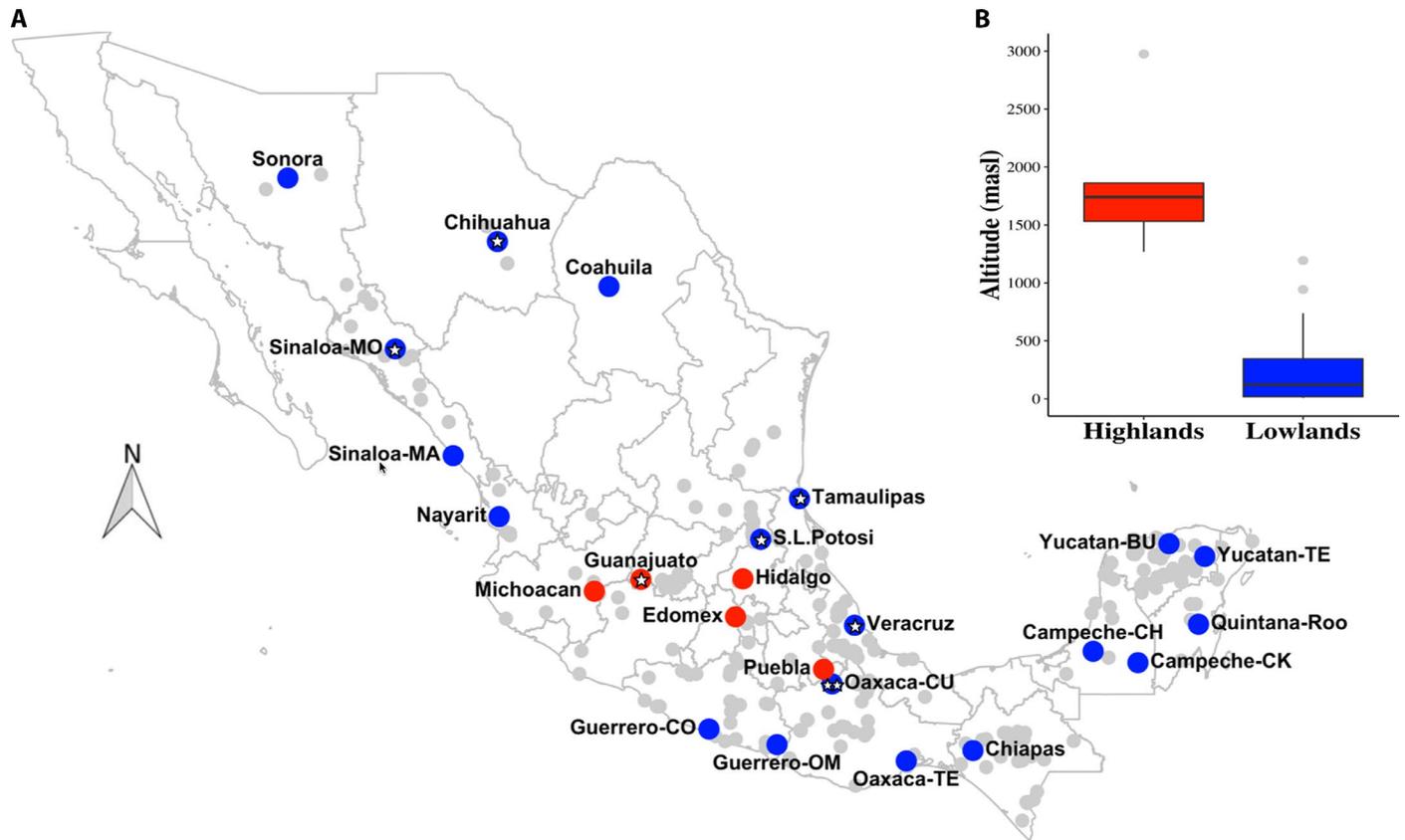


FIGURE 2. (A) Distribution of *Cucurbita moschata* in Mexico (gray points) and sampled localities (blue and red points). Localities were assigned as lowland (blue) and highland (red), according to an elevation gradient. (B) Box plot depicting the significant differences in elevation (m a.s.l.) between highland and lowland localities. Locality information is shown in Table 1.

region as mtDNA hereafter. All sequences were deposited in NCBI GenBank (MN381966–MN381976).

In addition, we amplified 12 nuclear microsatellite loci (CMTm187, CMTp175, CMTm120, CMTm144, CMTm261, CMTm11, CMTp129, CMTp193, CMTm221, CMTm54, CMTp88, and CMTp17) reported by Gong et al. (2008). To avoid linkage disequilibrium, we selected nuclear microsatellite loci from different chromosomes. We performed a multiplex PCR (six primers per multiplex; all forward primers were modified with a fluorescent dye: 6-FAM, HEX, and VIC) in a volume of 25 μ L with the following concentrations: 10 ng/ μ L of target DNA, Buffer 1X, 3 mM of $MgCl_2$, 0.2 mM of dNTPs, 0.1 μ M of each primer, and 2 U/ μ L of Taq DNA polymerase (Promega). We used a 2720 Thermal Cycler (Applied Biosystems), with the following program: 5 min at 95°C, then 35 cycles of 50 s at 95°C, 40 s at 60°C, and 60 s at 72°C, followed by 5 min at 72°C and a final step of 4°C. We confirmed PCR products by gel electrophoresis (2.5%) and sent amplicons for genotyping at the UIUC Core DNA Sequencing Facility, University of Illinois (<https://illinois.edu>).

We analyzed microsatellite electropherograms with PeakScanner software (Applied Biosystems) and constructed a matrix with the genotypes of each individual. To confirm the results, we performed the procedure twice for each sample. We used Micro-Checker version 2.2.3 (Van Oosterhout et al., 2004) to test for null alleles in each locus, allele dropout, and scoring errors. Locus CMTp17 did not amplify in all localities and showed evidence of being monomorphic;

thus, it was excluded from the analyses and results were obtained for 11 nuclear microsatellite loci.

Characterizing genetic diversity

We estimated the genetic diversity for each type of molecular marker and for each locality. For DNA sequences, we first determined whether sequences were polymorphic. For variable regions, we analyzed the relationship between haplotypes with the median-joining algorithm implemented in Network version 5.0.0.0 (Fluxus-engineering, 2015) using a 95% identity criterion to define connection limits. To root the network, we used sequences from two closely related taxa: domesticated *C. argyrosperma* ssp. *argyrosperma* and wild *C. argyrosperma* ssp. *sororia* from the state of Veracruz, Mexico. We estimated the number of segregating sites (S), haplotype diversity (H_d), and nucleotide diversity (π) with DnaSP version 5 (Librado and Rozas, 2009).

For the nuclear microsatellite loci, we calculated the allele frequencies and the number of private alleles per locality through direct count estimation. We also estimated the proportion of polymorphic loci per locality (P), expected heterozygosity (H_e), and observed heterozygosity (H_o) with Arlequin version 3.5.1.2 (Excoffier et al., 2005). We used HP-RARE version 1.0 to estimate allelic richness (A) and corrected it for sample size through rarefaction (Kalinowski, 2004). We performed an exact test for Hardy-Weinberg equilibrium (HWE) at each locus using Markov chain Monte Carlo

TABLE 1. Geographic information for localities of *Cucurbita moschata* collected in Mexico: locality ID, state, town, population, coordinates, altitude, and elevation category. One star indicates landraces provided by the germplasm collection INIFAP-Campo Experimental Bajío; two stars indicate samples provided by the germplasm collection of FESI-UNAM.

ID	Estate	Town	Population	Coordinates		Altitude (m a.s.l.)	Elevation category
				N	W		
Sonora	Sonora	Ures	Gudalupe de Ures	29.428056	-110.38917	386	Lowland
Chihuahua★	Chihuahua	El Saucillo	El Saucillo	28.000000	-105.25000	1193	Lowland
Coahuila	Coahuila	Cuatro Ciénegas de Carranza	Cuatro Ciénegas	26.988611	-102.06639	739	Lowland
Sinaloa-MO★	Sinaloa	Palmarito Mineral	Mocorito	25.583333	-107.75000	215	Lowland
Sinaloa-MA	Sinaloa	Mzatlán	Estación Presidio	23.183300	-106.33300	11	Lowland
Nayarit	Nayarit	Santiago Ixcuintla	Sauta	21.811389	-105.20667	16	Lowland
Michoacan	Michoacan	Cojumatlán	Cojumatlán	20.133333	-102.87500	1531	Highland
Tamaulipas★	Tamaulipas	Tampico	Tampico	22.216667	-97.850000	10	Lowland
S.L. Potosi★	S.L.P.	La Laguna	Tamazunchale	21.300000	-98.800000	259	Lowland
Guanajuato★	Guanajuato	Rcho San Vte	Celaya	20.400000	-101.73333	1741	Highland
Hidalgo	Hidalgo	Ixquimilpan	EL Alberto	20.412500	-99.234444	1862	Highland
Edomex	México	Isidro Fabela	Tlazala de Fabela	19.557778	-99.416667	2975	Highland
Guerrero-CO	Guerrero	Coyuca	El Campanario	17.033333	-100.06667	73	Lowland
Guerrero-OM	Guerrero	Ometepec	Las Vigas	16.683333	-98.400000	304	Lowland
Veracruz★	Veracruz	Pte Nacional	Pte. Nacional	19.366667	-96.500000	122	Lowland
Puebla	Puebla	Ajalpan	Tehuacán Santa Cruz	18.378333	-97.259444	1269	Highland
Oaxaca-TE	Oaxaca	Sto. Dom. Tepc.	Tehuantepec	16.316667	-95.233333	39	Lowland
Oaxaca-CU★★	Oaxaca	San Juan de los Cués	San Juan de los Cués	18.050000	-97.050000	943	Lowland
Chiapas	Chiapas	Jiquipilas	El Arroyón-Sinaloa	16.550000	-93.600000	636	Lowland
Campeche-CK	Campeche	Calakmul	Chichonal	18.783300	-90.650000	252	Lowland
Campeche-CH	Campeche	Champotón	Venustiano Carranza	18.527166	-89.552806	89	Lowland
Yucatan-BU	Yucatan	Buctzotz	Esperanza	21.201667	-88.792778	8	Lowland
Yucatan-TE	Yucatan	Temozón	Ek-Balam	20.916667	-87.916667	20	Lowland
Quintana-Roo	Quintana Roo	Felipe Carrillo Pto.	X-hazil	19.390000	-88.069833	11	Lowland

(MCMC) with 10,000 dememorization steps, 1000 batches, and 5000 iterations per batch. We estimated linkage disequilibrium (LD) and deviations from HWE for each locality through F_{IS} (inbreeding coefficient; Weir and Cockerham, 1984), using the EM algorithm implemented in Genepop version 4.0 (Raymond and Rousset, 1995).

Characterizing genetic structure

To assess the genetic differentiation among localities, we estimated pairwise F_{ST} . With DNA sequences, we estimated F_{ST} based on the number of pairwise differences between haplotypes with Arlequin. We estimated Nei's pairwise F_{ST} for microsatellite data using the package "hierfstat" (Goudet, 2005) for R version 1.4.2 (R Core Team, 2013). To depict the relationships between localities, we constructed a dendrogram with each F_{ST} matrix using the complete agglomeration method and "hclust" function from the R package "ape" (Paradis et al., 2004).

To further explore genetic structure for DNA sequences, we conducted a Bayesian analysis with BAPS version 6.0 (Corander et al., 2008). We used a nonspatial model, performed in five independent runs testing for K (number of clusters) = 1–8.

For nuclear microsatellite loci, we used STRUCTURE version 2.2 (Pritchard et al., 2000). Sampled locations from *C. moschata* share ancestry, and gene flow may be taking place; therefore, we defined the admixture model and correlated allele frequencies as priors. We conducted a previous run to determine the best amount of burn-in chains and MCMC simulations that would maximize the posterior likelihood. We explored values of K from 1 to 8 to maximize the posterior distributions of the model. Accordingly, for the

final run we used as parameters 500,000 burn-in chains, 1 million MCMC, and 15 repetitions for each K (1–8). To identify the optimal number of K we used the ΔK statistic on the rate of change between successive K -values, as proposed by Evanno et al. (2005). The Evanno test was performed with the online program STRUCTURE HARVESTER (Earl and von Holdt, 2012).

As an alternative approach to explore genetic structure for microsatellite data, we performed a discriminant analysis of principal components (DAPC; Jombart et al., 2008) using the R package "adegenet." DAPC is a multivariate method that summarizes the genetic differentiation between groups, while overlooking within-group variation (Jombart et al., 2010). We performed a cross-validation test to determine the number of principal components (PCs) to be retained. In accordance with the result, we retained 60 PCs and two discriminant functions (linear combinations of variables).

To accurately detect structure among localities within the Yucatan Peninsula plus Chiapas and within the rest of the samples (continental Mexico), we performed two independent runs with STRUCTURE using nuclear microsatellite loci following the same strategy as explained above.

Testing for association of genetic differentiation to elevation and geography

To define whether the distribution of genetic variance is better explained by geography or elevation for both molecular markers, we used Arlequin to perform an analysis of molecular variance (AMOVA). We tested two hypotheses: (1) test for genetic differentiation by elevation (highland and lowland); and (2) test for

genetic differentiation by geography, considering the Isthmus of Tehuantepec as a geographic barrier as has been described for the closely related *C. argyrosperma* (Sánchez-de la Vega et al., 2018).

To further test for the roles of geography and elevation in genetic differentiation, we performed a multiple regression on distance matrices (MRM) for both DNA sequences and nuclear microsatellite loci. We performed MRM using the R package “ecodist” (Goslee and Urban, 2007). MRM uses multiple regressions to test the association between a response distance matrix and two or more explanatory distance matrices. Therefore, it tests more than one explanatory variable and determines their individual significance (Lichstein, 2007). The significance and regression coefficients of the model are tested by permuting the response distance matrix, while holding the other matrices constant. In this case, we have two independent response matrices, linearized genetic distance for DNA sequences and linearized genetic distance for microsatellite loci ($F_{ST} / (1 - F_{ST})$), and our explanatory distance matrices were geographic and elevation distances between localities. We performed 1000 iterations for each MRM model to obtain the significance of the geographic and elevation matrices on the genetic distance, performing independent tests for DNA sequences and for nuclear microsatellite loci.

Finally, we assessed genetic structure in accordance with these spatial models (elevation and geography). For DNA sequences, we performed a spatial analysis with BAPS, which considered the geographic location of the samples ($K = 2$). We selected this K -value to test whether genetic structure was explained by elevation, considering the two categories previously defined (highland and lowland) or geography, with the Isthmus of Tehuantepec as a genetic barrier, as mentioned above.

RESULTS

Sampled localities can be grouped in two elevational categories (highland and lowland; $t = 5.256$; $df = 4.62$; $P = 0.004$). Five localities (Puebla, Michoacan, Guanajuato, Hidalgo, and Edomex) corresponded to highland with an elevation range of 1269–2975 m and an average of 1875.6 m. The rest of the localities were defined as lowland, with an elevation range of 8–1193 m and an average of 280.3 m (Fig. 2B).

Genetic diversity

The total length of the mtDNA sequence was 1027 bp with seven polymorphic sites. We identified 11 mtDNA haplotypes (A–K), which can be grouped in two haplogroups in relation to the position of the haplotypes from sister taxon *C. argyrosperma* in the haplotype network. The first haplogroup consists of haplotypes A, C, E, F, H, and I (Fig. 3A). The second haplogroup contains haplotypes B, D, G, J, and K (Fig. 3A). Haplotype A (blue in Fig. 3B) and haplotype B (pink in Fig. 3B) represent 55.37% and 30.08% of the total samples, respectively. Haplotype A was fixed in Sinaloa-MA, Guerrero-OM, Campeche-CK, and Yucatan-TE (Fig. 3B), while haplotype B was fixed in Sinaloa-MO, Edomex, and Puebla but was absent in the Yucatan Peninsula. The rest of the haplotypes were found in low frequencies; few of these haplotypes were found in >30% frequency, and in only one or two sampled locations: haplotype D (purple) in Tamaulipas, F (brown) in Oaxaca-CU, J (yellow) in Oaxaca-TE, and C (dark cyan) in

Quintana-Roo (Fig. 3B). Veracruz had six of the 11 haplotypes and Guerrero-CO had four haplotypes.

For the mtDNA region, mean sample size was 10.5 individuals (7–16). Average H_d was 0.316. Veracruz showed the highest genetic diversity ($H_d = 0.889$; Table 2). Average π was 0.0023, with values ranging from zero in eight localities to 0.0027 in Veracruz (Table 2).

For nuclear microsatellite loci, mean sample size was 18 individuals (14–22; Table 3). There was no evidence of null alleles or allele dropout for the 11 analyzed loci. We found that LD between all pairs of microsatellite loci was significant in only 8% of the pairs ($P < 0.05$); however, these patterns of LD were not consistent between localities, which suggests a lack of LD among analyzed loci.

We obtained a total of 122 microsatellite alleles (from four alleles in locus CMTp129 to 20 alleles in locus CMTm187; Appendix S1). All loci were polymorphic (average $P = 0.75 \pm 0.21$ [SD]), with a range from 0.27 in Puebla to 1.00 in five localities (Sonora, Nayarit, Michoacan, Guerrero-CO, and Campeche-CH). The mean frequency of private alleles was 0.0457. Half of the sampled localities showed private alleles, ranging from one in six localities (Coahuila, Sinaloa-MO, Guerrero-CO, Guerrero-OM, Puebla, and Yucatan-TE) to nine in Sinaloa-MA and Nayarit (Table 2). Mean allelic richness (A) after rarefaction was 2.96 (SD = 0.5), and it ranged from 1.87 in Yucatan-BU to 3.88 in Sinaloa-MA (Table 2). Mean genetic diversity (H_e) was 0.395 (SD = 0.06), with the lowest value found in Yucatan-BU ($H_e = 0.246$) and the highest in Nayarit ($H_e = 0.491$; Table 2).

All microsatellite loci showed significant deviations from HWE in at least one locality, but deviations were not consistent across samples. Most of the localities showed an excess of homozygosity ($P < 0.050$ or $P < 0.001$; Table 2) and relatively high inbreeding coefficients (mean $F_{IS} = 0.180$). Five localities did not show deviations from HWE (Sinaloa-MO, Tamaulipas, S.L. Potosi, Guanajuato, and Guerrero-CO; Table 2).

Genetic structure

Overall population differentiation for mtDNA was high ($F_{ST} = 0.469$). Genetic differentiation among pairs of localities ranged from 0.001 (Coahuila–Sonora) to 1 (Puebla–Sonora-MA, Guerrero-OM, Campeche-CK, and Yucatan-TE; Campeche-CK–Sinaloa-MO and Edomex; Yucatan-TE–Sinaloa-MO and Edomex; Appendix S2). The dendrogram showed three well-defined groups (Fig. 4A). The first group, except Sinaloa-MO, consists of highland localities that exhibited high frequencies of haplotype B (pink). The second group includes localities with a high frequency of haplotype A (blue), and the third group consists of localities with frequencies <50% for either A or B. The BAPS analysis for mtDNA, without considering spatial data, showed structure among sampled localities. The number of groups in the optimal partition was $K = 7$ ($\ln(P) = -82.07$; Appendix S3). All individuals from Sinaloa-MO, Edomex, and Puebla were assigned to Cluster 1 (red). Most individuals from the Yucatan Peninsula and Chiapas were assigned to Cluster 4 (green). A small proportion of individuals from six localities (Chihuahua, Tamaulipas, S.L. Potosi, Guerrero-CO, Oaxaca-TE, and Oaxaca-CU) were assigned to clusters 2, 3, 5, 6, and 7. Individuals from Veracruz (the most diverse population) were assigned to five genetic clusters (Appendix S3).

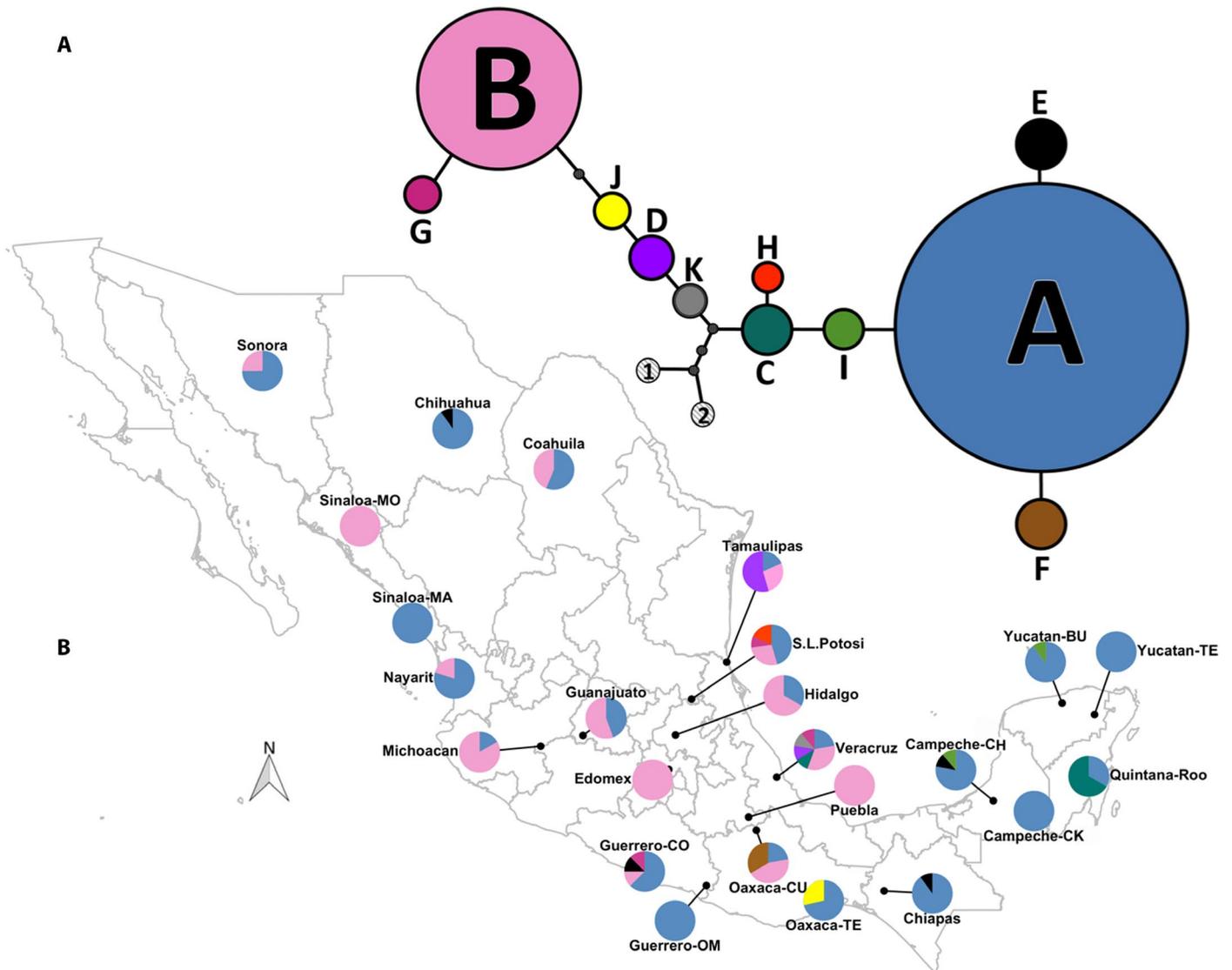


FIGURE 3. (A) Median-joining network of haplotypes found in *Cucurbita moschata* (mtDNA *trnL-trnF*); circle diameter is proportional to haplotype frequency. The network was rooted using *C. moschata*'s close relatives: 1 = *C. argyrosperma* ssp. *argyrosperma* (domesticated) and 2 = *C. argyrosperma* ssp. *sororia* (wild). (B) Geographic distribution and frequency of *C. moschata* haplotypes for each sampled location.

For nuclear microsatellite loci, $F_{ST} = 0.215$. Values among pairs of localities ranged from 0.041 (Nayarit–Sinaloa-MA) to 0.340 (Yucatan-BU–Guerrero-CO; Appendix S2), and Yucatan-BU was the most genetically differentiated locality. The dendrogram showed that the localities from the Yucatan Peninsula constitute a distinct group (Fig. 4B).

STRUCTURE analysis for nuclear data identified the optimal number of clusters (K) as 2 ($\ln(P) = -10,064.7$), but there was not a clear pattern among localities (Evanno's test and barplot $K = 2$; Appendix S4). Evanno's test has a tendency for optimal $K = 2$ (Janes et al., 2017), and the ΔK and posterior likelihood values were very similar for $K = 2$ and $K = 3$. We determined a clustering of $K = 3$ ($\ln(P) = -9456$; Fig. 5A) in our analysis, because it explains the data in accordance with the geographic distribution of this species. For $K = 3$, individuals from Sonora, Coahuila, Edomex, and Hidalgo were assigned to Cluster 1 (brown), and few individuals from Guerrero-OM were assigned to this genetic group. All individuals from Chihuahua, Sinaloa-MO, Michoacan, Tamaulipas, S.L. Potosi,

Guanajuato, and Guerrero-CO were assigned to Cluster 2 (green). Individuals from the Yucatan Peninsula and Chiapas were assigned to Cluster 3 (purple). Sinaloa-MA, Nayarit, Puebla, and Oaxaca (TE and CU) were the most diverse in terms of individual assignment to several clusters (Fig. 5A). Landraces from the Isthmus of Tehuantepec to the Yucatan Peninsula were genetically different from the rest of the sampled localities ($F_{ST} = 0.338$).

The DAPC results for nuclear data are similar to the results from STRUCTURE analysis with $K = 3$ (Appendix S5): localities from the Isthmus of Tehuantepec and the Yucatan Peninsula form a well-differentiated group. S.L. Potosi, Guanajuato, Sinaloa-MO, and Michoacan were clustered in one group. The rest of the localities did not show any clear genetic structure.

Localities within the Yucatan Peninsula and Chiapas showed structure with the nuclear microsatellite data. The best K was 3 ($\ln(P) = -2163.7$; Fig. 5B). The greater proportion of ancestry of Cluster 1 (blue) prevails in Chiapas. Most individuals from Quintana-Roo were assigned with high probability to Cluster

TABLE 2. Genetic diversity obtained for mtDNA and 11 nuclear microsatellite loci for *Cucurbita moschata* in Mexico: locality ID, number of individuals analyzed (n), number of haplotypes (h), haplotype diversity (H_d), nucleotide diversity (π), proportion of polymorphic loci (P), number of private alleles (PA), rarefied allelic richness (A), observed heterozygosity (H_o), expected heterozygosity (H_e), and inbreeding coefficient (F_{IS}). One star indicates landraces provided by the germplasm collection INIFAP-Campo Experimental Bajío; two stars indicate samples provided by the germplasm collection of FESI-UNAM. Statistical significance: $P < 0.05$ (*), statistical significance $P < 0.001$ (**), standard deviations (SD) are shown in parentheses in the bottom row.

ID	mtDNA				Nuclear microsatellite loci						
	n	h	H_d	π	n	P	PA	A	H_o	H_e	F_{IS}
Sonora	12	2	0.409	0.0020	20	1.00	2	3.02	0.328	0.408	0.348**
Chihuahua★	10	2	0.200	0.0002	14	0.55	0	3.01	0.484	0.473	0.063**
Coahuila	16	2	0.525	0.0026	20	0.82	1	2.90	0.356	0.397	0.269**
Sinaloa-MO★	10	1	0.000	0.0000	20	0.91	1	3.05	0.443	0.433	-0.021
Sinaloa-MA	11	1	0.000	0.0000	19	0.36	9	3.88	0.304	0.429	0.362**
Nayarit	10	2	0.356	0.0017	19	1.00	9	3.71	0.316	0.491	0.422**
Michoacan	12	2	0.303	0.0015	21	1.00	2	2.28	0.374	0.303	0.217*
Tamaulipas★	11	3	0.655	0.0018	17	0.91	0	3.02	0.536	0.396	-0.111
S.L. Potosi★	11	4	0.745	0.0027	19	0.73	0	3.51	0.493	0.465	-0.058
Guanajuato★	9	2	0.556	0.0027	20	0.64	0	3.37	0.440	0.431	-0.027
Hidalgo	12	2	0.485	0.0023	14	0.82	0	2.71	0.522	0.421	0.096*
Edomex	16	1	0.000	0.0000	16	0.36	0	2.43	0.357	0.315	0.090**
Guerrero-CO	8	4	0.643	0.0025	20	1.00	0	2.34	0.345	0.306	-0.024
Guerrero-OM	11	1	0.000	0.0000	16	0.82	1	2.76	0.352	0.375	0.328**
Veracruz★	9	6	0.889	0.0027	18	0.82	1	3.31	0.520	0.447	0.045*
Puebla	14	1	0.000	0.0000	19	0.27	1	3.50	0.268	0.471	0.493**
Oaxaca-TE	7	2	0.476	0.0019	18	0.73	4	3.20	0.479	0.403	0.029
Oaxaca-CU★★	9	3	0.722	0.0024	15	0.45	2	3.68	0.206	0.444	0.550**
Chiapas	10	2	0.200	0.0002	20	0.82	0	2.71	0.311	0.317	0.297**
Campeche-CK	9	3	0.417	0.0004	20	0.73	0	2.60	0.352	0.365	0.213**
Campeche-CH	10	1	0.000	0.0000	21	1.00	0	2.79	0.406	0.378	0.224**
Yucatan-BU	10	2	0.200	0.0002	20	0.73	0	1.87	0.354	0.246	0.098**
Yucatan-TE	10	1	0.000	0.0000	18	0.73	1	2.63	0.431	0.397	0.219**
Quintana-Roo	9	2	0.500	0.0005	22	0.73	0	2.96	0.479	0.383	0.101*
	256		0.58900	0.00239	446	0.75 (0.21)		2.96 (0.5)	0.394 (0.09)	0.395 (0.06)	0.180* (0.08)

TABLE 3. Analysis of molecular variance (AMOVA; Excoffier et al., 2005) performed with Arlequin version 3.5.1.2 for 24 sampled localities of *Cucurbita moschata* from Mexico using mitochondrial DNA (mtDNA AMOVA; F_{ST}) and 11 nuclear microsatellite loci (microsatellite AMOVA; ϕ_{ST}). AMOVA test for genetic differentiation by geography, considering the Isthmus of Tehuantepec as a geographic barrier; and test for genetic differentiation by elevation (highland and lowland) as defined in Table 1. Significant values of genetic structure (ϕ_{ST} and F_{ST}) and significant values of inbreeding coefficients (ϕ_{IS} , ϕ_{IT} , ϕ_{SC}) were found within individuals

Source of variation	df	Geography			Elevation		
		Percent variation	F coefficients	P	Percent variation	F coefficients	P
mtDNA AMOVA							
Among groups	1	30.8	F_{ST} 0.307	0.000 ± 0.000	46.94	F_{ST} 0.469	0.00 ± 0.00
Among populations within groups	22	31.32	F_{SC} 0.452	0.003 ± 0.001	19.52	F_{SC} 0.367	0.00 ± 0.00
Within populations	232	37.88			33.54		
Microsatellites AMOVA							
Among groups	1	3.99	ϕ_{ST} 0.039	0.000 ± 0.000	0.80	ϕ_{ST} 0.008	0.00 ± 0.00
Among populations within groups	22	19.9	ϕ_{IS} 0.157	0.000 ± 0.000	21.68	ϕ_{IS} 0.157	0.00 ± 0.00
Among individuals within populations	422	11.95	ϕ_{SC} 0.207	0.000 ± 0.000	12.17	ϕ_{SC} 0.218	0.00 ± 0.00
Within individuals	446	64.16	ϕ_{IT} 0.359	0.003 ± 0.001	65.35	ϕ_{IT} 0.346	0.20 ± 0.01

2 (magenta). Landraces from Campeche (CK and CH) showed high levels of admixture, in contrast to the localities of Yucatan (BU and TE), which showed low levels of admixture because individuals were assigned with a high probability to Cluster 3 (pink; Fig. 5B). For continental Mexico, $K = 2$ ($\ln(P) = -7367.9$) genetic structure among localities based on geography or elevation was not shown.

Test for association of genetic differentiation with elevation and geography

The AMOVA for mtDNA data showed that clustering by elevation explained 46.94% of the genetic variation among groups (highland vs. lowland), and 33.54% was due to within-individual variation (Table 3). The AMOVA performed with the same dataset considering

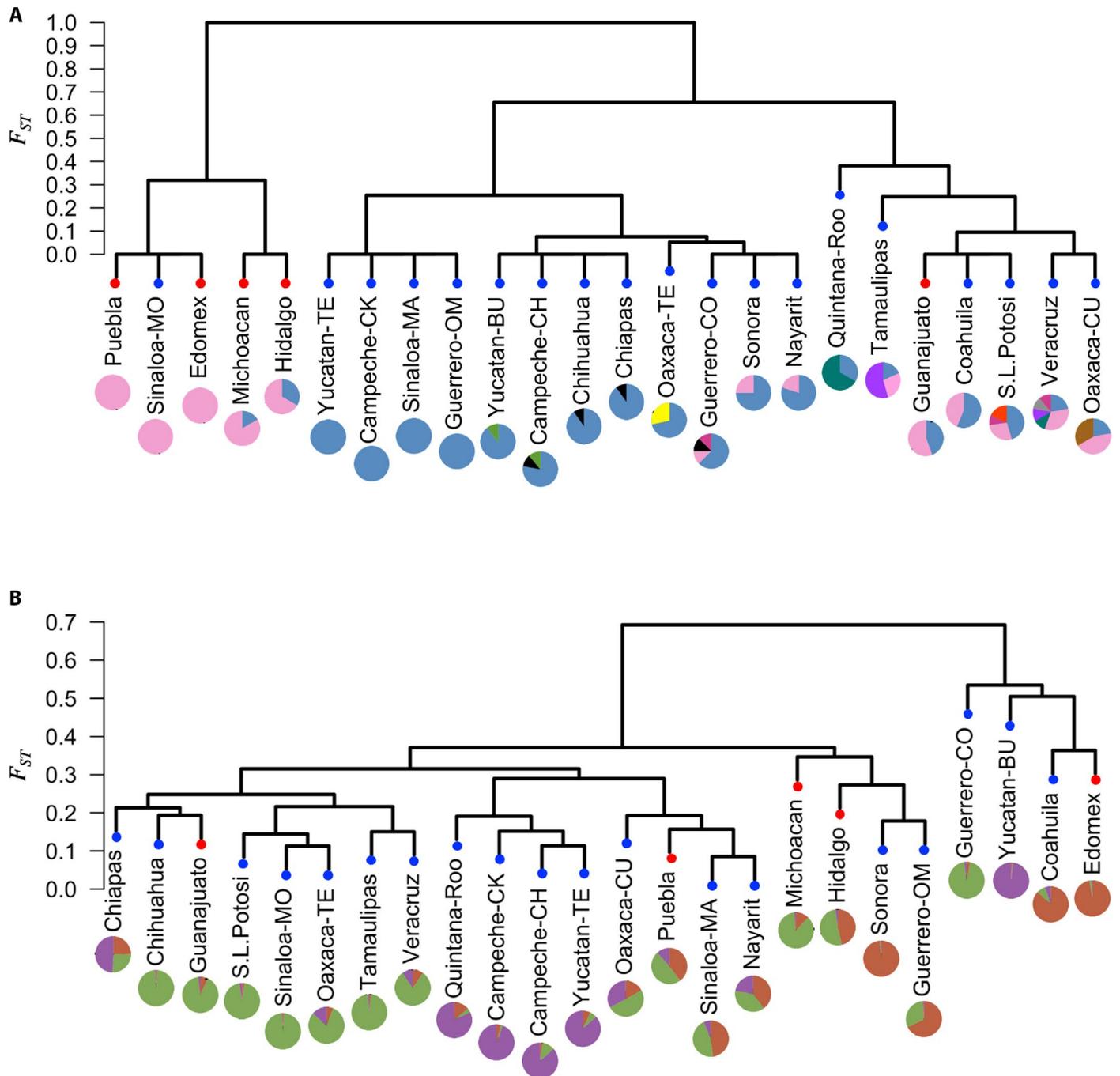


FIGURE 4. Dendrograms depicting pairwise F_{ST} values for 24 sampled locations of *Cucurbita moschata* from Mexico. (A) Mitochondrial region *trnL-trnF*; pie graphs correspond to haplotype frequency per population. (B) Eleven nuclear microsatellite loci; pie graphs correspond to the proportion of individuals per population assigned to each genetic cluster obtained with STRUCTURE for $K = 3$. Highland is indicated with red points and lowland with blue points. Values for pairwise F_{ST} are given in Appendix S2.

the Isthmus of Tehuantepec as a genetic barrier (Yucatan Peninsula and Chiapas vs. the rest of samples) explained 30.80% among group variation and 37.88% within-individual variation (Table 3). The BAPS test for mtDNA for $K = 2$ further supports that genetic structure is associated with elevation (highland and lowland; Appendix S3). Most individuals, especially from lowland localities (Sonora, Nayarit, S.L. Potosi, Guerrero-CO, Guerrero-OM, Chiapas, Campeche-CK, Campeche-CH, Yucatan-BU, Yucatan-TE, and Quintana-Roo) were

assigned to Cluster 1 (blue). By contrast, most individuals from highland localities from the Trans-Mexican Volcanic Belt (Puebla, Michoacan, Guanajuato, Hidalgo, and Edomex) were assigned to Cluster 2 (red). Even though Tamaulipas and Veracruz are lowland localities, they are close to the Trans-Mexican Volcanic Belt, which may explain their assignment to Cluster 2 (Appendix S3).

By contrast, in the AMOVA of nuclear microsatellite loci, elevation explained only 0.8% of among-group variation and 65.35% of

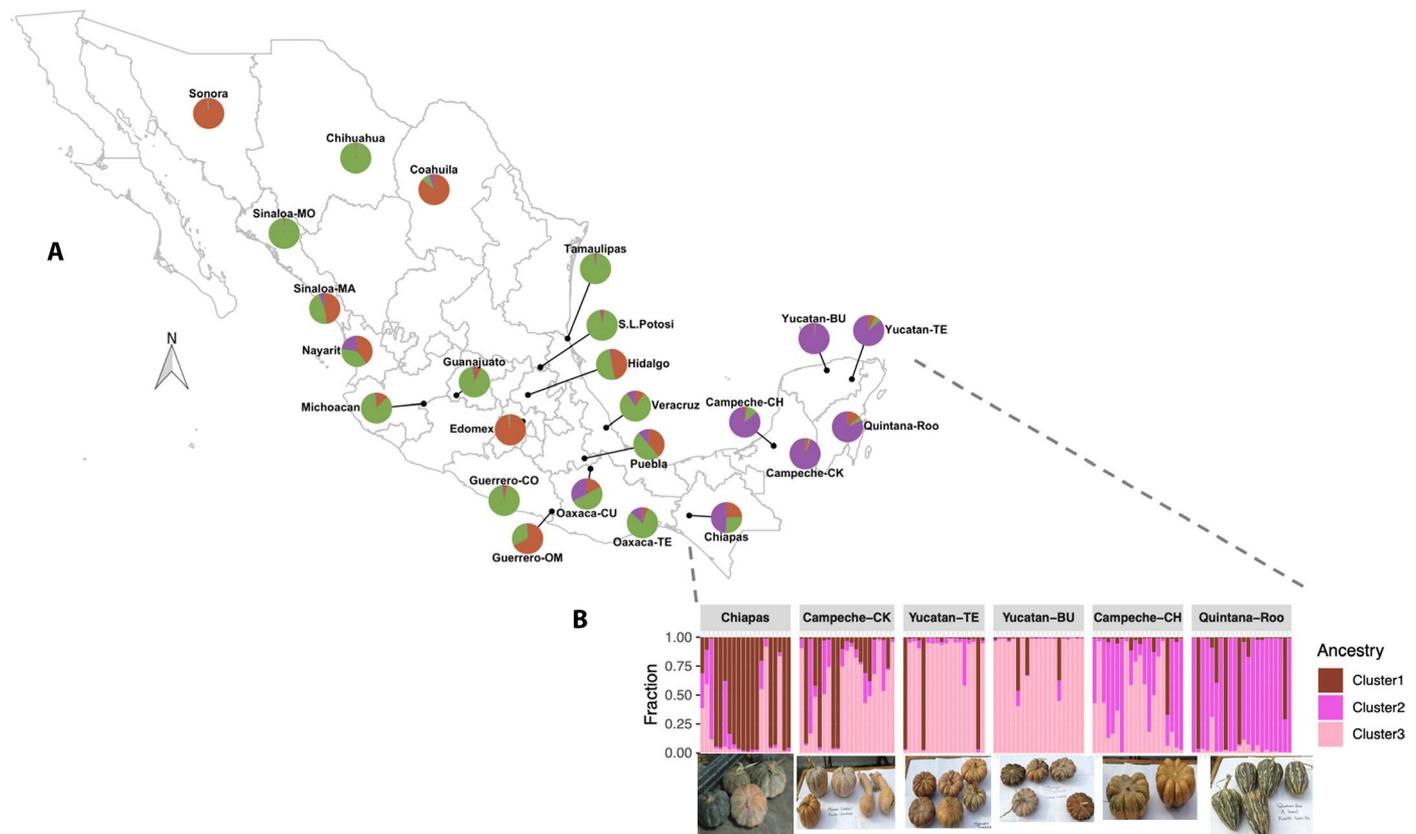


FIGURE 5. (A) Histogram showing the proportion of each individual per locality of *Cucurbita moschata* assigned to different genetic clusters obtained with STRUCTURE version 2.2 for $K = 3$. A high proportion of individuals from Sonora, Coahuila, and Edomex were assigned to Cluster 1 (brown); all individuals from Chihuahua, Sinaloa-MO, Michoacan, Tamaulipas, S.L. Potosi, Guanajuato, and Guerrero-CO were assigned to Cluster 2 (green), and most individuals from the Yucatan Peninsula were assigned to Cluster 3 (purple). (B) Histogram showing the genetic structure within localities from south-east Mexico (Isthmus of Tehuantepec to Yucatan Peninsula) obtained with STRUCTURE for $K = 3$, and fruit morphology of landraces for each locality.

within-individual variation, while grouping by geography (Yucatan Peninsula and Chiapas vs. the rest of samples) explained 3.99% of among-group variation and 64.16% of within-individual variation (Table 3).

On the other hand, the MRM test showed that both mtDNA and nuclear microsatellite loci pairwise genetic distance matrices have a significant association with elevation ($P = 0.004$ and 0.040 , respectively), and a nonsignificant association with geographic distance in any case ($P = 0.446$ and 0.550 , respectively).

DISCUSSION

Contrasting genetic diversity between chloroplast and mitochondria in *C. moschata*

Cucurbita moschata shows high genetic diversity in the mtDNA copy of the *trnL-trnF* region. This result agrees with the findings of Aguirre-Dugua et al. (2019), who report substitution rates (substitutions/site/million years) one to two orders of magnitude larger in mitochondrial copies of chloroplast origin ($rpL120-rpS22 = 0.0177$; $trnL-trnF = 0.0577$) when compared to their chloroplast homologous copies ($rpL120-rpS22 = 0.0032$; $trnL-trnF = 0.0061$).

A high substitution rate in mtDNA copies allows for the accumulation of genetic variation; this is the expected result for a fragment

that has lost its original function, while the lack of genetic variation in analyzed cpDNA regions may be related to a recent origin of this species (see Castellanos-Morales et al., 2018). A mitochondrial copy of *trnL-trnF* was an ideal marker for our study. Nonetheless, we recommend that future *Cucurbita* studies first analyze the identity of amplified sequences to avoid unsuitable comparisons with previous analyses based on this chloroplast region by Barboza et al. (2012) and Zheng et al. (2013).

High genetic variation in mtDNA of *C. moschata*

Mitochondrial haplotype diversity and nucleotide diversity (11 haplotypes, $H_d = 0.589$ and $\pi = 0.0023$) was higher in *C. moschata* than in *C. pepo* ssp. *pepo* (Mexican landraces), based on the same mitochondrial copy of the *trnL-trnF* region (2 haplotypes, $H_d = 0.142$ and $\pi = 0.0010$; Ruiz Mondragón, 2017). Moreover, we found that *C. moschata* in Mexico consists of two mtDNA lineages. Mitochondrial diversity found in our study indicates that Mexico is a region of high genetic diversity for this species. High genetic variation in accessions of *C. moschata* from Mexico has also been reported for chloroplast DNA (12 haplotypes for 38 accessions from Mexico; Barboza et al., 2012), further supporting this country as an area of high diversity for this species.

The high genetic diversity in mtDNA of *C. moschata* in Mexico may be related to several factors, including (1) that it is

a species with a wide distribution, which ranges from the United States to South America (OECD, 2012); (2) that it has a large effective population size, given that *C. moschata* is one of the most widely cultivated species of squash worldwide (Paris, 2016); and (3) that it has a long history since domestication in Mexico, given that the oldest archaeological record of *C. moschata*, ~7000 yr BP, was found in Tehuacán, in the state of Puebla (Whitaker et al., 1957).

High mitochondrial genetic diversity recovered in Veracruz and central Mexico may suggest that Mesoamerica could be one of the centers of domestication and/or diversification of *C. moschata*, as suggested by Whitaker and Davis in 1962. This is further supported by higher genetic variation in chloroplast sequences in Mexico when compared to Central America (Barboza et al., 2012), and by the oldest archaeological remains found relatively near (~200 km) the Veracruz sampled site. Moreover, similar patterns of high genetic variation at the domestication center have been found in other cultivated species in Mexico, such as the chili pepper (Kraft et al., 2014) and the common bean (Kwak et al., 2009). Nevertheless, future studies should include samples from Central and South America to test this idea.

Genetic diversity estimated with nuclear microsatellite loci, gene flow, and inbreeding in *C. moschata*

In terms of nuclear genetic variation, our results showed intermediate levels of allelic richness ($A = 2.96$) compared with other domesticated *Cucurbita* species, based on different sets of nuclear microsatellite loci (*C. maxima*: $A = 4.3$ [Każmińska et al., 2017]; *C. pepo* ssp. *pepo*: $A = 3.43$ – 3.8 [Gong et al., 2008, 2012; Formisano et al., 2012; Enríquez et al., 2018; Castellanos-Morales et al., 2019]; *C. argyrosperma* ssp. *argyrosperma*: $A = 2.7$ [Sánchez-de la Vega et al., 2018]; a hybrid *Cucurbita* rootstock [*C. moschata* × *C. maxima*]: $A = 2.2$ [Kong et al., 2014]). Total genetic variation of *C. moschata* ($H_E = 0.395$) was similar to values reported for sister taxa *C. argyrosperma* ssp. *argyrosperma* and *C. argyrosperma* ssp. *sororia* ($H_E = 0.41$ and $H_E = 0.43$, respectively; Sánchez-de la Vega et al., 2018), but slightly lower than values reported for other annual plants in microsatellite studies (i.e., $H_E = 0.46$; Nybom, 2004) and for other squash crops, for instance *C. pepo* ssp. *pepo* ($H_E = 0.500$ – 0.572 ; Enríquez et al., 2018; Castellanos-Morales et al., 2019), *C. maxima* ($H_E = 0.56$; Kaźmińska et al., 2017), and hybrid *Cucurbita* rootstocks ($H_E = 0.46$; Kong et al., 2014).

Levels of genetic diversity for microsatellite loci of *C. moschata* could be associated with possible gene flow with other squash taxa, given that wild and domesticated gourds often grow in close proximity in traditional agroecosystems in Mesoamerica. Even if the wild ancestor of *C. moschata* is unknown, *C. moschata* produce fertile offspring when crossed with several wild *Cucurbita* species (Whitaker and Bemis, 1964), and the closely related wild taxon *C. argyrosperma* ssp. *sororia* is sympatric in the Pacific coast and Gulf of Mexico. Spontaneous hybridization between *C. moschata*, *C. argyrosperma* ssp. *argyrosperma*, and *C. argyrosperma* ssp. *sororia* has been suggested in the west coast of Mexico (Montes-Hernández and Eguiarte, 2002). In addition, seed exchange and fruit selection within and among localities might also affect levels of nuclear genetic variation in *C. moschata*, as has been reported for *C. argyrosperma* ssp. *argyrosperma* (Sánchez-de la Vega et al., 2018).

Cucurbita moschata in Mexico has relatively high levels of inbreeding (average $F_{IS} = 0.18$) compared to other domesticated

squashes such as *C. argyrosperma* ssp. *argyrosperma* (average $F_{IS} = 0.033$; Sánchez-de la Vega et al., 2018) or *C. pepo* ssp. *pepo* (average $F_{IS} = 0.13$; Castellanos-Morales et al., 2019). However, the inbreeding coefficient in *C. moschata* was lower than that reported for other domesticated plants native to Mesoamerica, such as common bean *Phaseolus vulgaris* L. ($F_{IS} = 0.81$; Blair et al., 2011) and papaya *Carica papaya* ($F_{IS} = 0.58$; Matos et al., 2013). Domestication and crop improvement often involve increased inbreeding (Robinson, 2008; Moyers et al., 2018). For instance, in order to increase crop productivity (by ensuring pollination) or to reduce variability in the crop, farmers sometimes select geitonogamous individuals (Escalante et al., 1994; Montes-Hernández and Eguiarte, 2002), in which flowers may be fertilized with the pollen of a male flower from the same plant (Hessing, 1988). Nevertheless, the effects of inbreeding do not seem to be a crucial problem in domesticated *Cucurbita* species, because spontaneous hybridization and farm management promote genetic diversity (Robinson, 2008; Enríquez et al., 2018).

Stronger signals of genetic structure for mtDNA than for nuclear microsatellite loci in *C. moschata*

The analyzed mitochondrial region (*trnL-trnF*) reflects ancient uniparental history, with a pattern of strong differentiation ($F_{ST} = 0.469$) as a result of lineage divergence. BAPS results ($K = 7$) suggest that genetic structure may be driven by both lineage divergence and elevation, with most of the samples from highland constituting a well-differentiated and cohesive genetic cluster, including few individuals from other locations. By contrast, genetic structure in *C. moschata* for chloroplast DNA was correlated with morphological variation (Barboza et al., 2012). Future analysis should test for correlation between mtDNA genetic structure and morphological variation.

For nuclear microsatellite data, genetic differentiation in *C. moschata* ($F_{ST} = 0.215$) was lower than in *C. argyrosperma* ssp. *argyrosperma* ($F_{ST} = 0.261$; Sánchez-de la Vega et al., 2018), but higher than for *C. pepo* ssp. *pepo* ($F_{ST} = 0.16$; Castellanos-Morales et al., 2019). Our results are consistent with levels of genetic differentiation previously reported for outcrossing plants (average $F_{ST} = 0.22$; Nybom, 2004). Low values of pairwise F_{ST} (0.041–0.340) among localities and the presence of few private alleles suggest that gene flow has been common in this species. For instance, Chiapas, Oaxaca-CU, Nayarit, and Sinaloa show significant signals of admixture (Fig. 5A). Low differentiation among our localities can be attributed to seed exchange among local people and homogeneous morphological selection among farmers within certain geographic regions (see also Montes-Hernández et al., 2005). In addition, wild and cultivated species of *Cucurbita* are pollinated by specialized bees of the genera *Peponapis* and *Xenoglossa* that have limited dispersal ability (~0.7 km) and are found in low densities (Kohn and Casper, 1992; Enríquez et al., 2015). Limited geographic structure and low differentiation among sampled locations have been reported for *C. pepo* ssp. *pepo* (Castellanos-Morales et al., 2019) and for other domesticated species (i.e., beans; Papa and Gepts, 2003). Also, for landraces of relatively recent origin, not enough time may have passed for genetic drift to differentiate the populations.

Results from the STRUCTURE and DAPC analyses for nuclear microsatellite data are consistent. The localities are mainly differentiated between the Yucatan Peninsula landraces (beginning from Isthmus of Tehuantepec to the Yucatan Peninsula) and the

rest of the Mexican landraces (Figs. 4B and 5A). The Isthmus of Tehuantepec seems to act as a partial barrier to gene flow, as has been reported in other studies (see review in Ornelas et al., 2013), at least for nuclear genes. A similar genetic pattern has been described in the closely related domesticated squash *C. argyrosperma* ssp. *argyrosperma* (Sánchez-de la Vega et al., 2018). Although we did not conduct a statistical analysis, because we lack the formal morphological data, we observed that there are morphological differences between the Yucatan Peninsula landraces and the rest of the Mexican landraces (Figs. 1 and 5B). For instance, fruit size may represent an essential character because smaller fruits are preferred within the Yucatan Peninsula and Chiapas (Fig. 5B).

Elevation and geography shape genetic structure in *C. moschata*

We found two main patterns of genetic structure in *C. moschata* depending on the source of molecular data. Genetic differentiation for mtDNA mainly follows an elevation gradient, while genetic differentiation for nuclear microsatellite loci suggests a geographic pattern, in which the Isthmus of Tehuantepec seems to be the main barrier, and there is genetic structure within the two main clusters (Yucatan Peninsula and continental Mexico).

Domestication, selection, and local adaptation may lead to morphological and ecological differentiation among populations (Casas et al., 2007; Paris, 2016). These processes may have promoted the divergence of two main mitochondrial lineages of *C. moschata* in Mexico. One lineage (haplotype A) prevails in lowlands and particularly in localities from the Peninsula of Yucatan, as well as in localities from the south of Mexico. By contrast, the other lineage seems to be associated with highlands in the Trans-Mexican Volcanic Belt (Fig. 3B).

For mtDNA, AMOVA results show high levels of genetic variance explained by the elevational grouping (Table 3). The BAPS analysis also supported two main clusters: rather than a subdivision among geographic areas, these two clusters highlighted a clear altitudinal cline (Appendix S3). As previously mentioned, similar patterns of genetic structure between highland and lowland landraces have been noted in other domesticated taxa (e.g., maize and barley; Tanto et al., 2010; Van Heerwaarden et al., 2011); and genomic analyses have uncovered signals of local adaptation to these altitudinal gradients (Aguirre-Liguori et al., 2017; Fustier et al., 2017). Divergence in highland and lowland, together with human dispersal over large areas (Besnard et al., 2013), could explain the current wide geographic-altitudinal tolerance of *C. moschata* (Lira, 1995). To understand how *C. moschata* adapts to lowland and highland environments, local adaptation should be studied with genomic approaches—for instance, scanning single-nucleotide polymorphisms related to environmental parameters of their distribution, as for example conducted recently in wild maize by Aguirre-Liguori et al. (2017, 2019a, b).

For nuclear microsatellite loci, AMOVA results showed moderate levels of genetic variance explained by geographic grouping, and the Isthmus of Tehuantepec may be acting as a barrier to gene flow (Table 3). Clustering analysis showed structure within the Yucatan Peninsula. As previously mentioned, morphological variation may be related to this pattern of genetic structure (Fig. 5B), as it has been linked to chloroplast genetic variation (Barboza et al., 2012). Sánchez-de la Vega et al. (2018) suggested that genetic structure found in *C. argyrosperma* ssp. *argyrosperma* might be associated

with cultural preferences regarding fruit and seed selection associated with native American groups. In other domesticated species, genetic structure and morphological differentiation could be associated with fruit selection for a particular morphology (see Aguirre-Dugua et al., 2013; Blancas et al., 2016), and also with differences among human cultural management (Casas et al., 1996; Blancas et al., 2016; Sánchez-de la Vega et al., 2018). Future morphological and ethnobotanical analyses should be conducted to corroborate this hypothesis.

GENETIC RESOURCES AND CONCLUSIONS

Landraces of *Cucurbita moschata* in Mexico represent important plant genetic resources with versatile uses as food, important ecological roles, and high genetic diversity.

Divergence of *C. moschata* mtDNA lineages linked to an altitudinal gradient (highland and lowland) could be related to adaptation to different environmental conditions that may also favor its current wide range of cultivation. These lineages may hold the key for adaptation to future environmental conditions and should be preserved in germplasm collections. For instance, in Mexico, we found a variety that is grown in the highest elevation so far reported for this species (i.e., Edomex that was collected above 2800 m). Based on our findings, research should be conducted to identify specific genes or alleles with current or potential value for applications in food, agriculture, and forestry.

Our results indicate that Mexico is an area of high diversity for *C. moschata*. In particular, Veracruz and central Mexico showed high mitochondrial diversity. For future improvement programs for this species, we recommend that landraces from east-central and northeast Mexico (especially Tamaulipas, S.L. Potosi, and Veracruz) should be priorities for conservation. Landraces from the Yucatan Peninsula should also be carefully collected and deposited in germplasm collections for ex situ conservation because of their morphological and genetic uniqueness.

For a thorough understanding of the relationships among *C. moschata* and its sister taxa, it will be necessary to sample and jointly analyze the closely related wild taxon *C. argyrosperma* ssp. *sororia*.

ACKNOWLEDGMENTS

The authors thank CONABIO (project nos. KE004 and PE001 to R.L.-S.S.M.-H., and L.E.E.), Ciencia Básica CONACyT (project no. 2011.167826 to L.E.E.), and Problemas Nacionales CONACyT (project no. 247730 to D. Piñero) for funding this research. This manuscript is presented in partial fulfillment of the requirements to obtain a PhD by H. S. Hernández-Rosales in the Programa de Doctorado en Ciencias Biomédicas, Universidad Nacional Autónoma de México (UNAM). The first author thanks Consejo Nacional de Ciencia y Tecnología (CONACyT) for PhD scholarship (no. 385367). The paper was written during a sabbatical leave of L.E.E. with support of the program PASPA-DGAPA, UNAM, in the Department of Plant and Microbial Biology, University of Minnesota, in P. Tiffin's laboratory. The authors thank V. Souza and D. Piñero for their continuous support during the project. Additionally, we thank two anonymous reviewers for their comments that greatly improved the manuscript.

AUTHOR CONTRIBUTIONS

H.S.H.-R. and G.C.-M. contributed to fieldwork, lab work, molecular and population genetics analysis, and drafting the manuscript. G.S.-V. contributed to fieldwork and lab work. E.A.-P. contributed to correcting the manuscript. S.M.-H. was a project leader and generated the database for project design. R.L.-S. was a project leader and contributed to logistics. L.E.E. was a project leader, designed and coordinated the project and its logistics, and drafted and corrected the manuscript. All authors approved the final version of the manuscript.

DATA AVAILABILITY

The mitochondrial sequences can be accessed through GenBank (NCBI, <https://www.ncbi.nlm.nih.gov/genbank/>) accessions MN381966–MN381976.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

APPENDIX S1. Number of alleles found in landraces of *Cucurbita moschata* per nuclear microsatellites loci (Gong et al., 2008).

APPENDIX S2. Genetic differentiation among localities (pairwise F_{ST}) of *Cucurbita moschata* from Mexico. Lower diagonal pairwise matrix (in bold) corresponds to mitochondrial region (*trnL-trnF*). Upper diagonal pairwise matrix corresponds to nuclear microsatellite loci.

APPENDIX S3. Barplot from BAPS analysis performed with mitochondrial region (*trnL-trnF*) for localities of *Cucurbita moschata* in Mexico. Proportion of individual per locality assigned to each cluster, $K = 7$ (top). Proportion of individual per locality assigned to each cluster, $K = 2$ (bottom). Localities are arranged according to altitude, from lowlands to highlands, depicted with a black line.

APPENDIX S4. Bayesian clustering assignment based on 11 nuclear microsatellite loci amplified for *Cucurbita moschata* in Mexico. (A) Delta K -values obtained by the Evanno's method (Evanno et al., 2005), testing $K = 1$ to $K = 7$. (B) Barplot from STRUCTURE analysis depicting individual's assignment for $K = 2$.

APPENDIX S5. Discriminant analysis of principal components (DAPC) for 24 localities of *Cucurbita moschata* in Mexico. Yucatan Peninsula localities (outlined) are well differentiated from the rest of the samples.

LITERATURE CITED

- Aguilar, J., C. Illsley, and C. Marielle. 2003. Los sistemas agrícolas de maíz y sus procesos técnicos. In G. Esteva and C. Marielle [coords.], Sin maíz no hay país, 83–122. CONACULTA, Mexico City, Mexico.
- Aguirre-Dugua, X., E. Pérez-Negrón, and A. Casas. 2013. Phenotypic differentiation between wild and domesticated varieties of *Crescentia cujete* L. and culturally relevant uses of their fruits as bowls in the Yucatan Peninsula, Mexico. *Journal of Ethnobiology and Ethnomedicine* 9: 76.
- Aguirre-Dugua, X., G. Castellanos-Morales, L. M. Paredes-Torres, H. S. Hernández-Rosales, J. Barrera-Redondo, G. Sánchez-de la Vega, F. Tapia-Aguirre, et al. 2019. Evolutionary dynamics of transferred sequences between organellar genomes in *Cucurbita*. *Journal of Molecular Evolution* 87: 327–342.
- Aguirre-Liguori, J. A., M. I. Tenaillon, A. Vázquez-Lobo, B. S. Gaut, J. P. Jaramillo-Correa, S. Montes-Hernández, V. Souza, and L. E. Eguiarte. 2017. Connecting genomic patterns of local adaptation and niche suitability in teosintes. *Molecular Ecology* 26: 4226–4240.
- Aguirre-Liguori, J. A., B. S. Gaut, J. P. Jaramillo-Correa, M. I. Tenaillon, S. Montes-Hernández, F. García-Oliva, S. J. Hearne, and L. E. Eguiarte. 2019a. Divergence with gene flow is driven by local adaptation to temperature and soil phosphorus concentration in teosinte subspecies (*Zea mays parviglumis* and *Zea mays mexicana*). *Molecular Ecology* 28: 2814–2830.
- Aguirre-Liguori, J. A., S. Ramírez-Barahona, P. Tiffin, and L. E. Eguiarte. 2019b. Climate change is predicted to disrupt patterns of local adaptation in wild and cultivated maize. *Proceedings of the Royal Society B: Biological Sciences* 286: 20190486.
- Altieri, M. A. 1995. *Agroecology: the science of sustainable agriculture*. 2nd ed. Westview Press, Boulder, CO, USA.
- Alverson, A. J., X. Wei, D. W. Rice, D. B. Stern, K. Barry, and J. D. Palmer. 2010. Insights into the evolution of mitochondrial genome size from complete sequences of *Citrullus lanatus* and *Cucurbita pepo* (Cucurbitaceae). *Molecular Biology and Evolution* 27: 1436–1448.
- Balvino-Olvera, F. J., K. F. Sánchez-Gómez, J. A. Lobo, G. Avila-Sakar, R. Cruz-Reyes, G. Sánchez-Montoya, Y. Herrerías-Diego, et al. 2017. Latitudinal structured populations of the Mexican wild squash *Cucurbita argyrosperma* subsp. *sororia* revealed by microsatellite markers. *Crop Pasture Science* 68: 850–858.
- Baranek, M., G. Stift, J. Vollmann, and T. Lelley. 2000. Genetic diversity within and between the species *Cucurbita pepo*, *C. moschata* and *C. maxima* as revealed by RAPD markers. *Cucurbit Genetics Cooperative Report* 23: 73–77.
- Barboza, N., F. J. Albertazzi, J. A. Sibaja-Cordero, F. Mora-Umaña, C. Astorga, and P. Ramírez. 2012. Analysis of genetic diversity of *Cucurbita moschata* (D.) germplasm accessions from Mesoamerica revealed by PCR SSCP and chloroplast sequence data. *Scientia Horticulturae* 134: 60–71.
- Besnard, G., B. Khadari, M. Navascués, M. Fernández-Mazuecos, A. El Bakkali, N. Arrigo, D. Baali-Cherif, et al. 2013. The complex history of the olive tree: from late quaternary diversification of Mediterranean lineages to primary domestication in the northern Levant. *Proceedings of the Royal Society B: Biological Sciences* 280: 2012.2833.
- Blair, M. W., L. M. Díaz, H. R. Gill-Langarica, R. Rosales-Serna, N. Mayek-Perez, and J. A. Acosta-Gallegos. 2011. Genetic relatedness of Mexican common bean cultivars revealed by microsatellite markers. *Crop Science* 51: 2655–2667.
- Blancas, J., A. Casas, A. I. Moreno-Calles, and J. Caballero J. 2016. Cultural motives of plant management and domestication. In R. Lira, A. Casas, and J. Blancas [eds.], *Ethnobotany of Mexico: interactions of people and plants in Mesoamerica*, 233–255. Springer, New York, NY, USA.
- Bukasov, S. M. 1981. *Las plantas cultivadas de México, Guatemala y Colombia*. Centro Agronomía Tropical Investigación y Enseñanza. Turrialba, Costa Rica.
- Casas, A., M. C. Vázquez, J. L. Viveros, and J. Caballero. 1996. Plant management among the Nahua and the Mixtec in the Balsas river basin, Mexico: an ethnobotanical approach to the study of plant domestication. *Human Ecology* 24: 455–479.
- Casas, A., A. Otero-Arnaiz, E. Pérez-Negrón, and A. Valiente-Banuet. 2007. *In situ* management and domestication of plants in Mesoamerica. *Annals of Botany* 100: 1101–1115.
- Castellanos-Morales, G., L. M. Paredes-Torres, N. Gámez, H. S. Hernández-Rosales, G. Sánchez-de la Vega, J. Barrera-Redondo, E. Aguirre-Planter, et al. 2018. Historical biogeography and phylogeny of *Cucurbita*: insights from ancestral area reconstruction and niche evolution. *Molecular Phylogenetics and Evolution* 128: 38–54.
- Castellanos-Morales, G., K. Y. Ruiz-Mondragón, H. S. Hernández-Rosales, G. S. la Vega, E. Aguirre-Planter, S. Montes-Hernández, R. Lira-Saade, and L. E.

- Eguiarte. 2019. Tracing back the origin of pumpkins (*Cucurbita pepo* ssp. *pepo* L.) in Mexico. *Proceedings of the Royal Society B: Biological Sciences* 286: 20191440.
- Cerón González, L., J. P. Legaria Solano, C. Villanueva Verduzco, and J. Sahagún Castellanos. 2010. Diversidad genética en cuatro especies mexicanas de calabaza (*Cucurbita* spp.). *Revista Fitotecnica Mexicana* 33: 189–196.
- Chomicki, G., H. Schaefer, and S. S. Renner. 2019. Origin and domestication of Cucurbitaceae crops: insights from phylogenies, genomics and archaeology. *New Phytologist*. <https://doi.org/10.1111/nph.16015>.
- Corander, J., P. Marttinen, J. Sirén, and J. Tang. 2008. Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. *BMC Bioinformatics* 9: 539.
- Decker-Walters, D. S., and T. W. Walters. 2000. Squash. In K. F. Kiple and K. C. Ornelas [eds.], *The Cambridge World History of Food*. Cambridge University, USA.
- Dillehay, T. D., J. Rossen, T. C. Andres, and D. E. Williams. 2007. Pre-ceramic adoption of peanut, squash, and cotton in Northern Peru. *Science* 316: 1890–1893.
- Doyle, J. J., and J. L. Doyle. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19: 1115.
- Earl, D. A., and B. M. von Holdt. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* 4: 359–361.
- Eguiarte, L. E., H. S. Hernández-Rosales, J. Barrera-Redondo, G. Castellanos-Morales, L. M. Paredes-Torres, G. Sánchez-de la Vega, K. Y. Ruiz-Mondragón, et al. 2018. Domesticación, diversidad y recursos genéticos y genómicos de México: el caso de las calabazas. *TIP Revista Especializada en Ciencias Químico-Biológicas* 21: 85–101.
- Enriquez, E., R. Ayala, V. H. Gonzalez, and J. Núñez-Farfán. 2015. Alpha and beta diversity of bees and their pollination role on *Cucurbita pepo* L. (Cucurbitaceae) in the Guatemalan cloud forest. *The Pan-Pacific Entomologist* 91: 211–222.
- Enriquez, E., P. Landaverde-González, R. Lima-Cordón, E. Solórzano-Ortiz, R. Tapia-López, and J. Nuñez-Farfán. 2018. Population genetics of traditional landraces of *Cucurbita pepo* L., 1753 in the cloud forest in Baja Verapaz, Guatemala. *Genetic Resources and Crop Evolution* 65: 979–991.
- Escalante, A. M., G. Coello, L. E. Eguiarte, and D. Pinero. 1994. Genetic structure and mating systems in wild and cultivated populations of *Phaseolus coccineus* and *P. vulgaris* (Fabaceae). *American Journal of Botany* 81: 1096–1103.
- Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* 14: 2611–2620.
- Ewing, B., L. D. Hillier, M. C. Wendl, and P. Green. 1998. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Research* 8: 175–185.
- Excoffier, L., G. Laval, and S. Schneider. 2005. Arlequin (version 3.5.1.2): An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics* 1: 47–50.
- FAO. Food and Agriculture of the United Nations. 2019. FAOSTAT. Crops. Pumpkins, squash and gourds. Last Update: January 18, 2019. Accessed July 10, 2019. URL <http://www.fao.org/faostat/en/#data/QC/>.
- Ferriol, M., and B. Pico. 2008. Pumpkin and winter squash. In J. Prohens and F. Nuez [eds.], *Handbook of Plant Breeding, vegetables I*. Springer, New York, USA.
- Ferriol, M., B. Picó, and F. Nuez. 2003a. Genetic diversity of a germplasm collection of *Cucurbita pepo* using SRAP and AFLP markers. *Theoretical and Applied Genetics* 107: 271–282.
- Ferriol, M., B. Picó, and F. Nuez. 2003b. Genetic diversity of some accessions of *Cucurbita maxima* from Spain using RAPD and SBAP markers. *Genetic Resources and Crop Evolution* 50: 227–238.
- Ferriol, M., B. Picó, and F. Nuez. 2004a. Morphological and molecular diversity of a collection of *Cucurbita maxima* landraces. *Journal of the American Society for Horticultural Science* 129: 60–69.
- Ferriol, M., B. Pico, P. Fernandez de Cordova, and F. Nuez. 2004b. Molecular diversity of a germplasm collection of squash (*Cucurbita moschata*). *Crop Science* 44: 653–664.
- Fluxus-engineering. 2015. Network v. 5.0.0. Available from: http://www.fluxus-engineering.com/sharenet_rn.htm.
- Formisano, G., C. Roig, C. Esteras, M. R. Ercolano, F. Nuez, A. J. Monforte, and M. B. Picó. 2012. Genetic diversity of Spanish *Cucurbita pepo* landraces: an unexploited resource for summer squash breeding. *Genetic Resources and Crop Evolution* 59: 1169–1184.
- Fustier, M.-A., J.-T. Brandenburg, B. J. Lapeyronnie, L. E. Eguiarte, Y. Vigouroux, D. Manicacci, and M. I. Tenaillon. 2017. Signatures of local adaptation in lowland and highland teosintes from whole-genome sequencing pooled samples. *Molecular Ecology* 26: 2738–2756.
- Gliessman, S. R. 1983. Allelopathic interactions in crop-weed mixtures: applications for weed management. *Management* 9: 991–999.
- Goldman, A. 2004. *The complete squash: a passionate grower's guide to pumpkins, squash, and gourds*. Artisan, NY, USA.
- Gong, L., G. Stift, R. Kofler, M. Pachner, and T. Lelley. 2008. Microsatellites for the genus *Cucurbita* and an SSR-based genetic linkage map of *Cucurbita pepo* L. *Theoretical and Applied Genetics* 117: 37–48.
- Gong, L., H. S. Paris, M. H. Nee, G. Stift, M. Pachner, J. Vollmann, and T. Lelley. 2012. Genetic relationships and evolution in *Cucurbita pepo* (pumpkin, squash, gourd) as revealed by simple sequence repeat polymorphisms. *Theoretical and Applied Genetics* 124: 875–891.
- Gong, L., H. S. Paris, G. Stift, M. Pachner, J. Vollmann, and T. Lelley. 2013. Genetic relationships and evolution in *Cucurbita* as viewed with simple sequence repeat polymorphisms: the centrality of *C. okeechobeensis*. *Genetic Resources and Crop Evolution* 60: 1531–1546.
- Gordon, D., C. Abajian, and P. Green. 1998. Consed: a graphical tool for sequence finishing. *Genome Research* 8: 195–202.
- Goslee, S. C., and D. L. Urban. 2007. The ecodist package for dissimilarity-based analysis of ecological data. *Journal of Statistical Software* 22: 1–19.
- Goudet, J. 2005. Hierfstat, a package for R to compute and test hierarchical F-statistics. *Molecular Ecology Notes* 5: 184–186.
- Hamilton, M. 1999. Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. *Molecular Ecology* 8: 521–523.
- Hessing, M. B. 1988. Geitonogamous pollination and its consequences in *Geranium caespitosum*. *American Journal of Botany* 75: 1324–1333.
- Hurd, P. D., E. G. Linsley, and T. W. Whitaker. 1971. Squash and gourd bees (*Peponapis*, *Xenoglossa*) and the origin of cultivated *Cucurbita*. *Evolution* 25: 218–234.
- Inan, N., M. Yildiz, S. Sensoy, S. Kafkas, and K. Abak. 2012. Efficacy of ISSR and SRAP techniques for molecular characterization of some *Cucurbita* genotypes including naked (hull-less) seed pumpkin. *Journal of Animal and Plant Sciences* 22: 126–136.
- Janes, J. K., J. M. Miller, J. R. Dupuis, R. M. Malefant, J. C. Gorrell, C. I. Cullingham, and R. L. Andrew. 2017. The $K = 2$ conundrum. *Molecular Ecology* 26: 3594–3602.
- Jombart, T. 2008. ADEGENET: A R package for the multivariate analysis of genetic markers. *Bioinformatics* 24: 1403–1405.
- Jombart, T., S. Devillard, and F. Balloux. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics* 11: 94.
- Kalinowski, S. T. 2004. Counting Alleles with rarefaction: private alleles and hierarchical sampling designs. *Conservation Genetics* 5: 539–543.
- Kates, H. R., P. S. Soltis, and D. E. Soltis. 2017. Evolutionary and domestication history of *Cucurbita* (pumpkin and squash) species inferred from 44 nuclear loci. *Molecular Phylogenetics and Evolution* 111: 98–109.
- Każmińska, K., K. Sobieszek, M. Targońska-Karasek, A. Korzeniewska, K. Niemirowicz-Szczytt, and G. Bartoszewski. 2017. Genetic diversity assessment of a winter squash and pumpkin (*Cucurbita maxima* Duchesne) germplasm collection based on genomic *Cucurbita*-conserved SSR markers. *Scientia Horticulturae* 219: 37–44.
- Kohn, J. R., and B. B. Casper. 1992. Pollen-mediated gene flow in *Cucurbita foetidissima* (Cucurbitaceae). *American Journal of Botany* 79: 57–62.
- Kong, Q., J. Chen, Y. Liu, Y. Ma, P. Liu, S. Wu, Y. Huang, and Z. Bie. 2014. Genetic diversity of *Cucurbita* rootstock germplasm as assessed using simple sequence repeat markers. *Scientia Horticulturae* 175: 150–155.

- Kraft, K. H., C. H. Brown, G. P. Nabhan, E. Luedeling, J. d. J. Luna Ruiz, G. Coppens d'Eeckenbrugge, R. J. Hijmans, and P. Gepts. 2014. Multiple lines of evidence for the origin of domesticated chili pepper, *Capsicum annuum*, in Mexico. *Proceedings of the National Academy of Sciences* 111: 6165–6170.
- Kwak, M., J. A. Kami, and P. Gepts. 2009. The putative Mesoamerican domestication center of *Phaseolus vulgaris* is located in the Lerma-Santiago basin of Mexico. *Crop Science* 49: 554–563.
- Librado, P., and J. Rozas. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451–1452.
- Lichstein, J. W. 2007. Multiple regression on distance matrices: a multivariate spatial analysis tool. *Plant Ecology* 188: 177–131.
- Lira, R. 1995. Estudios taxonómicos y ecogeográficos de las Cucurbitaceae latinoamericanas de importancia económica. Systematic and ecogeographic studies on crop gene pools. 9. International Plant Genetic Resources. Rome, Italia.
- Lira, R., L. E. Eguarte, S. Montes, D. Zizumbo-Villarreal, P. C. G. Marín, and M. Quesada. 2016. *Homo sapiens*–*Cucurbita* interaction in Mesoamerica: domestication, dissemination, and diversification. In R. Lira, A. Casas, and J. Blancas [eds.], *Ethnobotany of Mexico: interactions of people and plants in Mesoamerica*, 389–401. Springer, New York, NY, USA.
- Lozada-Aranda, M., I. Rojas Barrera, A. Mastretta Yanes, A. Ponce-Mendoza, C. Burgeff, M. A. Orjuela-R., and O. Oliveros. 2017. Las milpas de México. *Oikos* 17: 10–12.
- Matos, E. L. S., E. J. Oliveira, O. N. Jesus, and J. L. L. Dantas. 2013. Microsatellite markers of genetic diversity and population structure of *Carica papaya*. *Annals of Applied Biology* 163: 298–310.
- Montes-Hernandez, S., and L. E. Eguarte. 2002. Genetic structure and indirect estimates of gene flow in three taxa of *Cucurbita* (Cucurbitaceae) in western Mexico. *American Journal of Botany* 89: 1156–1163.
- Montes-Hernandez, S., L. C. Merrick, and L. E. Eguarte. 2005. Maintenance of squash (*Cucurbita* spp.) landrace diversity by farmers' activities in Mexico. *Genetic Resources and Crop Evolution* 52: 697–707.
- Moyers, B. T., P. L. Morrell, and J. K. McKay. 2018. Genetic costs of domestication and improvement. *Journal of Heredity* 109: 103–116.
- Nee, M. 1990. The domestication of *Cucurbita* (Cucurbitaceae). *Economic Botany* 44: 56–68.
- Nybom, H. 2004. Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Molecular Ecology* 13: 1143–1155.
- Organization for Economic Co-operation and Development [OECD]. 2012. Consensus document on the biology of *Cucurbita* L. (squashes, pumpkins, zucchinis, and gourds). Series on harmonization of regulatory oversight in biotechnology. Paris, France.
- Ornelas, J. F., V. Sosa, D. E. Soltis, J. M. Daza, C. González, P. S. Soltis, C. Gutiérrez-Rodríguez, et al. 2013. Comparative phylogeographic analyses illustrate the complex evolutionary history of threatened cloud forests of northern Mesoamerica. *PLoS ONE* 8: e56283.
- Papa, R., and P. Gepts. 2003. Asymmetry of gene flow and differential geographical structure of molecular diversity in wild and domesticated common bean (*Phaseolus vulgaris* L.) from Mesoamerica. *Theoretical and Applied Genetics* 106: 239–250.
- Paradis, E., J. Claude, and K. Strimmer. 2004. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 20: 289–290.
- Paris, H. S. 2016. Genetic resources of pumpkins and squash *Cucurbita* spp. In R. Grumet, N. Katzir, and J. Garcia-Mas [eds.], *Genetics and genomics of Cucurbitaceae*. Plant genetics and genomics: crops and models, vol 20. Springer, Cham.
- Paris, H. S., N. Yonash, V. Portnoy, N. Mozes-Daube, G. Tzuri, and N. Katzir. 2003. Assessment of genetic relationships in *Cucurbita pepo* (Cucurbitaceae) using DNA markers. *Theoretical and Applied Genetics* 106: 971–978.
- Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959.
- R Core Team. 2013. R: a language and environment for statistical computing. Vienna, Austria. <https://www.R-project.org>.
- Raymond, M., and F. Rousset. 1995. GENEPOP: a population genetics software for exact tests and ecumenicism, version 1.2. *Journal of Heredity* 86: 248–249.
- Robinson, R. W. 2008. Rationale and methods for producing hybrid cucurbit seed. *Journal of New Seeds* 1: 1–47.
- Ruiz Mondragón, K. Y. 2017. Filogeografía y diversidad genética en *Cucurbita pepo* L. 1753 en México. Bachelor degree thesis. Facultad de Ciencias, Universidad Nacional Autónoma de México, México.
- Sánchez-de la Vega, G., G. Castellanos-Morales, N. Gámez, H. S. Hernández-Rosales, A. Vázquez-Lobo, E. Aguirre-Planter, J. P. Jaramillo-Correa, et al. 2018. Genetic resources in the “Calabaza Pipiana” squash (*Cucurbita argyrosperma*) in Mexico: genetic diversity, genetic differentiation and distribution models. *Frontiers in Plant Science* 9: 1–18.
- Sanjurjo, O. I., D. R. Piperno, T. C. Andres, and L. Wessel-Beaver. 2002. Phylogenetic relationships among domesticated and wild species of *Cucurbita* (Cucurbitaceae) inferred from a mitochondrial gene: implications for crop plant evolution and areas of origin. *Proceedings of the National Academy of Sciences* 99: 535–540.
- Shaw, J., E. B. Lickey, E. E. Schilling, and R. L. Small. 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. *American Journal of Botany* 94: 275–288.
- Taberlet, P., L. Gielly, G. Pautou, and J. Bouvet. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology* 17: 1105–1109.
- Tamura, K., D. Peterson, N. Peterson, G. Stecher, M. Nei, and S. Kumar. 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution* 28: 2731–2739.
- Tanto Hadado, T., D. Rau, E. Bitocchi, and R. Papa. 2010. Adaptation and diversity along an altitudinal gradient in Ethiopian barley (*Hordeum vulgare* L.) landraces revealed by molecular analysis. *BMC Plant Biology* 10: 121.
- Van Heerwaarden, J., J. Doebley, W. H. Briggs, J. C. Glaubitz, M. M. Goodman, J. de Jesus Sanchez Gonzalez, and J. Ross-Ibarra. 2011. Genetic signals of origin, spread, and introgression in a large sample of maize landraces. *Proceedings of the National Academy of Sciences* 108: 1088–1092.
- Van Oosterhout, C., W. F. Hutchinson, D. P. M. Wills, and P. Shipley. 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* 4: 535–538.
- Venables, W. N., and B. D. Ripley. 2002. Modern applied statistics with S. Springer, New York.
- Weir, B. S., and C. C. Cockerham. 1984. Estimating *F*-Statistics for the analysis of population structure. *Evolution* 38: 1358.
- Whitaker, T. W. 1947. American origin of the cultivated cucurbits. I. Evidence from the Herbals. II. Survey of old and recent botanical evidence. *Annals of the Missouri Botanical Garden* 34: 101–111.
- Whitaker, T. W. 1968. Ecological aspects of the cultivated *Cucurbita*. *HortScience* 3: 9–11.
- Whitaker, T. W., and W. P. Bemis. 1964. Evolution in the genus *Cucurbita*. *Evolution* 18: 553–559.
- Whitaker, T. W., and H. C. Cutler. 1965. Cucurbits and cultures in the Americas. *Economic Botany* 19: 344–349.
- Whitaker, T. W., and G. N. Davis. 1962. Cucurbits. Botany, cultivation, and utilization. Interscience Publishers Inc. New York, USA.
- Whitaker, T. W., H. C. Cutler, and R. S. Macneish. 1957. Cucurbit materials from 3 caves near Ocampo, Tamaulipas. *American Antiquity* 22: 352–358.
- Whitaker, T. W., and J. B. Bird. 1949. Identification and significance of the Cucurbit materials from Huaca Prieta, Peru. *American Museum Novitates*. Nueva York, USA.
- Wu, J., Z. Chang, Q. Wu, H. Zhan, and S. Xie. 2011. Molecular diversity of Chinese *Cucurbita moschata* germplasm collections detected by AFLP markers. *Scientia Horticulturae* 128: 7–13.
- Zheng, Y. H., A. J. Alverson, Q. F. Wang, and J. D. Palmer. 2013. Chloroplast phylogeny of *Cucurbita*: evolution of the domesticated and wild species. *Journal of Systematics and Evolution* 51: 326–334.

Zizumbo-Villarreal, D., and S. Terán. 1985. Las semillas de la cultura. Los agrosistemas tradicionales una alternativa de desarrollo. *Boletín de la Escuela de Ciencias Antropológicas de la Universidad Autónoma de Yucatán* 72: 3–18.

Zizumbo-Villarreal, D., P. Colunga García-Marín, F. Pat May, J. Martínez Castillo, and J.O. Mijangos Cortés. 2010. Recursos fito-genéticos para la alimentación y la agricultura. In García R. Durán- and M. Méndez-González [eds.], Biodiversidad y desarrollo humano en Yucatán, 334–339. CONABIO, Yucatán, México.

SUPPLEMENTARY MATERIALS

Capítulo 1: GENÉTICA DE POBLACIONES Y FILOGEOGRAFÍA DE *Cucurbita moschata*, Duchesne ex Lam. Duchesne ex Poir. (CUCURBITACEAE) EN MÉXICO.

Artículo de investigación: Hernández-Rosales, H. S., G. Castellanos-Morales, G. Sánchez-de la Vega, E. Aguirre-Planter, S. Montes-Hernández, R. Lira-Saade, y L. E. Eguiarte. 2020. Phylogeographic and population genetic analyses of *Cucurbita moschata* reveal divergence of two mitochondrial lineages linked to an elevational gradient. *American Journal of Botany*. 107: 510-525.

Appendix S1. Number of alleles found in landraces of *Cucurbita moschata* per nuclear microsatellites loci (Gong et al., 2008).

Appendix S2. Genetic differentiation among localities (pairwise F_{st}) of *Cucurbita moschata* from Mexico. Lower diagonal pairwise matrix (in bold) corresponds to mitochondrial region (*trnL-trnF*). Upper diagonal pairwise matrix corresponds to nuclear microsatellite loci.

Appendix S3. Barplot from BAPS analysis performed with mitochondrial region (*trnL-trnF*) for localities of *Cucurbita moschata* in Mexico. Proportion of individual per locality assigned to each Cluster $K = 7$ (top). Proportion of individual per locality assigned to each Cluster $K = 2$ (bottom). Localities are arranged according to altitude from lowlands to highlands depicted with a black line.

Appendix S4. Bayesian clustering assignment based on eleven nuclear microsatellite loci amplified for *Cucurbita moschata* in Mexico. A) Delta K values obtained by the Evanno's method (Evanno et al., 2005), testing $K = 1$ to $K = 7$. B) Barplot from Structure analysis depicting individual's assignment for $K = 2$.

Appendix S5. Discriminant analysis of principal components (DAPC) for 24 localities of *Cucurbita moschata* in Mexico. Yucatan Peninsula localities (outlined) are well differentiated from the rest of the samples.

Appendix S1. Number of alleles found in landraces of *Cucurbita moschata* per nuclear microsatellites loci (Gong et al., 2008).

locus	Expect size (pb)	n. a.
CMTm187	128	20
CMTp175	155	19
CMTm120	121	7
CMTm144	150	16
CMTm261	228	6
CMTm11	108	11
CMTp129	147	4
CMTp193	186	10
CMTm221	108	9
CMTp17	84	1
CMTm54	154	12
CMTp88	167	8
	Total	122

Appendix S2. Genetic differentiation among localities (pairwise F_{ST}) of *Cucurbita moschata* from Mexico. Lower diagonal pairwise matrix (in bold) corresponds to mitochondrial region (*trnL-trnF*). Upper diagonal pairwise matrix corresponds to nuclear microsatellite loci.

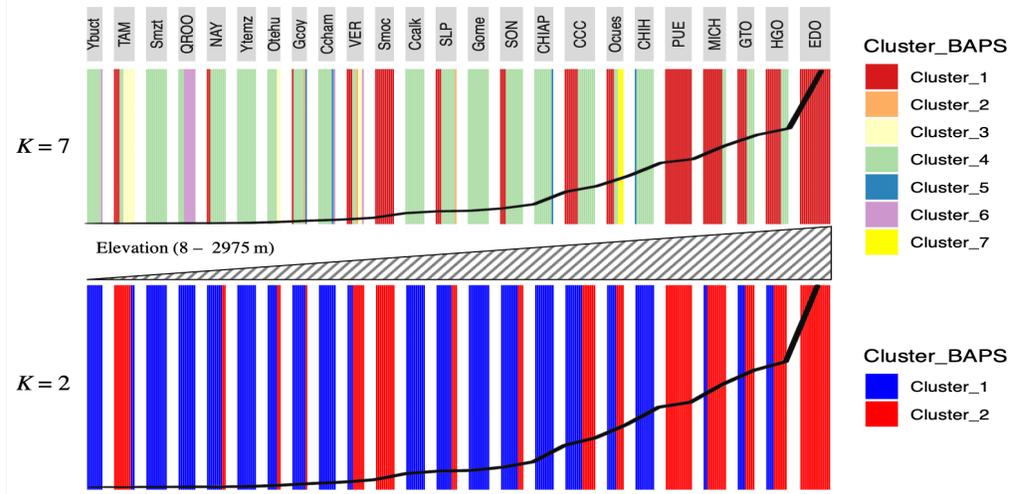
	Sonora	Chihuahua	Coahuila	Sinaloa-MO	Sinaloa-MA	Nayarit
Sonora		0.117	0.154	0.112	0.045	0.068
Chihuahua	0.056		0.171	0.091	0.112	0.114
Coahuila	0.001	0.237		0.188	0.115	0.116
Sinaloa-MO	0.707	0.900	0.469		0.097	0.099
Sinaloa-MA	0.170	0.010	0.347	1.000		0.041
Nayarit	0.000	0.008	0.039	0.778	0.124	
Michoacan	0.466	0.700	0.225	0.070	0.811	0.532
Tamaulipas	0.335	0.481	0.248	0.535	0.600	0.364
S.L.Potosi	0.026	0.192	0.000	0.472	0.317	0.050
Guanajuato	0.092	0.379	0.000	0.394	0.535	0.149
Hidalgo	0.234	0.499	0.030	0.245	0.624	0.295
Edomex	0.761	0.923	0.533	0.000	1.000	0.825
Guerrero-CO	0.000	0.023	0.024	0.667	0.198	0.000
Guerrero-OM	0.170	0.010	0.347	1.000	0.000	0.124
Veracruz	0.150	0.331	0.046	0.352	0.465	0.184
Puebla	0.745	0.917	0.514	0.000	1.000	0.812
Oaxaca-TE	0.052	0.076	0.158	0.802	0.254	0.037
Oaxaca-CU	0.230	0.434	0.095	0.369	0.569	0.271
Chiapas	0.056	0.000	0.237	0.900	0.010	0.008
Campeche-CK	0.009	0.000	0.154	0.802	0.090	0.000
Campeche-CH	0.157	0.000	0.334	1.000	0.000	0.111
Yucatan-BU	0.056	0.000	0.237	0.900	0.010	0.008
Yucatan-TE	0.157	0.000	0.334	1.000	0.000	0.111
Quintana-Roo	0.399	0.508	0.367	0.762	0.655	0.420

	Michoacan	Tamaulipas	S.L.Potosi	Guanajuato	Hidalgo
Sonora	0.12	0.144	0.098	0.107	0.131
Chihuahua	0.164	0.12	0.1	0.119	0.142
Coahuila	0.224	0.156	0.124	0.174	0.154
Sinaloa-MO	0.119	0.063	0.063	0.111	0.093
Sinaloa-MA	0.102	0.122	0.079	0.105	0.119
Nayarit	0.096	0.107	0.079	0.107	0.097
Michoacan		0.162	0.117	0.113	0.195
Tamaulipas	0.360		0.066	0.131	0.115
S.L.Potosi	0.254	0.170		0.085	0.11
Guanajuato	0.084	0.209	0.000		0.17
Hidalgo	0.000	0.250	0.079	0.000	
Edomex	0.126	0.610	0.551	0.483	0.319
Guerrero-CO	0.428	0.238	0.000	0.083	0.213
Guerrero-OM	0.811	0.600	0.317	0.535	0.624
Veracruz	0.151	0.049	0.000	0.000	0.034
Puebla	0.109	0.588	0.528	0.457	0.297
Oaxaca-TE	0.577	0.338	0.081	0.241	0.368
Oaxaca-CU	0.152	0.180	0.056	0.024	0.048
Chiapas	0.700	0.481	0.192	0.379	0.499
Campeche-CK	0.593	0.369	0.094	0.257	0.388
Campeche-CH	0.803	0.586	0.301	0.518	0.612
Yucatan-BU	0.700	0.481	0.192	0.379	0.499
Yucatan-TE	0.803	0.586	0.301	0.518	0.612
Quintana-Roo	0.586	0.381	0.260	0.380	0.447

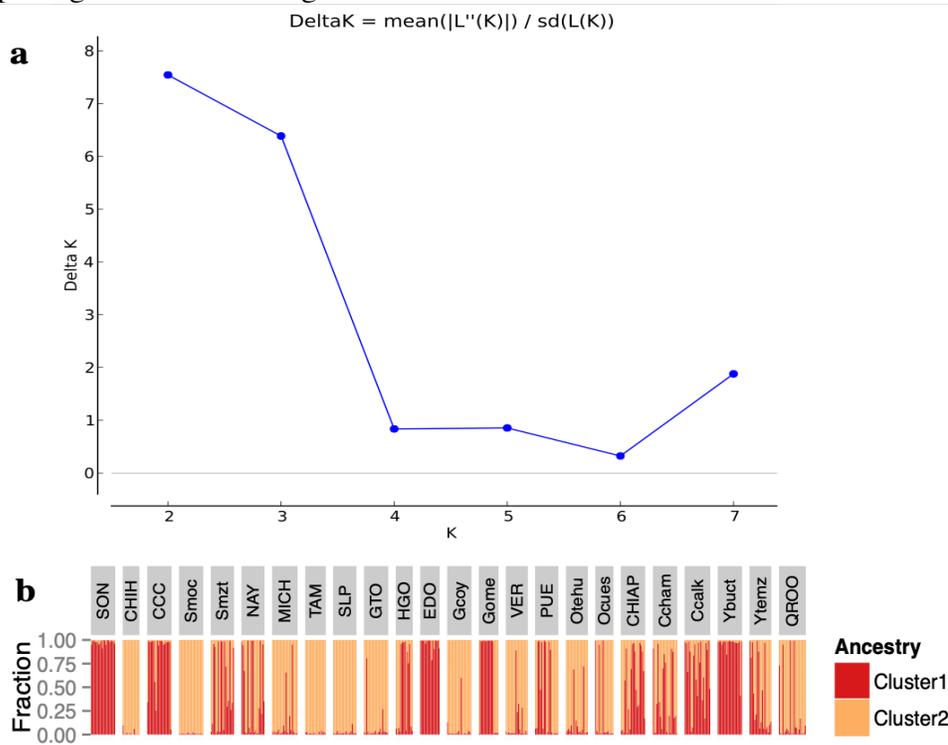
	Edomex	Guerrero- CO	Guerrero-OM	Veracruz	Puebla	Oaxaca-TE
Sonora	0.14	0.156	0.047	0.129	0.09	0.095
Chihuahua	0.219	0.196	0.132	0.136	0.12	0.106
Coahuila	0.166	0.26	0.131	0.137	0.12	0.173
Sinaloa-MO	0.229	0.183	0.139	0.117	0.08	0.051
Sinaloa-MA	0.149	0.202	0.078	0.122	0.07	0.088
Nayarit	0.154	0.194	0.09	0.1	0.04	0.082
Michoacan	0.258	0.25	0.173	0.15	0.13	0.133
Tamaulipas	0.248	0.197	0.157	0.086	0.09	0.079
S.L.Potosi	0.227	0.14	0.147	0.069	0.08	0.067
Guanajuato	0.209	0.18	0.157	0.12	0.1	0.11
Hidalgo	0.198	0.247	0.113	0.133	0.06	0.093
Edomex		0.323	0.108	0.236	0.14	0.231
Guerrero-CO	0.737		0.223	0.185	0.22	0.164
Guerrero-OM	1.000	0.198		0.169	0.09	0.129
Veracruz	0.442	0.047	0.465		0.1	0.092
Puebla	0.000	0.718	1.000	0.415		0.066
Oaxaca-TE	0.852	0.000	0.254	0.176	0.838	
Oaxaca-CU	0.459	0.152	0.569	0.000	0.432	0.278
Chiapas	0.923	0.023	0.010	0.331	0.917	0.076
Campeche-CK	0.847	0.000	0.090	0.211	0.834	0.000
Campeche-CH	1.000	0.181	0.000	0.447	1.000	0.235
Yucatan-BU	0.923	0.053	0.010	0.331	0.917	0.076
Yucatan-TE	1.000	0.181	0.000	0.447	1.000	0.235
Quintana-Roo	0.814	0.281	0.655	0.185	0.800	0.358

	Oaxaca- CU	Chiapas	Campeche CK	Campeche CH	Yucatan BU	Yucatan TE	Quintana-Roo
Sonora	0.092	0.112	0.109	0.125	0.194	0.109	0.137
Chihuahua	0.121	0.123	0.139	0.136	0.217	0.122	0.137
Coahuila	0.138	0.16	0.191	0.129	0.252	0.17	0.195
Sinaloa-MO	0.098	0.107	0.101	0.127	0.228	0.109	0.123
Sinaloa-MA	0.052	0.108	0.079	0.1	0.15	0.078	0.12
Nayarit	0.05	0.104	0.08	0.079	0.157	0.062	0.101
Michoacan	0.106	0.145	0.151	0.177	0.273	0.115	0.172
Tamaulipas	0.115	0.105	0.134	0.115	0.284	0.14	0.129
S.L.Potosi	0.061	0.107	0.096	0.109	0.222	0.099	0.098
Guanajuato	0.111	0.114	0.15	0.154	0.254	0.117	0.156
Hidalgo	0.129	0.144	0.108	0.115	0.233	0.141	0.129
Edomex	0.243	0.191	0.238	0.19	0.299	0.203	0.244
Guerrero- CO	0.188	0.193	0.187	0.201	0.34	0.202	0.193
Guerrero- OM	0.151	0.125	0.146	0.125	0.21	0.145	0.181
Veracruz	0.089	0.13	0.134	0.121	0.266	0.13	0.121
Puebla	0.076	0.082	0.081	0.087	0.197	0.078	0.092
Oaxaca-TE	0.081	0.07	0.072	0.095	0.208	0.091	0.097
Oaxaca-CU		0.121	0.07	0.103	0.179	0.07	0.088
Chiapas	0.434		0.102	0.077	0.253	0.086	0.107
Campeche- CK	0.312	0.000		0.066	0.141	0.053	0.06
Campeche- CH	0.553	0.000	0.077		0.155	0.073	0.087
Yucatan-BU	0.434	0.000	0.000	0.000		0.13	0.222
Yucatan-TE	0.553	0.000	0.077	0.000	0.000		0.084
Quintana- Roo	0.340	0.508	0.381	0.641	0.508	0.641	

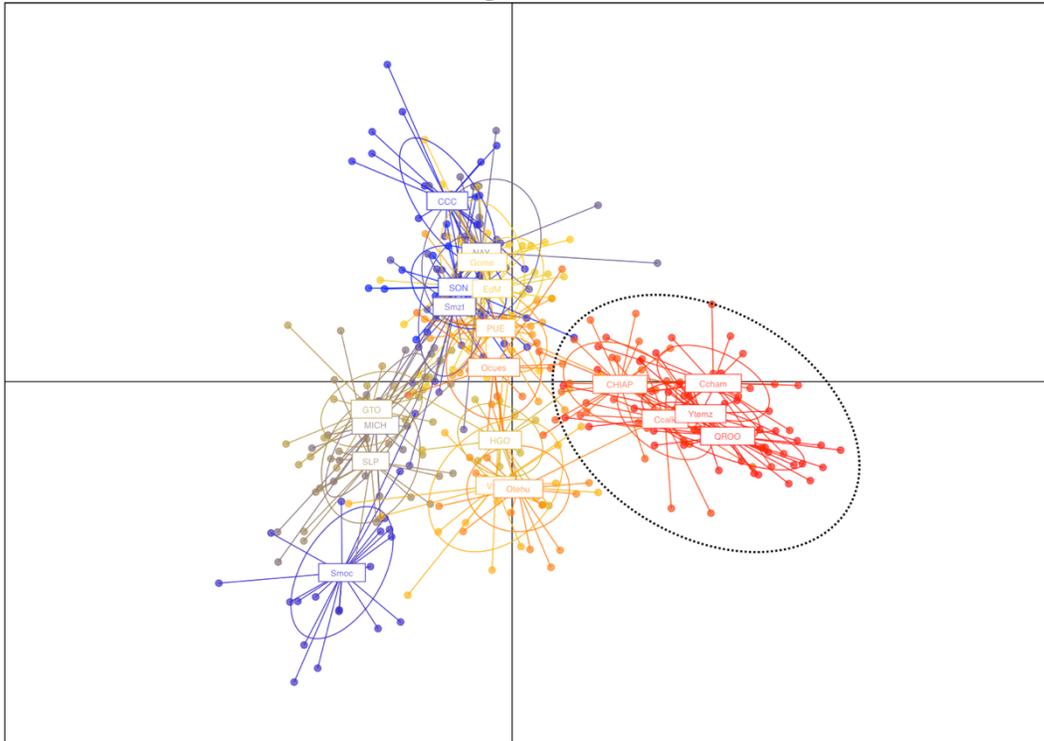
Appendix S3. Barplot from BAPS analysis performed with mitochondrial region (*trnL*–*trnF*) for localities of *Cucurbita moschata* in Mexico. Proportion of individual per locality assigned to each Cluster $K = 7$ (top). Proportion of individual per locality assigned to each Cluster $K = 2$ (bottom). Localities are arranged according to altitude from lowlands to highlands depicted with a black line.



Appendix S4. Bayesian clustering assignment based on eleven nuclear microsatellite loci amplified for *Cucurbita moschata* in Mexico. A) Delta K values obtained by the Evanno's method (Evanno et al., 2005), testing $K = 1$ to $K = 7$. B) Barplot from Structure analysis depicting individual's assignment for $K = 2$.



Appendix S5. Discriminant analysis of principal components (DAPC) for 24 localities of *Cucurbita moschata* in Mexico. Yucatan Peninsula localities (outlined) are well differentiated from the rest of the samples.



Capítulo 2: GENÉTICA DE POBLACIONES COMPARADA Y ANÁLISIS DE FLUJO GENÉTICO DE ESPECIES DOMESTICADAS DEL GÉNERO *Cucurbita*.

Artículo en preparación: Juntas, pero no revueltas: Diferenciación genética y tasas de migración de tres especies domesticadas del género *Cucurbita* (CUCURBITACEAE) en México.

Este manuscrito está estructurado de acuerdo con las directrices de la revista mexicana *Botanical Sciences* y se plantea enviarse una vez que se hayan realizado las correcciones necesarias y se cuente con la aprobación de cada uno de los coautores. El objetivo fue estimar el flujo genético entre especies domesticadas del género *Cucurbita*. Se recopilaron, estandarizaron y utilizaron datos de microsatélites nucleares de calabazas domesticadas en México: *C. argyrosperma* (Vega et al., 2018), *C. pepo* (Castellanos-Morales et al., 2019) y *C. moschata* (Hernández-Rosales et al., 2020). Se identificaron cinco loci de microsatélites nucleares que son consistentes entre las tres especies. Se obtuvieron los estadísticos clásicos de diversidad genética (P , A , H_O , H_E , F_{IS}). Los valores de diferenciación y estructura genética se calcularon con la distancia genética de Nei y se realizó un análisis de varianza molecular (AMOVA). Finalmente, para la estimación de tasas de flujo genético actual se consideraron dos modelos (global y local) y se utilizó el programa BayesAss. Los valores de diversidad genética son congruentes con los reportes previos para cada especie. Las especies se encuentran genéticamente bien diferenciadas. Las tasas de flujo genético son bajas a nivel global y local entre las especies analizadas. El flujo genético no es un evento significativo aún cuando los agricultores cultivan en conjunto distintas especies de calabaza dentro de la misma milpa, por lo que el acervo genético de cada especie depositado en sus variedades locales y razas criollas se conserva. No se descarta el uso de cruza interespecífica con fines de mejoramiento genético, sin embargo, la reproducción en el futuro tiene que ser un proceso controlado considerando las bases genéticas.

Artículo de investigación.

Juntas, pero no revueltas: Diferenciación genética y tasas de migración de tres especies domesticadas del género *Cucurbita* (Cucurbitaceae) en México.

Helena S. Hernández-Rosales¹, Gabriela Castellanos-Morales², Guillermo Sánchez-de la Vega¹, Erika Aguirre-Planter¹, Daniel Piñero³, Rafael Lira-Saade^{4, 5} y Luis E. Eguiarte^{1,5}.

¹Laboratorio de Evolución Molecular y Experimental, Departamento de Ecología Evolutiva, Instituto de Ecología, Universidad Nacional Autónoma de México, Circuito Exterior s/n, Ciudad de México 04510, México.

²Departamento de Conservación de la Biodiversidad, El Colegio de la Frontera Sur, Unidad Villahermosa, Villahermosa, Tabasco, México.

³Departamento de Ecología Evolutiva, Instituto de Ecología, Universidad Nacional Autónoma de México.

⁴ UBIPRO, Facultad de Estudios Superiores Iztacala, Universidad Nacional Autónoma de México, Av. de los Barrios #1, Col. Los Reyes Iztacala, Tlalnepantla, Edo. de Mex, 54090, México

⁵Autores de correspondencia: rlira@unam.mx, fruns@unam.mx

Resumen

Antecedentes: El uso de marcadores moleculares permite evaluar las tasas de migración entre especies y entre poblaciones. Dentro de los sistemas tradicionales de cultivo en México, es común encontrar más de una especie domesticada de *Cucurbita* creciendo en la misma temporada de cultivo. Se reconoce la existencia de flujo genético entre calabazas domesticadas, sin embargo, pocos estudios han estimado las tasas de migración y la dirección del flujo genético entre estas.

Preguntas: ¿Cuáles son los niveles de diferenciación entre las especies domesticadas de *Cucurbita*? ¿Cómo es el flujo genético a nivel global y a nivel local entre especies domesticadas de *Cucurbita* en México, y en qué dirección se dan con mayor intensidad?

Especies estudiadas: *Cucurbita argyrosperma* subesp. *argyrosperma* L., *Cucurbita moschata* Duchesne, *Cucurbita pepo* subesp. *pepo* L.

Métodos: Se estudió la diversidad genética de cinco loci de microsatélites nucleares de tres especies de *Cucurbita* domesticadas. La diferenciación genética entre taxa y a nivel local se evaluó con la medida de G_{ST} y se implementó un análisis de estructura genética. Se estimaron las tasas de migración reciente entre taxa y a nivel local implementando inferencia bayesiana con el programa BayesAss.

Resultados: Los niveles de diversidad genética son relativamente altos en todas las especies ($H_E = 0.357-0.566$). Las tres especies se encuentran genéticamente bien diferenciadas ($G_{ST} = 0.457-0.508$), con señales de introgresión limitadas. Las tasas de flujo genético entre las especies son bajas en ambas direcciones tanto a nivel global ($m < 0.0161$) y como a nivel local ($m < 0.0186$).

Conclusión: Aunque las especies están aisladas genéticamente unas de otras, la presencia de híbridos espontáneos apoya la posibilidad de llevar a cabo fertilización artificial con fines de mejoramiento. El flujo genético entre especies, aunque moderado aparentemente ha permitido conservar los acervos genéticos representados por las razas y variedades locales. Las barreras naturales que evitan la introgresión intraespecífica entre especies de calabazas podrían reducir la posible contaminación de los transgenes de los cultivos mejorados hacia las especies silvestres y variedades locales.

1 Un organismo híbrido generalmente se define como la consecuencia de la fertilización
2 cruzada entre individuos que provienen de poblaciones o especies distintas y se pueden
3 distinguir en función de uno o más caracteres heredables (Harrison, 1990) aunque en
4 algunos casos la única evidencia de la presencia de híbridos es a partir de individuos con
5 características intermedias (Ellstrand et al., 1996). La hibridación artificial ha sido
6 importante en el mejoramiento de cultivos y por ello ha sido estudiada con mayor
7 frecuencia en plantas que en animales desde el inicio de la domesticación (Rieseberg,
8 1997). La hibridación natural ha sido importante en la historia evolutiva de muchas plantas,
9 aunque es más rara en animales (Dowling and Secor, 1997). El término introgresión se
10 refiere a la transferencia de genes entre poblaciones distintas (Rieseberg, 1997). Una de las
11 consecuencias evolutivas de la hibridación e introgresión es el incremento de la diversidad
12 genética (González-Rodríguez, 2011). Detectar el grado de introgresión y la dirección del
13 flujo genético entre cultivos de distintas especies que coexisten muy cercanamente, es
14 fundamental para el manejo y mejoramiento de cultivos y para la conservación de los
15 recursos fitogenéticos (Cerrato et al., 2020).

16 Cuando la hibridación o introgresión suceden a bajas frecuencias en las poblaciones
17 naturales, la única forma práctica de detectar estos eventos es a partir del análisis de la
18 variación y estructura genética con el uso de marcadores moleculares (González-Rodríguez,
19 2011). Una manera popular de inferir el grado de introgresión en términos de flujo genético
20 es a partir de la medida de diferenciación genética F_{ST} de Wright (Wright, 1965) y/o
21 estimando la tasa de migración (m). La F_{ST} evalúa las diferencias las frecuencias alélicas
22 entre poblaciones (y por tanto la diferenciación genética entre ellas). La tasa de migración
23 (m) es la fracción de individuos de una población que migra (o transfiere sus genes) de otra

24 población y se puede usar para indicar la tasa de flujo genético en modelos de estructura
25 poblacional (Hamilton, 2009). La tasa de migración y la F_{ST} se relacionan ya que la tasa a
26 la que un alelo se fija en una población puede ser contrarrestada por el flujo genético que
27 llega de otras poblaciones a una tasa m (Wright, 1969). Estos factores, en genomas
28 diploides, llegan a un equilibrio, aproximando, $F_{ST} \approx \frac{1}{4} (4 Nm + 1)$. Entonces, cuando la
29 $F_{ST} = 1$ es el máximo de la varianza en las frecuencias de alelos (no se comparten alelos
30 entre poblaciones), no hay flujo genético y las poblaciones son genéticamente totalmente
31 distintas. Por otro lado, cuando la $F_{ST} = 0$ significa que no hay varianza en la frecuencia de
32 alelo (todas las poblaciones tienen exactamente las mismas frecuencias alélicas) o sea, las
33 poblaciones son idénticas u homogéneas y el flujo genético es constante. En términos
34 generales, si $F_{ST} \neq 0$ se sugiere que hay flujo genético. Sin embargo, una desventaja de esta
35 medida es que no permite detectar la dirección de migración de alelos entre poblaciones.

36 Entre los cultivos más valorados a nivel mundial se encuentran las especies de
37 plantas domesticadas de la familia Cucurbitaceae que incluye, por mencionar algunos, al
38 melón (*Cucumis melo* L.), a la sandía (*Citrullus lanatus* Thunb.) y a las calabazas
39 (*Cucurbita* L. spp.) (Chomicki et al., 2019). El género *Cucurbita* se considera como uno de
40 los géneros con mayor diversidad morfológica (especialmente en sus frutos) dentro de las
41 angiospermas (Bisognin, 2002) y comprende alrededor de 20 taxa nativos de América y la
42 mayor parte se distribuyen en México (Paris, 2016). Este género destaca entre las
43 cucurbitáceas por incluir el mayor número de especies domesticadas: *C. pepo* L., *C.*
44 *argyrosperma* K. Koch subesp. *argyrosperma*, *C. maxima* Duchesne, *C. ficifolia* Bouché
45 y *C. moschata* Duchesne (Paris, 2016; Castellanos-Morales et al., 2018).

46 En México, las calabazas se cultivan tradicionalmente en los sistemas agrícolas
47 tradicionales de origen Mesoamericano denominados “milpas” (Hernández Xolocotzi,
48 1988). Estas calabazas se siembran en asociación con maíz (*Zea mays* L.) principalmente,
49 pero también pueden estar acompañadas de frijoles (*Phaseolus* L. spp.) y chiles (*Capsicum*
50 L.), entre otras especies de plantas (Lozada-Aranda et al. 2017). De la misma manera, es
51 común que dentro del mismo agroecosistema se siembren variedades locales de más de una
52 especie de calabaza, incluso en la misma temporada de cultivo (p. ej., *C. argyrosperma*
53 subesp. *argyrosperma*–*C. moschata*; *C. pepo* subesp. *pepo*–*C. moschata*) (Montes-
54 Hernandez and Eguiarte, 2002; Montes-Hernández et al., 2005).

55 Todas las especies de calabazas, incluyendo los taxa silvestres, son diploides y
56 cuentan con 20 pares de cromosomas ($2n = 40$) (Whitaker and Davis, 1962). Son plantas
57 monoicas auto-compatibles, y se considera que son polinizadas principalmente por especies
58 de abejas especialistas, pertenecientes principalmente a los géneros *Peponapis* spp.
59 y *Xenoglossa* spp. (Hurd et al., 1971). También puede ser polinizadas por otras abejas,
60 incluyendo las *Apis mellifera* (Lira, 1995; Lira et al., 2016). En general las especies de
61 calabazas están aisladas reproductivamente unas de las otras (Paris, 2016). Sin embargo, la
62 relación de compatibilidad genética entre los cinco taxa domesticados ha sido estudiado
63 ampliamente (Whitaker, 1951; Merrick, 1990; Lira, 1995) y se ha demostrado la presencia
64 de híbridos potencialmente fértiles (Whitaker and Bemis, 1964; Cruz-Reyes et al., 2015).

65 En México se han reportado cruzas espontáneas entre variedades locales de distintas
66 especies de calabazas domesticadas (Decker-Walters et al., 1990). Los experimentos de
67 cruzas artificiales indican que existe elevada probabilidad de obtener individuos híbridos y
68 que el éxito de las cruzas dependerá en gran medida del cruce de los genotipos parentales y

69 la porción de la variación (cultivares, razas o variedades regionales, ecotipos de las especies
70 silvestres, etc.) que se utilice de cada una de ellas (Lira, 1995; OECD, 2012). El estudio de
71 cruza confirma la estrecha relación entre las especies domesticadas *C. moschata*, *C.*
72 *argyrosperma* y revalida la idea de un ancestro en común con *C. pepo* (Lira, 1995).

73 La obtención de híbridos es un mecanismo importante para el mejoramiento de
74 cultivos enfocado en producción y calidad de frutos. Sin embargo, uno de los principales
75 objetivos de estos experimentos y como ocurre con muchos otros géneros que incluyen
76 especies cultivadas, es la posibilidad de transferir resistencia a enfermedades a las especies
77 domesticadas (Lira, 1995). De hecho, ya se han generado algunos híbridos con este fin. Por
78 ejemplo, se ha transferido la resistencia al *powdery mildew* –una enfermedad causada por
79 hongos– a *C. pepo* subesp. *pepo* desde las especies silvestres *C. lundelliana* y *C.*
80 *okeechobeensis* subesp. *martinezii* empleando a *C. moschata* como intermediario (Paris and
81 Cohen, 2002), entre otras hibridaciones (Maluf et al., 1997; Matos et al., 2013).

82 Pocos estudios han estimado el flujo genético entre distintas especies domesticadas
83 de calabaza en México dentro de los sistemas tradicionales de policultivos (Montes-
84 Hernandez and Eguiarte, 2002; González et al., 2010). La dirección y la intensidad de
85 migración entre taxa, no solo en calabazas, sino en otras especies dentro de sistemas
86 agrícolas tradicionales también ha sido poco estudiada. Los objetivos del presente trabajo
87 son: (1) Analizar de manera integral los niveles de diversidad y estructura genética con
88 datos de microsatélites nucleares de tres especies domesticadas de *Cucurbita* (*C. pepo*
89 subesp. *pepo*, *C. argyrosperma* subesp. *argyrosperma* y *C. moschata*), cultivadas en todo
90 México. (2) Estimar la diferenciación genética entre taxa en dos niveles: global y local y (3)
91 Evaluar la tasa y dirección de migración entre especies, así como a nivel local.

92 **Materiales y métodos.**

93 *Especies estudiadas.* De las cinco especies de calabazas domesticadas, *C. pepo* subesp.
94 *pepo*, *C. argyrosperma* subesp. *argyrosperma* y *C. moschata* sobresalen por su importancia
95 económica y alimenticia en México (Montes-Hernández et al., 2005; Chomicki et al.,
96 2019). La evidencia arqueobotánica, junto con diversos estudios moleculares y la amplia
97 diversidad de variedades locales indican que México es centro de domesticación y/o
98 diversificación de estas especies (Whitaker et al., 1957; Smith, 1997; Kates et al., 2017;
99 Castellanos-Morales et al., 2018, 2019; Sánchez-de La Vega et al., 2018; Hernández-
100 Rosales et al., 2020).

101 *Cucurbita argyrosperma* subesp. *argyrosperma*, comúnmente conocida como
102 calabaza pipiana en México, es apreciada por el uso de sus semillas en la gastronomía
103 mexicana (Lira, 1995; Villanueva, 2007). La distribución de este taxa comprende desde el
104 suroeste de Estados Unidos, casi todo México hasta Centro América (OECD, 2012).
105 Contrario a otras calabazas domesticadas, su cultivo fuera de América es limitado, por lo
106 que se considera que su producción se da principalmente a nivel local en cultivos de
107 temporada (Lira, 1995). La variación más importante que se observa en esta especie se
108 encuentra en la gran diversidad en tamaño, forma y color de sus semillas entre los cultivos
109 locales (Sánchez-de la Vega et al., 2018). Esta especie es importante en la agricultura local
110 de México, donde las variedades locales se seleccionan y conservan de manera tradicional
111 (Lira, 1995). Lo frutos de esta especie, en contraste con el resto de las especies
112 domesticadas, son poco apreciados por su pulpa, lo que limita su cultivo a nivel global. Sin
113 embargo, en los Estados Unidos se han desarrollado algunos cultivares comerciales
114 (OECD, 2012).

115 *Cucurbita moschata*, mejor conocida como calabaza de castilla es la segunda
116 especie con mayor diversidad morfológica del género (después de *C. pepo*) y es cultivada
117 ampliamente en distintas regiones a nivel mundial (OECD, 2012; Paris, 2016). El amplio
118 rango de distribución de esta especie y su adaptación a distintas condiciones ecológicas se
119 debe a la exitosa historia evolutiva de sus cultivares (Lira, 1995; Hernández-Rosales et al.,
120 2019). En Mesoamérica, sus variedades y razas locales son comúnmente cultivadas, en
121 asociación con otras especies de calabazas en las milpas, especialmente junto con su
122 especie cercana *C. argyrosperma* subesp. *argyrosperma* (Lira, 1995). En México representa
123 un importante recurso fitogenético, debido a la gran diversidad de razas locales existente en
124 esta área y al extenso uso de sus flores, frutos y semillas (Paris, 2016; Hernández-Rosales,
125 et al., 2020).

126 *Cucurbita pepo* subesp. *pepo* es la especie con mayor diversidad morfológica y
127 número de variedades comerciales desarrolladas a nivel global para su consumo como
128 verdura y dada su gran importancia económica a nivel mundial, es la especie más estudiada
129 del género (Paris, 2016; Eguiarte et al., 2018; Castellanos-Morales et al., 2019; Chomicki et
130 al., 2019). Los restos arqueológicos más antiguos (~9,500 años) han sido encontrados en
131 México y los estudios moleculares más recientes apuntan a esta región como el centro de
132 domesticación de esta especie (Castellanos-Morales et al., 2019). En México es común
133 encontrarlas en monocultivos extensivos, sin embargo, su cultivo también se mantiene en
134 agroecosistemas tradicionales, y se presenta una gran diversidad entre sus variedades
135 locales (Lira, 1995).

136

137 **Obtención de los datos**

138 Se analizaron bases de frecuencias alélicas estimadas a partir de marcadores de
139 microsatélites nucleares de las tres especies de estudio (Sánchez-de la Vega et al., 2018;
140 Castellanos-Morales et al., 2019; Hernández-Rosales et al., 2020) (Fig. 1; Apéndice 1).
141 Para esto, se estandarizó la lectura del tamaño alélico de los microsatélites en las tres bases
142 de datos a partir de sus electroferogramas. Estos datos provienen del Laboratorio de
143 Evolución Molecular y Experimental del Instituto de Ecología de la Universidad Nacional
144 Autónoma de México (UNAM) y forman parte del proyecto CONABIO núm. KE004
145 “Diversidad genética de las especies de *Cucurbita* en México e hibridación entre plantas
146 genéticamente modificadas y especies silvestres de *Cucurbita*”.

147 **Estimación de frecuencias alélicas y análisis de variación genética.**

148 Para evaluar los niveles de diversidad y los patrones de estructura genética global de las
149 tres especies de estudio: (1) se eligieron solo aquellas poblaciones muestreadas de cada
150 especie que correspondieran a variedades locales y/o razas criollas; (2) todas las localidades
151 muestreadas de cada especie se consideraron como un solo grupo de datos y (3) se
152 seleccionaron aquellos loci de microsatélites nucleares que fueran compartidos entre las
153 tres especies (Figura 1; Apéndice 1).

154 Las frecuencias alélicas fueron estimadas por conteo directo. La prueba exacta de
155 equilibrio Hardy-Weinberg de cada locus se llevó a cabo con el programa GenoDive
156 versión 3.01 (Meirmans and Van Tienderen, 2004). También con este programa se
157 obtuvieron los estadísticos básicos: heterocigosis esperada (H_E), heterocigosis observada
158 (H_O), coeficiente de endogamia (G_{IS}) y número efectivo de alelos (A). El número de alelos

159 privados por especie se calculó con la librería de R PopGenReport (Adamack and Gruber,
160 2014).

161

162 **Diferenciación y estructura genética.**

163 Se determinó la diferenciación genética entre taxa con la distancia genética de Nei (D , Nei,
164 1978) con el programa GenoDive. También con GenoDive, se realizó un análisis de
165 varianza molecular (AMOVA) para determinar los porcentajes de variación dentro y entre
166 taxa. Se evaluó la estructura genética entre y dentro de cada taxón con el programa
167 Structure versión 2.2 (Pritchard et al., 2000). Considerando la ancestría en común que
168 comparten las tres especies y los reportes de fertilización cruzada entre taxa, se utilizó un
169 modelo de frecuencias alélicas independientes y un modelo de ancestría de *admixture*
170 (mezcla). La prueba final se llevó a cabo con una longitud de 500,000 *burn-in* y 1 millón de
171 Cadenas de Markov Monte Carlo (MCMC). Se probaron de 1–12 grupos genéticos (K) con
172 15 repeticiones para cada uno. Para complementar la identificación de grupos genéticos que
173 reflejan los datos se realizó un análisis discriminante de componentes principales (DAPC)
174 (Jombart et al., 2010) implementado en la librería de R “adegenet” (Jombart, 2008).

175

176 **Estimación de tasas de flujo genético interespecífico reciente.**

177 Se estimaron las tasas de migración, considerando dos modelos: global y local. La tasa de
178 migración global se estimó entre las tres especies de estudio. Para estimar las tasas de
179 migración a nivel local, se eligieron cinco pares de localidades en las que se colectaron en

180 la misma temporada de siembra dentro de la misma milpa a *C. argyrosperma* subesp.
181 *argyrosperma* y *C. moschata* (Tabla 2). Se utilizó el programa BayesAss versión 1.3
182 (Wilson and Rannala, 2003). Este método Bayesiano calcula la probabilidad posterior de la
183 ancestría migrante de cada individuo, la frecuencia alélica y el coeficiente de endogamia de
184 cada población a partir de repeticiones de MCMC. Para cada grupo de datos se corrieron en
185 total 100×10^6 iteraciones de MCMC y se descartaron las primeras 10×10^6 iteraciones
186 como *burn-in*. Para determinar adecuadamente los valores de delta para la tasa de
187 migración (m), frecuencias de alelos (a) y coeficiente de endogamia (f) de cada
188 comparación, se probaron distintos valores hasta alcanzar los porcentajes de aceptación
189 recomendados por los autores del programa (para m entre 20%–60%; para a y f entre 65%–
190 76%). Se comparó la convergencia de la corrida final de cada prueba con el programa
191 Tracer versión 1.7 (Rambaut et al., 2018). Los valores finales de m , a , f y de los porcentajes
192 de aceptación de cada valor para la prueba global y las pruebas locales se encuentran en el
193 Material Suplementario (Tabla S1).

194

195 **Resultados.**

196 **Diversidad genética.**

197 Solo cinco loci de microsatélites nucleares se comparten entre las tres especies.
198 (CMTm120, CMTm11, CMTp193, CMTm54, CMTp88 (Gong et al., 2008)), los cuales se
199 utilizaron para estimar los estadísticos básicos de diversidad genética. Se obtuvieron 105
200 alelos para el total de la muestra: el 60% corresponde a alelos privados (63 alelos) y el
201 11.42% (12 alelos) corresponde a alelos compartidos entre todas las especies. Los taxa que

202 comparten menos alelos son *C. argyrosperma* subesp. *argyrosperma* y *C. pepo* subesp.
203 *pepo* (8 alelos; Figura 2).

204 La riqueza alélica y el número de alelos privados fueron considerablemente más
205 altos para *C. pepo* subesp. *pepo* ($A = 3.18$; A.P. = 35), mientras que lo valores de *C.*
206 *moschata* ($A = 1.96$; A.P. = 15) y *C. argyrosperma* subesp. *argyrosperma* ($A = 1.72$; A.P. =
207 13) son relativamente similares (Tabla 1 y Figura 2). En términos de diversidad genética
208 (H_E), se mantiene el mismo patrón entre las especies (*C. pepo* subesp. *pepo* = 0.566; *C.*
209 *moschata* = 0.431 y *C. argyrosperma* subesp. *argyrosperma* = 0.357) (Tabla 1). Los tres
210 taxa presentaron valores significativos de deficiencia de heterocigos (Tabla 1).

211

212 **Diferenciación y estructura genética.**

213 Los tres taxa están genéticamente bien diferenciados (Tabla 2). Las especies con mayor
214 diferenciación genética son *C. argyrosperma* subesp. *argyrosperma* y *C. pepo* subesp. *pepo*
215 ($G_{ST} = 0.508$), mientras que las especies menos diferenciadas son *C. moschata* y *C. pepo*
216 subesp. *pepo* ($G_{ST} = 0.457$). El análisis de varianza molecular muestra que el mayor
217 porcentaje de variación se explica entre taxa (85.3%; Tabla S2). El número óptimo de
218 grupos genéticos de acuerdo con el análisis de Structure es $K = 3$ ($\ln(P) = -10365.5$). La
219 mayoría de los individuos de cada una de las especies fue asignado a un grupo genético
220 distinto (verde = *C. argyrosperma* subesp. *argyrosperma*; azul = *C. moschata*; rojo = *C.*
221 *pepo* subesp. *pepo*; Figura 3). Dentro de cada uno de los grupos genéticos, se pueden
222 apreciar individuos híbridos. Al menos 13 individuos identificados del grupo verde (*C.*
223 *argyrosperma* subesp. *argyrosperma*) tienen en su composición genética alrededor de un

224 10% del grupo genético azul (*C. moschata*) y por lo menos un individuo parece estar mal
225 identificado taxonómicamente. Para el caso del grupo genética azul (*C. moschata*) dos
226 individuos parecen estar mal identificados taxonómicamente, y alrededor de 17 individuos
227 tiene un 10% su composición genética asignada a los otros dos grupos. En el caso del grupo
228 rojo (*C. pepo* subesp. *pepo*) también hay evidencias de individuos mal identificados
229 taxonómicamente, y en contraste con los grupos anteriores, pocos individuos presentan
230 señales de mezcla híbrida (Fig. 3). El análisis de DAPC apoya que cada una de las especies
231 conforma un grupo bien definido (Fig. S1). Ambos análisis muestran señales de
232 introgresión entre los tres taxa.

233

234 **Tasas de flujo genético reciente entre taxa.**

235 El análisis global de migración sugiere que cierto flujo genético ocurre o ha sucedido entre
236 las tres especies, sin embargo, la fracción de migrantes es bajo en todos los casos (Tabla 3;
237 Fig. 4). La tasa de migración total para cada especie se calculó como 1- la proporciones de
238 no migrantes. El valor más alto corresponde a *C. pepo* subesp. *pepo* ($m = 0.0254$), seguida
239 de *C. moschata* ($m = 0.0178$) y finalmente la tasa total de migración más baja es la de *C.*
240 *argyrosperma* subesp. *argyrosperma* ($m = 0.0172$) (Fig. 4).

241 La mayor tasa de migración, también estimada con BayesAss, sucede de *C.*
242 *moschata* a *C. pepo* subesp. *pepo* ($m = 0.0161$), seguida de *C. pepo* subesp. *pepo* a *C.*
243 *argyrosperma* subesp. *argyrosperma* ($m = 0.0121$) y de *C. pepo* subesp. *pepo*
244 a *C. moschata* ($m = 0.0105$). La tasa de migración estimada entre las especies cercanas *C.*

245 *argyrosperma* subesp. *argyrosperma* y *C. moschata* resultó ser baja en ambas direcciones
246 ($m = 0.0051-0.0073$) (Tabla 3).

247

248 **Tasas de flujo genético reciente entre *C. argyrosperma* subesp. *argyrosperma* y *C.***
249 ***moschata* a nivel local.**

250 El flujo genético estimadas con BayesAss dentro de cada una de las cinco localidades
251 donde se cultivan simultáneamente *C. argyrosperma* subesp. *argyrosperma* y *C. moschata*
252 se mantiene en ambas direcciones y en términos generales las tasas de migración son bajas
253 ($m\bar{x} = 0.0165-0.0154$; Tabla 4). La localidad en la que se estimó un mayor flujo genético
254 entre estas dos especies fue la de Ometepec, Guerrero donde la tasa de migración de *C.*
255 *argyrosperma* subesp. *argyrosperma* a *C. moschata* es de 0.0186; mientras que la tasa de *C.*
256 *moschata* a *C. argyrosperma* subesp. *argyrosperma* fue de 0.0197 (Tabla 4). La localidad
257 con la menor la tasa de flujo genético fue de Cojumatlán, Michoacán, en donde la tasa de *C.*
258 *argyrosperma* subesp. *argyrosperma* a *C. moschata* es de 0.0148 y de *C. moschata* a *C.*
259 *argyrosperma* subesp. *argyrosperma* es de 0.0141 (Tabla 4). En estas poblaciones donde
260 se cultivan las dos especies, promedio se estimó mayor migración de *C.*
261 *argyrosperma* subesp. *argyrosperma* a *C. moschata* ($m = 0.0162$) que al contrario ($m =$
262 0.1300).

263

264 **Discusión.**

265 **Diversidad genética y endogamia.**

266 Los microsatélites nucleares son altamente polimórficos (Goldstein et al., 1995) y aunque
267 para este estudio se analizaron pocos marcadores se consiguió estudiar la diversidad
268 genética dentro y entre las especies domesticadas del género *Cucurbita*. El número efectivo
269 de alelos privados es una medida que reflejan que tan distinta es una especie respecto a otra
270 (Kalinowski, 2004). En relación a esta medida, *C. pepo* subesp. *pepo* es más distinta (35
271 alelos privados) respecto a *C. moschata* y *C. argyrosperma* subesp. *argyrosperma*. Que
272 esta especie tenga más alelos privados puede deberse a: (1) el linaje al cual pertenece
273 divergió antes (1.3 Ma; Castellanos-Morales et al., 2018) que el linaje al cual pertenecen las
274 especies hermanas *C. moschata* y *C. argyrosperma argyrosperma* subesp. *argyrosperma*
275 (1.06 Ma; Castellanos-Morales et al., 2018) por lo que ha tenido más tiempo para acumular
276 un mayor número de variantes. (2) Los microsatélites analizados fueron diseñados para *C.*
277 *pepo* y *C. moschata* (Gong et al., 2008) y se espera que detecten mas alelos para estas
278 especies.

279 Por otro lado, los alelos compartidos entre pares de especies y entre todas ellas
280 puede deberse a: (1) Ancestría, las especies del mismo género que estan estrechamente
281 relacionadas (como es el caso especies hermanas *C. moschata* y *C. argyrosperma* subesp.
282 *argyrosperma*) comparten más alelos porque provienen de un ancestro en común. (2)
283 Homoplasia. Las especies que pertenecen al mismo género pero no estan estrechamente
284 relacionadas comparten estados alélicos sin que estos lo sean por ancestría (Hamilton,
285 2009).

286 Nuestro análisis comparativo usando los mismos loci, demuestra que *C.*
287 *pepo* subesp. *pepo* y sus variedades locales que se cultivan en México presentan la mayor
288 diversidad genética respecto a las variedades de *C. moschata* y *C. argyrosperma* subesp.

289 *argyrosperma*. La alta diversidad de variedades locales en México y Guatemala de *C.*
290 *pepo* subesp. *pepo* ya había sido descrita anteriormente en otros estudios (Enríquez et al.,
291 2018; Castellanos-Morales et al., 2019). Hasta el momento *C. pepo* subesp. *pepo* parece ser
292 el taxa domesticado con mayor diversidad genética del género *Cucurbita*. Contrario a
293 lo reportado por Sánchez-de la Vega et al. (2018) quienes analizaron a *C. a.* subesp.
294 *argyrosperma* ($H_E = 0.410$; nueve loci) y Hernández-Rosales et al. (2020) que estudiaron a
295 *C. moschata* ($H_E = 0.395$; 11 loci), en este estudio, usando solo cinco loci de microsatélites,
296 la diversidad genética de *C. moschata* ($H_E = 0.430$) resultó mayor a la de *C. a.* subesp.
297 *argyrosperma* ($H_E = 0.357$). Lo anterior puede atribuirse tanto al número y cuales
298 microsatélites que se usaron para evaluar la diversidad, ya que estos marcadores fueron
299 diseñados para *C. pepo* y *C. moschata* (Gong et al., 2008) y no son los más variables para
300 la subesp. *argyrosperma*, como lo reportaron Sanchez-de la Vega y colaboradores (2018).
301 En términos generales, las tres especies analizadas presentan niveles de intermedios a altos
302 de diversidad genética respecto a lo que comúnmente se espera en estudios de
303 microsatélites para plantas anuales ($H_E = 0.46$) (Nybom, 2004).

304 La alta diversidad genética de las calabazas que se cultivan en México que
305 reportamos en este trabajo concuerda con las principales fuentes de evidencia para
306 determinar los centros de domesticación y diversificación (Zohary, 1999). Los trabajos de
307 filogenias moleculares, registros arqueobotánicos, la presencia de parientes silvestres, la
308 amplia variedad de razas criollas y nuestros resultados apoyan el supuesto de que México
309 es centro de domesticación y diversificación de estas especies de *Cucurbita* (Lira, 1995;
310 Smith, 1997; Bisognin, 2002; Sanjur et al., 2002; Paris, 2012; Castellanos-Morales et al.,
311 2019; Kates et al., 2017 y otros).

312 Las tres especies analizadas presentaron niveles significativos de endogamia
313 (promedio $F_{IS} = 0.298$; Tabla1). Nuestros resultados son congruentes con lo reportado en
314 los estudios previos (Sánchez-de La Vega et al., 2018; Castellanos-Morales et al., 2019;
315 Hernández-Rosales et al., 2020). La endogamia es una vía utilizada para fijar caracteres
316 fenotípicos favorables en las especies domesticadas y proceso en el cual los agricultores
317 resguardan, promueven, seleccionan y controlan los procesos reproductivos de unos pocos
318 individuos (Casas y Parra 2016) (Clawson, 1985; Altieri and Merrick, 1987). No obstante,
319 la deficiencia de heterócigos en cada taxa podría deberse al efecto Wahlund (Wahlund,
320 1928), pues cada conjunto de datos por especie esta representado por localidades y parcelas
321 distintas, así cada una de ellas al presentar diferencias en sus frecuencias alélicas generan
322 estructura genética dentro de cada taxa. Además, tanto en *C. argyrosperma* subesp.
323 *argyrosperma* y *C. moschata* se ha reportado una importante diferenciación genética entre
324 las calabazas cultivadas en el sureste de México y las cultivadas en el resto del país
325 (Sánchez-de la Vega et al. 2018; Hernández-Rosales et al. 2020).

326

327 **Diferenciación, estructura y flujo genético global.**

328 A fin a los valores de distancia genética de Nei con isoenzimas reportados por Montes-
329 Hernández y Eguiarte (2002) en localidades de un área pequeña de Jalisco ($F_{ST}'s = 0.122-$
330 0.301), nuestros resultados con loci de microsatélites nucleares también muestran que los
331 tres taxa analizados se encuentran altamente diferenciados ($G_{ST} = 0.457-0.508$). Nuestros
332 resultados de AMOVA son congruentes con la diferenciación genética encontrada entre
333 taxa, ya que más del 85.3% de la variación se debe a las diferencias entre especies. Las

334 señales de diferenciación entre especies cultivadas de *Cucurbita* es congruente con lo que
335 reportó Cerón-Gonzales et al. (2010) utilizando marcadores moleculares de tipo RAPD
336 ($G_{ST} = 0.770$).

337 Como cada especie fue considerada una sola población, el análisis de estructura
338 genética y DAPC identificó claramente a cada taxa como un grupo genético bien definido.
339 Sin embargo, en ambos análisis se puede observar que hay algunos individuos de *C.*
340 *moschata* que se asemejan a *C. p.* subsp. *pepo* y otros un poco a *C. a.* subsp. *argyrosperma*,
341 y de la misma manera algunos individuos de *C. a.* subsp. *argyrosperma* se asemejan a *C.*
342 *moschata* y *C. pepo* (Figura S1). Para aseverar que los patrones observados son
343 consecuencia de la hibridación entre especies se debe explorar y eliminar la posibilidad de
344 una clasificación incompleta de linajes (Zhou et al., 2017).

345 Los histogramas mayores a $K = 4$ (resultados no mostrados) muestran señales de
346 estructura genética al interior de cada especie, congruente con lo se ha reportado
347 anteriormente (Hernández-Rosales et al., 2020; Sánchez-de la Vega et al., 2018;
348 Castellanos-Morales et al., 2019). Además, las bajas tasas de flujo genético, también al
349 considerar a cada especie como una población, son más bajas de lo que se ha reportado para
350 estas especies con isoenzimas en un Jalisco (Montes-Hernández y Eguiarte, 2002).

351 Nuestros resultados confirman que las especies están aisladas una de otras ya que
352 presentan poco flujo genético y se encuentran genéticamente bien diferenciadas (OECD,
353 2012). De acuerdo a nuestro análisis global de estructura genética no podemos descartar la
354 presencia de híbridos espontáneos. Sin embargo, existe una alta probabilidad que los frutos

355 híbridos no sean fértiles y en caso de tener semillas bien desarrolladas estas no seas viables
356 (Lira, 1995).

357 Los resultados obtenidos en este estudio, corroboran de manera indirecta los niveles
358 de la compatibilidad entre las especies de *Cucurbita* descrita por (Whitaker and Bemis,
359 1964), Lira (1995) y Lebeda (2006). Estos autores proponen que las relaciones de
360 compatibilidad, la cantidad y frecuencia de flujo genético esta mediada por el acervo
361 genético primario (p. ej. entre subespecies), secundario (p. ej. entre especies hermanas), y
362 terciario (p. ej. especies del mismo género más alejadas genéticamente). Es decir, cada una
363 de las especies estudiadas se cruzarán más fácilmente con su acervo genético primario, que
364 con el acervo secundario y en menor medida con el terciario. (i) *C. pepo* subesp. *pepo* tiene
365 mayor facilidad de producir híbridos con los diversos cultivares (comestibles y
366 ornamentales) y con los taxa silvestres más relacionados (subesp. *fraterna* y subesp. *ovifera*
367 var. *texana* y var. *ozarkana*), mientras que solo hasta el segundo nivel tiene compatibilidad
368 con otras especies domesticadas como *C. moschata* y *C. argyrosperma* subesp.
369 *argyrosperma*. (ii) *C. argyrosperma* subesp. *argyrosperma* produce fácilmente híbridos con
370 su pariente silvestre *C. argyrosperma* subesp. *sororia*, en un segundo nivel con *C.*
371 *moschata* y hasta un tercer nivel (mucho más limitado) con *C. pepo*. Finalmente, (iii) *C.*
372 *moschata* mantiene la relación anterior, es decir, puede producir híbridos relativamente
373 fácilmente con *C. argyrosperma* (acervo genético secundario) y en menor medida con *C.*
374 *pepo* (Whitaker and Bemis, 1964).

375 Desafortunadamente, el muestreo de nuestro análisis carece de las condiciones
376 adecuadas para realizar inferencias más sofisticadas sobre el flujo genético e introgresión
377 entre estas especies. La distancia geográfica entre los sitios de colecta de cada especies es

378 mucho más amplia que el rango de movilidad de los principales polinizadores *Xenoglossa*
379 y *Peponapis* (~0.7 Km) (Kohn and Casper, 1992), aun cuando estos insectos son muy
380 eficientes al mover grandes cantidades de polen y favorecen el entrecruzamiento,
381 comúnmente se encuentran en bajas densidades (Enríquez et al., 2018). Finalmente,
382 experimentos de cruza artificiales reportados en la literatura han demostrado que existen
383 mecanismos naturales para evitar la introgresión entre los taxa cultivados como las barreras
384 pre-cigóticas y post-cigóticas (Lebeda et al. 2006).

385

386 **Flujo genético a nivel local.**

387 Nuestros datos confirman que el cultivo en conjunto de *C. argyrosperma* subesp.
388 *argyrosperma* y *C. moschata* en México es frecuente en el mismo agroecosistema y en la
389 misma temporada incluso en comunidades no reportadas anteriormente.(Montes-Hernández
390 et al., 2005).

391 Contrario a Montes-Hernández y Eguiarte (2002), quienes reportaron niveles
392 moderados de introgresión entre estas especies utilizando isoenzimas ($Nem\ s = 5.74-11.54$),
393 las tasas de flujo genético en las cinco localidades que se evaluaron en este estudio con loci
394 de microsatélites nucleares son bajas ($Nm = 1.72-2.54$). Como se menciona anteriormente,
395 en ambos casos la especie contraria representa el segundo nivel de acervo genético, por lo
396 que nuestros resultados respaldan el supuesto que, incluso a nivel local, es poco probable la
397 generación de híbridos espontáneos (Lira, 1995; Lebeda et. al., 2006). Aunque son taxa
398 estrechamente relacionados, la fenología de floración es muy parecido y su distribución en
399 México es simpátrica, la incompatibilidad genética, las barreras pre-cigóticas y post-

400 cigóticas (y otros factores probablemente aun no descritos), permiten mantener el
401 aislamiento reproductivo (Lira, 1995). Sin embargo, esto último ha sido investigado poco
402 en las especies del género *Cucurbita* lo que abre un campo de estudio.

403 **Conservación de los recursos genéticos de *Cucurbita* en México.**

404 Las variedades criollas de especies domesticadas del género *Cucurbita* que se siembran de
405 forma tradicional en México representan un importante recurso y reservorio de diversidad
406 genética. El poco flujo genético entre especies permite conservar los acervos genéticos de
407 cada especie. Las barreras naturales que impiden altos niveles de introgresión
408 intraespecífica son una gran ventaja para evitar la posible contaminación de los transgenes
409 de los cultivos mejorados hacia las especies silvestres.

410 Experimentos de fertilización artificial (Whitaker and Bemis, 1964; Kirkpatrick and
411 Wilson, 1988; Maluf et al., 1997; Brown et al., 2003) demuestran que es relativamente fácil
412 obtener híbridos con fines de mejoramiento genético asociado a la transferencia de
413 resistencia a enfermedades. No obstante, para que los programas de hibridación
414 intraespecíficos sean exitosos, los cruzamientos deben llevarse a cabo mayormente con los
415 reservorios genéticos primarios y secundarios de cada especie (Lira, 1995).

416 **Financiamiento.**

417 Este trabajo ha sido financiado por el proyecto CONABIO núm. KE004 “Diversidad
418 genética de las especies de *Cucurbita* en México e hibridación entre plantas genéticamente
419 modificadas y especies silvestres de *Cucurbita*”.

420

421 **Literatura citada.and**

- 422 Adamack, A. T., and B. Gruber. 2014. PopGenReport: Simplifying basic population
423 genetic analyses in R. *Methods in Ecology and Evolution* 5: 384–387.
- 424 Altieri, M. A., and L. Merrick. 1987. In situ conservation of crop genetic resources through
425 maintenance of traditional farming systems. *Economic Botany* 41: 86–96.
- 426 Bisognin, D. A. 2002. Origin and evolution of cultivated cucurbits. *Ciência Rural* 32:4.
- 427 Brown, R. N., A. Bolanos-Herrera, J. R. Myers, and M. M. Jahn. 2003. Inheritance of
428 resistance to four cucurbit viruses in *Cucurbita moschata*. *Euphytica* 129: 253–258.
- 429 Castellanos-Morales, G., L. M. Paredes-Torres, N. Gámez, H. S. Hernández-Rosales, G.
430 Sánchez-de la Vega, J. Barrera-Redondo, E. Aguirre-Planter, et al. 2018. Historical
431 biogeography and phylogeny of *Cucurbita*: Insights from ancestral area reconstruction
432 and niche evolution. *Molecular Phylogenetics and Evolution* 128: 38–54.
- 433 Castellanos-Morales, G., K. Y. Ruiz-Mondragón, H. S. Hernández-Rosales, G. S. la Vega,
434 E. Aguirre-Planter, S. Montes-Hernández, R. Lira-Saade, and L. E. Eguiarte. 2019.
435 Tracing back the origin of pumpkins (*Cucurbita pepo* ssp. *pepo* L.) in Mexico. *Proc.*
436 *R. Soc. B.* 286: 20191440.
- 437 Cerón González, L., J. P. Legaria Solano, C. Villanueva Verduzco, and J. Sahagún
438 Castellanos. 2010. Diversidad genética en cuatro especies mexicanas de calabaza
439 (*Cucurbita* ssp.). *Revista Fitotecnia Mexicana* 33: 189–196.
- 440 Cerrato, E. J., D. Cosano, D. Esquivel, R. Otero, C. Jiménez-Sanchidrián, and J. R. Ruiz.
441 2020. A multi-analytical study of a wall painting in the Satyr domus in Córdoba,

442 Spain. *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy* 232:
443 118148.

444 Chomicki, G., H. Schaefer, and S. S. Renner. 2019. Origin and domestication of
445 Cucurbitaceae crops: insights from phylogenies, genomics and archaeology. *New*
446 *Phytologist*: 0–3.

447 Clawson, D. L. 1985. Harvest Security and Intraspecific Diversity in Traditional Tropical
448 Agriculture. *Economic Botany* 39, 56–67.

449 Cruz-Reyes, R., G. Ávila-Sakar, G. Sánchez-Montoya, and A. M. Quesada. 2015.
450 Experimental assessment of gene flow between transgenic squash and a wild relative
451 in the center of origin of cucurbits. *Ecosphere* 6: 1–13.

452 Decker-Walters, D. S., T. W. Walters, U. Posluszny, and P. G. Kevan. 1990. Genealogy
453 and gene flow among annual domesticated species of *Cucurbita*. *Canadian Journal of*
454 *Botany* 68(4): 782–789.

455 Dowling, T. E., and C. L. Secor. 1997. The role of hybridization and introgression in the
456 diversification of animals. *Annual Review of Ecology and Systematics* 28:593–619.

457 Ellstrand, N. C., R. Whitkus, and L. H. Rieseberg. 1996. Distribution of spontaneous plant
458 hybrids. *Proceedings of the National Academy of Sciences of the United States of*
459 *America* 93 (10) 5090-5093.

460 Enríquez, E., P. Landaverde-González, R. Lima-Cordón, E. Solórzano-Ortíz, R. Tapia-
461 López, and J. Nuñez-Farfán. 2018. Population genetics of traditional landraces of
462 *Cucurbita pepo* L., 1753 in the cloud forest in Baja Verapaz, Guatemala. *Genetic*

- 463 *Resources and Crop Evolution* 65: 979–991.
- 464 Goldstein, D. B., A. R. Linares, L. L. Cavalli-Sforza, and M. W. Feldman. 1995. An
465 evaluation of genetic distances for use with microsatellite loci. *Genetics* 139 (1): 1
466 463-471.
- 467 Gong, L., G. Stift, R. Kofler, M. Pachner, and T. Lelley. 2008. Microsatellites for the genus
468 *Cucurbita* and an SSR-based genetic linkage map of *Cucurbita pepo* L. *Theoretical*
469 *and Applied Genetics* 117: 37–48.
- 470 González, L. C., J. P. L. Solano, C. V. Verduzco, and J. S. Castellanos. 2010. Diversidad
471 genética en cuatro especies Mexicanas de calabaza (*Cucurbita* spp.). *Revista*
472 *Fitotecnia Mexicana* 33: 189–196.
- 473 Hamilton, M. 2009. Populations genetics. Wiley-Blackwell. Malasia. pp 68-73, 105-149 y
474 235–281.
- 475 Harrison, R. G. 1990. Hybrid zones: widows on evolutionary process. *Oxford Surveys in*
476 *Evolutionary Biology* Vol.7 pp.69–128.
- 477 Hernández-Rosales, H. S., G. Castellanos-Morales, G. Sánchez-de la Vega, E. Aguirre-
478 Planter, S. Montes-Hernández, R. Lira-Saade, and L. E. Eguiarte. 2020.
479 Phylogeographic and population genetic analyses of *Cucurbita moschata* reveal
480 divergence of two mitochondrial lineages linked to an elevational gradient. *American*
481 *Journal of Botany* 107: 510–525.
- 482 Hernández Xolocotzi, E. 1988. La agricultura tradicional en México. *Comercio Exterior* 38
483 (8): 673–678.

484 Hurd, P. D., E. G. Linsley, and T. W. Whitaker. 1971. Squash and Gourd Bees (*Peponapis*,
485 *Xenoglossa*) and the Origin of the Cultivated *Cucurbita*. *Evolution* 25(1), 218–234.

486 Jombart, T. 2008. Adegnet: a R package for the multivariate analysis of genetic markers.
487 *Bioinformatics* 24: 1403–1405.

488 Jombart, T., S. Devillard, and F. Balloux. 2010. Discriminant analysis of principal
489 components: a new method for the analysis of genetically structured populations. *BMC*
490 *Genetics* 11: 94.

491 Kalinowski, S. T. 2004. Counting Alleles with Rarefaction: Private Alleles and
492 Hierarchical Sampling Designs. *Conservation Genetics* 5: 539–543.

493 Kates, H. R., P. S. Soltis, and D. E. Soltis. 2017. Evolutionary and domestication history of
494 *Cucurbita* (pumpkin and squash) species inferred from 44 nuclear loci. *Molecular*
495 *Phylogenetics and Evolution* 111: 98–109.

496 Kirkpatrick, K. J., and H. D. Wilson. 1988. Interspecific Gene Flow in *Cucurbita*: *C.*
497 *texana* vs. *C. pepo*. *American Journal of Botany*.

498 Kohn, J. R., and B. B. Casper. 1992. Pollen-mediated gene flow in *Cucurbita foetidissima*
499 (*Cucurbitaceae*). *American Journal of Botany* 75: 519–527.

500 Maluf, W. R., J. J. Pereira, and A. R. Figueira. 1997. Inheritance of resistance to the papaya
501 ringspot virus-watermelon strain from two different accessions of winter squash
502 *Cucurbita maxima* Duch. *Euphytica* 94: 163–168.

503 Matos, E. L. S., E. J. Oliveira, O. N. Jesus, and J. L. L. Dantas. 2013. Microsatellite
504 markers of genetic diversity and population structure of *Carica papaya*. *Annals of*

- 505 *Applied Biology* 163: 298–310.
- 506 Meirmans, P. G., and P. H. Van Tienderen. 2004. GENOTYPE and GENODIVE: Two
507 programs for the analysis of genetic diversity of asexual organisms. *Molecular*
508 *Ecology Notes* 4: 792–794.
- 509 Montes-Hernández, S., and L. E. Eguiarte. 2002. Genetic structure and indirect estimates of
510 gene flow in three taxa of *Cucurbita* (Cucurbitaceae) in western Mexico. *American*
511 *Journal of Botany* 89: 1156–1163.
- 512 Montes-Hernández, S., L. C. Merrick, and L. E. Eguiarte. 2005. Maintenance of squash
513 (*Cucurbita* spp.) landrace diversity by farmers' activities in Mexico. *Genetic*
514 *Resources and Crop Evolution* 52: 697–707.
- 515 Nybom, H. 2004. Comparison of different nuclear DNA markers for estimating
516 intraspecific genetic diversity in plants. *Molecular Ecology* 13: 1143–1155.
- 517 Paris, H. S. 2016. Genetic Resources of Pumpkins and Squash *Cucurbita* spp. In R. Grumet,
518 N. Katzir, and J. Garcia-Mas [eds.], *Genetics and genomics of Cucurbitaceae*. Plant
519 genetics and genomics: crops and models, vol 20. Springer, Cham.
- 520 Paris, H. S., and R. Cohen. 2002. Powdery mildew-resistant summer squash hybrids having
521 higher yields than their susceptible, commercial counterparts. *Euphytica* 124: 121–
522 128.
- 523 Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of Population Structure
524 Using Multilocus Genotype Data. *Genetics* 155: 945–959.
- 525 Rambaut, A., A. J. Drummond, D. Xie, G. Baele, and M. A. Suchard. 2018. Posterior

526 summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology* 67(5):
527 901–904.

528 Rieseberg, L. H. 1997. Hybrid origins of plant species. *Annual Review of Ecology and*
529 *Systematics* 28: 359–389.

530 Sánchez-de La Vega, G., G. Castellanos-Morales, N. Gámez, H. S. Hernández-Rosales, A.
531 Vázquez-Lobo, E. Aguirre-Planter, J. P. Jaramillo-Correa, et al. 2018. Genetic
532 resources in the “calabaza pipiana” squash (*Cucurbita argyrosperma*) in Mexico:
533 Genetic diversity, genetic differentiation and distribution models. *Frontiers in Plant*
534 *Science* 9: 1–13.

535 Smith, B. D. 1997. The Initial Domestication of *Cucurbita pepo* in the Americas 10,000
536 Years Ago. *Science* 276: 932–934.

537 Wahlund, S. 1928. Composition of populations from the perspective of the theory of
538 heredity. *Hereditas* 11: 65–105.

539 Whitaker, T. W., and W. P. Bemis. 1964. Evolution in the Genus *Cucurbita*. *Evolution* 18:
540 553.

541 Whitaker, T. W., H. C. Cutler, and R. S. Macneish. 1957. Cucurbit Materials from 3 Caves
542 near Ocampo, Tamulipas. *American Antiquity* 22: 352–358.

543 Wilson, G. A., and B. Rannala. 2003. Bayesian inference of recent migration rates using
544 multilocus genotypes. *Genetics* 163: 1177–1191.

545 Wright, S. 1965. The Interpretation of Population Structure by F-Statistics with Special
546 Regard to Systems of Mating. *Evolution* 19: 395–420.

547 Zhou, Y., L. Duvaux, G. Ren, L. Zhang, O. Savolainen, and J. Liu. 2017. Importance of
548 incomplete lineage sorting and introgression in the origin of shared genetic variation
549 between two closely related pines with overlapping distributions. *Heredity* 118: 211–
550 220.

551 Zohary, D. 1999. Monophyletic vs. polyphyletic origin of the crops on which agriculture
552 was founded in the Near East. *Genetic Resources and Crop Evolution* 46, 133–142.

553

554 **Tablas.**

555

556 **Tabla 1.** Estadísticos de diversidad genética para cinco loci de microstélites nucleares de
 557 tres especies domesticadas de *Cucurbita* (n = tamaño de muestra; NPop = número de
 558 poblaciones por muestra; *A* = riqueza alélica rarefaccionada; A. P. = alelo privados; *H_O* =
 559 heterocigosis observada; *H_E* = heterocigosis esperada; *G_{IS}* = coeficiente de
 560 endogamia; *Nm* = tasa de migración en el equilibrio promedio dentro de cada especie; **p* <
 561 0.05).

	n	NPop	<i>A</i>	A.P.	<i>H_O</i>	<i>H_E</i>	<i>G_{IS}</i>	<i>Nm</i>
<i>C. argyrosperma</i> subesp. <i>argyrosperma</i>	322	19	1.72	13	0.274	0.357	0.232*	1.72
<i>C. moschata</i>	446	24	1.96	15	0.291	0.43	0.324*	1.78
<i>C. pepo</i> subesp. <i>pepo</i>	291	13	3.18	35	0.375	0.566	0.337*	2.54
	$\Sigma=1059$	$\Sigma=36$	$\bar{x}=1.90$	S=63	$\bar{x}=0.314$	$\bar{x}=0.451$	$\bar{x}=0.305^*$	

562

563

564 **Tabla 2.** Diferenciación genética (*G_{ST}*) entre pares de taxa domesticados del género
 565 *Cucurbita* estimada a partir de cinco loci de microsatélites (** = *p* < 0.05).

	<i>C. argyrosperma</i> subesp. <i>argyrosperma</i>	<i>C. moschata</i>
<i>C. moschata</i>	0.502**	
<i>C. pepo</i> subesp. <i>pepo</i>	0.508**	0.457**

566

567

568

569 **Tabla 3.** Tasas de flujo genético reciente entre taxa (modelo global), estimadas con
 570 BayesAss. Los valores en negrita indican la fracción de individuos que permanecen dentro
 571 de la muestra, el resto de los valores indican las proporciones de migrantes ($m = \text{tasa de}$
 572 migración) con cada una de los taxa. La desviación estándar de la distribución posterior
 573 marginal para cada estimación se muestra entre paréntesis (\pm s. d.).

	Migra de:		
Migra a:	<i>C. argyrosperma</i> subesp. <i>argyrosperma</i>	<i>C. moschata</i>	<i>C. pepo</i> subesp. <i>pepo</i>
<i>C. argyrosperma</i> subesp. <i>argyrosperma</i>	0.9828 (0.0041)	0.0051 (0.0023)	0.0121 (0.0035)
<i>C. moschata</i>	0.0073 (0.0024)	0.9822 (0.0036)	0.0105 (0.0027)
<i>C. pepo</i> subesp. <i>pepo</i>	0.0093 (0.0033)	0.0161 (0.0042)	0.9746 (0.0052)

574

575

576

577

578

579

580

581

582

583

584

585

586 **Tabla 4.** Tasas de flujo genético entre *C. argyrosperma subesp. argyrosperma* y *C.*
587 *moschata* de cinco localidades en México, estimadas con BayesAss. Los valores en negrita
588 indican la proporción de migrantes de cada especie. La desviación estándar de la
589 distribución posterior marginal para cada estimación se muestra entre paréntesis (\pm s. d.).
590

	Migra de:		
Migra a:	<i>C. argyrosperma</i> subesp. <i>argyrosperma</i>	<i>C. moschata</i>	Localidad
<i>C. a. subesp. argyrosperma</i>	0.9854 (0.0138)	0.0146 (0.0138)	Coahuila CCC
<i>C. moschata</i>	0.0152 (0.0142)	0.9848 (0.0142)	
<i>C. a. subesp. argyrosperma</i>	0.9803 (0.0185)	0.0197 (0.0185)	Guerrero Ometepec
<i>C. moschata</i>	0.0186 (0.0177)	0.9814 (0.0177)	
<i>C. a. subesp. argyrosperma</i>	0.9849 (0.0145)	0.0151 (0.0145)	Oaxaca Tehuántepec
<i>C. moschata</i>	0.0172 (0.0165)	0.9828 (0.0165)	
<i>C. a. subesp. argyrosperma</i>	0.9863 (0.0130)	0.0137 (0.0130)	Yucatán Buctzotz
<i>C. moschata</i>	0.0151 (0.0145)	0.9849 (0.0145)	
<i>C. a. subesp. argyrosperma</i>	0.9860 (0.0135)	0.0140 (0.0135)	Yucatán Temozón
<i>C. moschata</i>	0.0164 (0.0159)	0.9836 (0.0159)	
	$\bar{x} = \mathbf{0.0165}$	$\bar{x} = \mathbf{0.0154}$	

591
592
593
594

595 **Índice de figuras.**

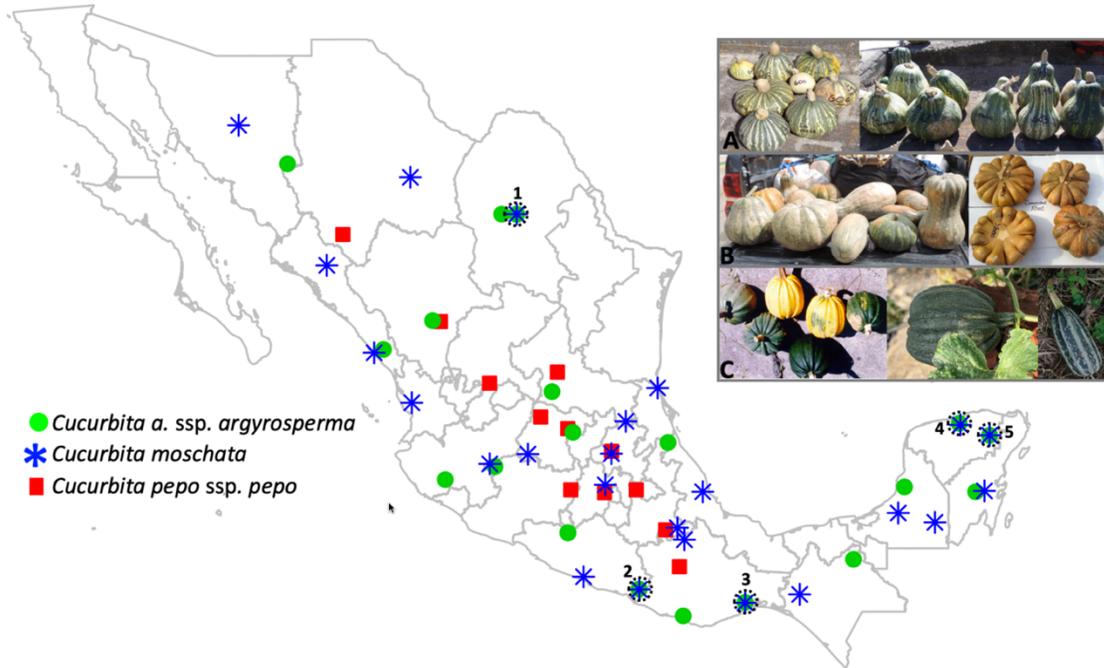
596 **Figura 1.** Mapa de distribución de las localidades analizadas de las variedades locales de
597 tres taxa domesticados del género *Cucurbita* (ver Apéndice 1). Los círculos punteados de
598 *C. moschata* y *C. argyrosperma* subesp. *argyrosperma* indican las localidades en las que se
599 llevo a cabo los análisis de migración a nivel local: 1) Coahuila CCC, 2) Guerrero
600 Ometepec, 3) Oaxaca Tehuantepec, 4) Yucatán Buctotz y 5) Yucatán Temozón. A la
601 derecha se muestra parte de la diversidad morfológica de las variedades locales que crecen
602 en México: A) *C. argyrosperma* subesp. *argyrosperma*, B) *C. moschata* y C) *C.*
603 *pepo* subesp. *pepo*.

604 **Figura 2.** Diagrama de Venn de alelos compartidos de cinco loci de microsatélites de tres
605 taxa domesticados de *Cucurbita* en México. Al interior del diagrama el número de alelos
606 compartidos entre todas las especies. En la parte intermedia se muestra el número de alelos
607 compartidos entre cada par de taxa. Al exterior el número de alelos privados para cada
608 especie.

609 **Figura 3.** Análisis de Structure ($K = 3$) usando cinco loci de microsatélites. . Cada taxa
610 analizado esta representado por un grupo genético bien diferenciado (verde = *C.*
611 *argyrosperma* subesp. *argyrosperma*; azul = *C. moschata*; rojo = *C. pepo* subesp. *pepo*).

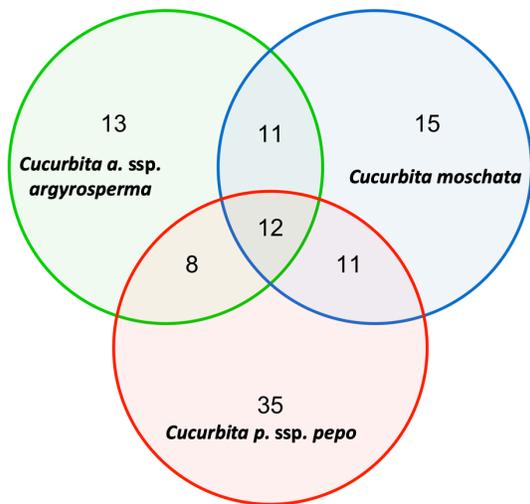
612 **Figura 4.** Tasas de flujo genético reciente entre taxa domesticados de *Cucurbita* en
613 México, estimadas con BayesAss. El valor al centro de los círculos indica la tasa de
614 migración de cada especie (m) calculado como $1 -$ la proporción de no migrantes. Las
615 flechas indican la dirección del flujo genético entre taxa. Los valores indican la proporción

616 de migrantes y la desviación estándar de la distribución posterior marginal para cada
617 estimación se muestra entre paréntesis (\pm s. d.).



619

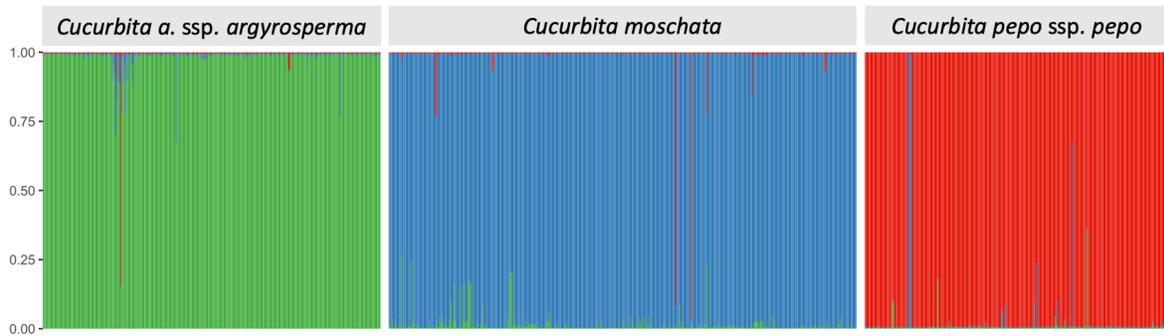
620 **Figura 5.** Mapa de distribución de las localidades analizadas de las variedades locales de tres taxa
 621 domesticados del género *Cucurbita* (ver Apéndice 1). Los círculos punteados de *C. moschata* y *C.*
 622 *argyrosperma* subesp. *argyrosperma* indican las localidades en las que se llevo a cabo los análisis
 623 de migración a nivel local: 1) Coahuila ,CCC, 2) Guerrero, Ometepec, 3) Oaxaca, Tehuantepec, 4)
 624 Yucatán, Buctzotz y 5) Yucatán, Temozón. A la derecha se muestra parte de la diversidad
 625 morfológica de las variedades locales que crecen en México: A) *C. argyrosperma* subesp.
 626 *argyrosperma*, B) *C. moschata* y C) *C. pepo* subesp. *pepo*.



627

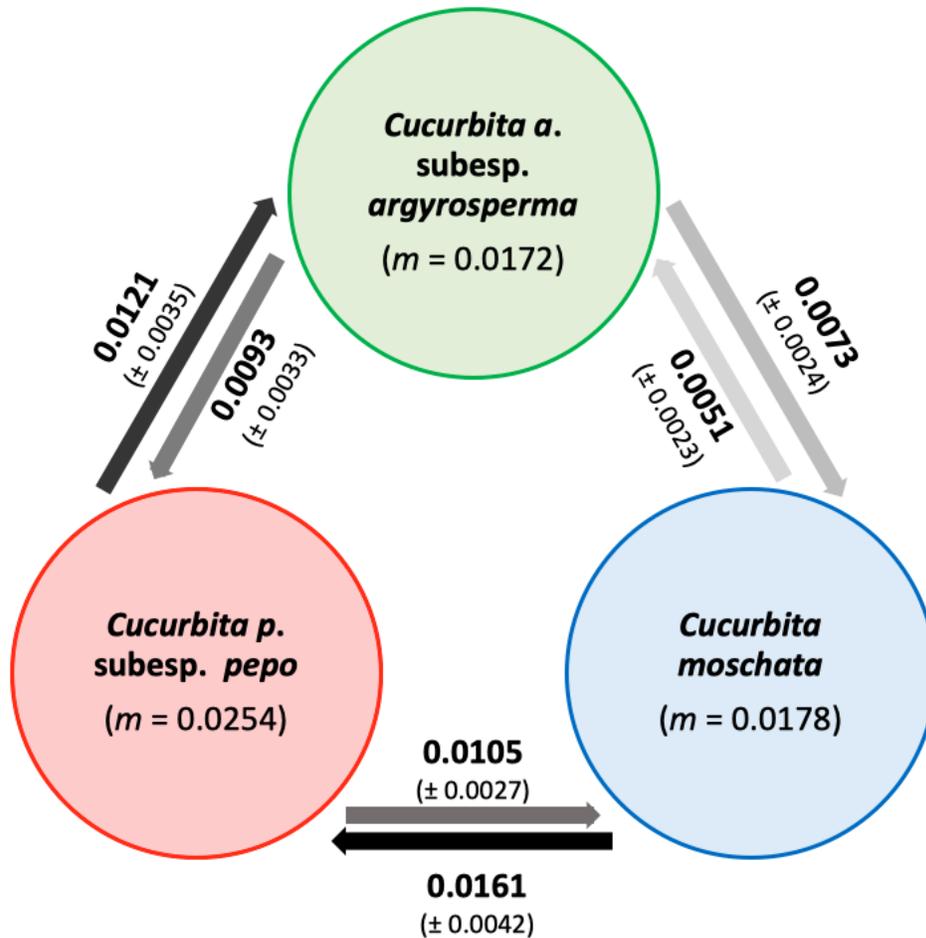
628 **Figura 6.** Diagrama de Venn de alelos compartidos de cinco loci de microsatélites de tres taxa
 629 domesticados de *Cucurbita* en México. Al interior del diagrama el número de alelos compartidos
 630 entre todas las especies. En la parte intermedia se muestra el número de alelos compartidos entre
 631 par de taxa. Al exterior el número de alelos privados para cada especie.

632



633

634 **Figura 7.** Análisis de Structure ($K = 3$) usando cinco loci de microsatélites. . Cada taxa
 635 analizado esta representado por un grupo genético bien diferenciado (verde = *C.*
 636 *argyrosperma subesp. argyrosperma*; azul = *C. moschata*; rojo = *C. pepo subesp. pepo*).



637

638 **Figura 8.** Tasas de flujo genético reciente entre taxa domesticados de *Cucurbita* en
 639 México, estimadas con BayesAss. El valor al centro de los círculos indica la tasa de
 640 migración de cada especie (m) calculado como $1 -$ la proporción de no migrantes. Las
 641 flechas indican la dirección del flujo genético entre taxa. Los valores indican la proporción
 642 de migrantes y la desviación estándar de la distribución posterior marginal para cada
 643 estimación se muestra entre paréntesis (\pm s. d.).

644

645

646

647

648

649

650 **Material suplementario.**

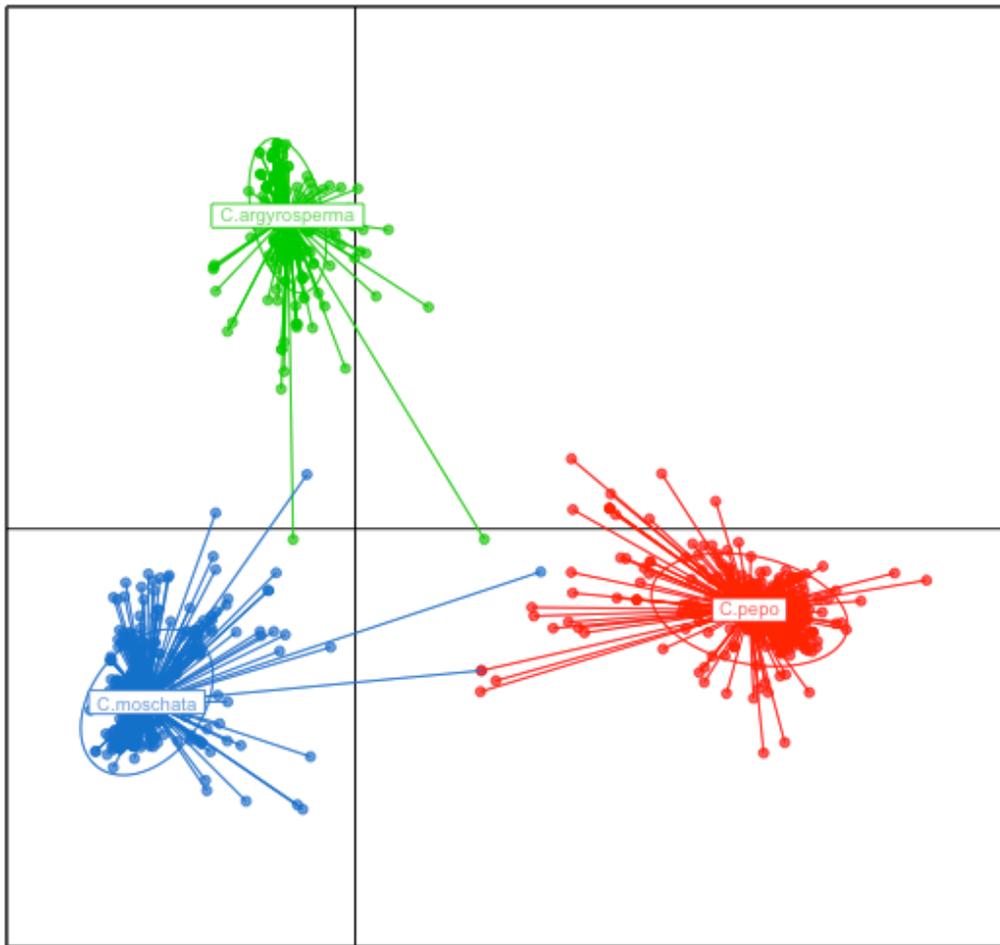
651 **Tabla S 1.** En negrita los valores finales de delta para la tasa de migración (m), frecuencias
 652 de alelos (a) y coeficiente de endogamia (f) de cada comparación para estimar las de tasas
 653 de flujo genético utilizando el programa BayesAss. Entre paréntesis, los porcentajes de
 654 aceptación para cada valor, de acuerdo a los autores del programa se aceptan los valores de
 655 m entre 20%–60% y para a y f entre 65%–76%.

	m	a	f
Global (<i>C. argyrosperma</i> <i>subesp. argyrosperma</i> <i>C. moschata</i> <i>C. pepo</i> <i>subesp. pepo</i>)	0.02 (0.47)	0.006 (0.76)	0.06 (0.71)
Local (<i>C. argyrosperma subesp. argyrosperma</i> <i>C. moschata</i>)			
Coahuila CCC	0.10 (0.49)	0.10 (0.75)	0.03 (0.73)
Guerrero Ometepec	0.25 (0.30)	0.15 (0.69)	0.30 (0.58)
Oaxaca Tehuantepec	0.10 (0.51)	0.10 (0.74)	0.30 (0.74)
Yucatán Buctzotz	0.10 (0.50)	0.10 (0.72)	0.35 (0.74)
Yucatán Temozón	0.10 (0.48)	0.10 (0.71)	0.40 (0.72)

656

657 **Tabla S 2.** Componentes de un análisis de varianza molecular (AMOVA) para cinco
 658 microsatélites nucleares de tres especies domesticadas del género *Cucurbita* en México (*C.*
 659 *argyrosperma subesp. argyrosperma*, *C. moschata* y *C. pepo*).

Fuente de variación	% variación	Coefficientes de F
Entre los individuos	0.094	0.906
Entre los individuos dentro de cada especie	0.053	0.36
Entre especies	0.853	0.853



660

661 **Figura S 1.** Análisis de discriminantes de componentes principales (DAPC) para todas las
 662 poblaciones de cada uno de los taxa domesticados de *Cucurbita*. Cada especie conforma un
 663 grupo genético bien diferenciado.

Capítulo 3: GENÓMICA DE POBLACIONES E IDENTIFICACIÓN DE LOCI BAJO SELECCIÓN DE *Cucurbita moschata*.

Genómica de poblaciones e identificación de loci con señales de selección de *Cucurbita moschata*.

El estudio de las especies domesticadas desde el enfoque de la genómica de poblaciones y genómica comparada permite elucidar el origen de las especies cultivadas y la identificación de genes con potencial para el mejoramiento de cultivos (Eguiarte et al., 2018; Barrera-Redondo et al., 2019). Las especies del género *Cucurbita* han resultado excelentes modelos para el estudio moderno del proceso de domesticación y para entender el papel de la adaptación y selección artificial (Eguiarte et al., 2018).

Las técnicas de secuenciación de nueva generación (también llamados métodos de secuenciación paralela masiva) en conjunto con el desarrollo de paqueterías bioinformáticas han permitido hoy en día tener al menos cuatro genomas de referencia y miles de SNPs de las especies de *Cucurbita* domesticadas y datos de algunas silvestres (Sun et al., 2017; Montero-Pau et al., 2018; Barrera-Redondo et al., 2019; Nguyen et al., 2020; Barrera-Redondo et al., 2021; Martínez-González et al., 2021). Sin embargo, hoy en día no se sabe mucho acerca de los patrones generales de la genómica de poblaciones y los posibles genes con potencial para el fitomejoramiento de *Cucurbita moschata*.

Se analizaron 96 individuos de 24 localidades de *Cucurbita moschata* (analizadas por Hernández-Rosales et al. [2020]) con un método de representación reducida (tGBS) para obtener SNPs. Con estos datos fue posible explorar los patrones generales de la genómica de poblaciones de esta especie, así como identificar loci con señales de selección. Esta información permitirá eventualmente plantear hipótesis evolutivas de la historia, origen y domesticación de esta especie, así mismo su relación con otras especies domesticadas y/o silvestres. Otro objetivo fue identificar polimorfismos con señales de selección a través de una prueba de *Outliers* basada en el parámetro F_{ST} (Foll y Gaggiotti, 2008a).

Obtención de las muestras e identificación de SNPs.

Se analizaron las 24 poblaciones estudiadas por Hernández-Rosales et al. (2020) y la información de cada una de ellas puede ser consultada en dicha referencia. Estas localidades representan variedades locales cultivadas dentro de la distribución natural de *C. moschata* en México. Para obtener una mejor representación de la diversidad genética de las localidades y evitar que los individuos estuvieran estrechamente relacionados se eligió a una planta proveniente de una semilla por fruto, y en total se analizaron cuatro individuos por cada localidad. Se realizaron 96 extracciones de ADN con calidad genómica (Yield ug = 1000 y pureza A260/A280 = 1.7-2.0) y se genotiparon utilizando la tecnología de secuenciación tGBS® (*tunable Genotyping by Sequencing*) (Ott et al., 2017). Las muestras de ADN de estas poblaciones fueron enviadas a secuenciar en la empresa Data2Bio (<https://www.data2bio.com/>) que utiliza la plataforma *Ion Torrent* e implementa dos enzimas de restricción (Sau3AI/BfuCI and Nsp) y un grupo de cebadores anónimos para realizar bibliotecas de 200 a 300 pb.

Se generaron aproximadamente ~115.4 M de lecturas sin procesar. Se utilizó como guía el *pipeline* implementado por Barrera-Redondo y Sánchez-de la Vega (2021) para la identificación de polimorfismos de una sola base. Primero, las lecturas fueron recortadas (*trimmed*; se descartaron aproximadamente 30pb) y filtradas por su calidad PHRED (≥ 20 en el 80% de sus bases) con el programa Trimmomatic (Bolger et al., 2014). Posteriormente, las secuencias se mapearon contra el genoma de referencia de *C. moschata* (Sun et al., 2017) con el programa Segemehl (Hoffmann et al., 2014) y solo se conservaron las secuencias con mapeo único.

La identificación de polimorfismos de una sola base (SNPs) no neutrales se realizó con el programa Plink (Purcell et al., 2007) y se utilizaron los siguientes criterios: profundidad mínima de secuenciación 3X; se eliminaron los SNPs con más del 50% de *missing data* y posteriormente se eliminaron los individuos con más del 50% de *missing data*; se eliminaron los SNPs con frecuencia mínima de alelos (MAF) menor a 0.02. En total tuvimos 9,720 SNPs.

Consecutivamente con el primer set de 9,720 SNPs se implementó una prueba exacta de equilibrio de Hardy-Weinberg con un valor de corte menor a $p \leq 0.01$ y una prueba de desequilibrio de ligamiento ($r^2 > 0.1$) para obtener los loci neutrales. En total tuvimos 1,797 SNPs neutrales.

Diversidad, diferenciación y estructura genética poblacional.

Los estadísticos clásicos de diversidad: heterocigosis esperada (H_E), heterocigosis observada (H_O), coeficiente de endogamia (F_{IS}) y número privado de alelos (A_p) se calcularon con los 1,797 SNPs neutrales utilizando el módulo *populations* del programa Stacks (Catchen et al., 2013). Se calculó la diferenciación entre localidades con el índice de distancia genética de Nei (Nei, 1978) con la librería de R “adegenet” (Jombart, 2008).

Con los SNPs neutrales se evaluó la estructura genética de las poblaciones mediante el programa Admixture (Alexander et al., 2009) y pusimos a prueba valores de $K = 2-12$. Para definir el número de grupos genéticos que explican los datos se realizó una prueba de *Cross-validation*. Finalmente se realizó un análisis de componentes principales (PCA) con la librería de R “adegenet” (Jombart, 2008).

Identificación de loci con señales de selección.

Para llevar a cabo la prueba *Outliers* basado en el parámetro F_{ST} se utilizó la base de 9,720 SNPs y se implementó en el programa Bayescan (Foll y Gaggiotti, 2008b). A través de diferencias en las frecuencias alélicas entre poblaciones para cada locus, se identificaron aquellos que no se ajustan a patrones esperados bajo un modelo neutral. BayeScan es un método que busca adaptaciones locales para datos de frecuencias alélicas, modelando por separado en una población específica y para locus específicos (α), el análisis es sensible a la selección que actúa en algún locus en particular. Para cada polimorfismo, BayeScan estima la distribución posterior bajo neutralidad $\alpha = 0$, de manera separada permite la selección $\alpha \neq 0$ y computa la probabilidad posterior como medida de apoyo para el modelo de adaptación local relativa a la demografía neutral.

Se realizaron dos corridas independientes con los parámetros por default: con un *prior odds ratio* (PO) de 10 que corresponde a un *prior*, en el cual el modelo de selección es 1/10 tal cuál, como un modelo neutral para cualquier SNP, 20 *pilot* se corrieron, un *burn-in* de 50,000 iteraciones, seguidas de 50,000 iteraciones de salida, con un intervalo de 10 en 5,000 iteraciones de estimación posterior. Se analizaron ambas corridas para demostrar que los datos convergieran con la librería *coda* implementada en R (Plummer et al., 2006). Para controlar los falsos positivos, se calculó un valor esperado de $\log_{10}PO$ que se ajusta al 1% en la tasa de

descubrimiento de falsos positivos (FDR). Para determinar los SNPs con señales de adaptación se utilizó la interpretación de Jeffrey (1961).

Resultados.

Los estadísticos de diversidad se estimaron a partir de los 1,797 SNPs neutrales. La diversidad genética promedio en términos de la heterocigosis esperada (H_E) de las 24 localidades analizadas de *C. moschata* fue de $H_E = 0.1324$ (ver Tabla 1). La localidad con la menor diversidad genética fue el estado de Tamaulipas ($H_E = 0.0692$), mientras que la localidad con mayor diversidad fue el estado de San Luis Potosí ($H_E = 0.1609$). La localidad de Yucatán-Buctzotz mostró un mayor número de alelos privado ($A_p = 43$), mientras que la de Nayarit fue la que presentó menos ($A_p = 2$). Todos los valores de diversidad genética para cada localidad se presentan en la Tabla 1.

En cuanto a los valores de coeficiente de endogamia (F_{IS}), el promedio para la especie fue de 0.2423 (Tabla 1). La mayoría de las poblaciones presentaron niveles de endogamia positivos, con un rango que va de $F_{IS} = 0.00557$ – 0.16598. Sin embargo, ningún valor de F_{IS} resultó significativamente diferente de cero.

La diferenciación promedio para la especie en términos de F_{ST} fue de 0.275. La diferenciación genética entre pares de localidades fluctuó entre 0.284 (Veracruz–Quintana-Roo) y 0.046 (Oaxaca-CU–Sinaloa-MO). Se construyó un dendrograma a partir de la matriz de distancias genéticas de Nei, donde se puede observar que las localidades de la Península de Yucatán (Campeche-CK y CH, Yucatán-TE y BU y Quintana-Roo) conforman un grupo bien definido (Figura 1).

El *Cross-validation* del análisis de Admixture identificó $K = 3$ para las muestras analizadas (Figura 2). La mayoría de los individuos de las localidades de Chihuahua, Coahuila, Sinaloa-MA, Nayarit y Guerrero-OM fueron asignados al grupo genético 1 (rojo). Los individuos de la población de Veracruz muestran una mezcla de los tres grupos genéticos. El grupo genético 2 (amarillo) se encuentra bien representado en los individuos de las poblaciones de Michoacán, Tamaulipas, Hidalgo y Guerrero-CO. Los individuos de las localidades de la Península de Yucatán (Campeche-CK y CH, Yucatán-TE y BU y Quintana-Roo) fueron asignados principalmente al grupo genético 3 (naranja).

De acuerdo con el análisis de componentes principales (PCA), la suma de los componentes 1 y 2 explican el 28.66% de la variación de 1,797 SNPs, sin embargo, el agrupamiento de las localidades es poco claro. Por el contrario, la suma de los componentes 1 y 3

(que explican el 27% de la variación; Figura 3) permite identificar un mejor agrupamiento de las localidades de acuerdo con la discriminación de variedades propuesta por Hernández-Rosales et al. (2020). En la Figura 3 se puede apreciar que existe diferenciación genética de las localidades de la Península de Yucatán, las localidades de Oaxaca y Veracruz conservan el orden geográfico latitudinal, mientras que no hay patrón claro entre las localidades del resto de México.

Tabla 1. Estadísticos de diversidad genómica para 24 localidades de *C. moschata* y 1,797 SNPs neutrales. (A_p = alelos privados; H_o = heterocigosis observada; H_E = heterocigosis esperada; F_{IS} = coeficiente de endogamia).

Localidad	A_p	H_o	H_E	F_{IS}
Sonora	5	0.1135	0.1331	0.0968
Chihuahua	29	0.1316	0.1339	0.0513
Coahuila	16	0.1186	0.1141	0.0374
Sinaloa-MO	14	0.1413	0.1276	0.0270
Sinaloa-MA	36	0.0962	0.1481	0.1660
Nayarit	2	0.1381	0.1390	0.0627
Michoacán	11	0.1234	0.1288	0.0574
Tamaulipas	24	0.0998	0.0692	-0.0005
S.L.Potosí	5	0.1617	0.1609	0.0558
Guanajuato	8	0.1276	0.1333	0.0596
Hidalgo	19	0.1107	0.1161	0.0553
Edomex	8	0.1549	0.1515	0.0528
Guerrero-CO	24	0.1197	0.1020	0.0056
Guerrero-OM	7	0.1220	0.1365	0.0793
Veracruz	13	0.1403	0.1287	0.0302
Puebla	17	0.1097	0.1247	0.0730
Oaxaca-TE	21	0.1314	0.1360	0.0560
Oaxaca-CU	16	0.1000	0.1317	0.1095
Chiapas	26	0.1384	0.1445	0.0693
Campeche-CK	25	0.1404	0.1435	0.0576
Campeche-CH	13	0.1478	0.1430	0.0509
Yucatán-BU	43	0.1345	0.1299	0.0389
Yucatán-TE	24	0.1401	0.1554	0.0891
Quintana-Roo	21	0.1357	0.1475	0.0807
	17.8	0.1282	0.1324	0.2423

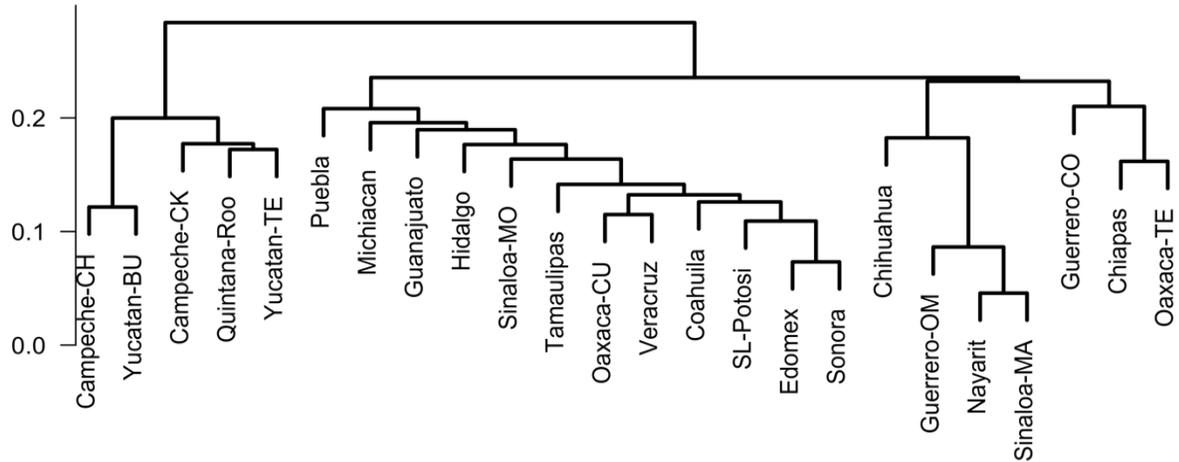


Figura 1. Dendrograma construido con los valores de diferenciación genética (distancias de Nei) entre pares de localidades analizadas de *C. moschata* utilizando 1,797 SNPs neutrales.

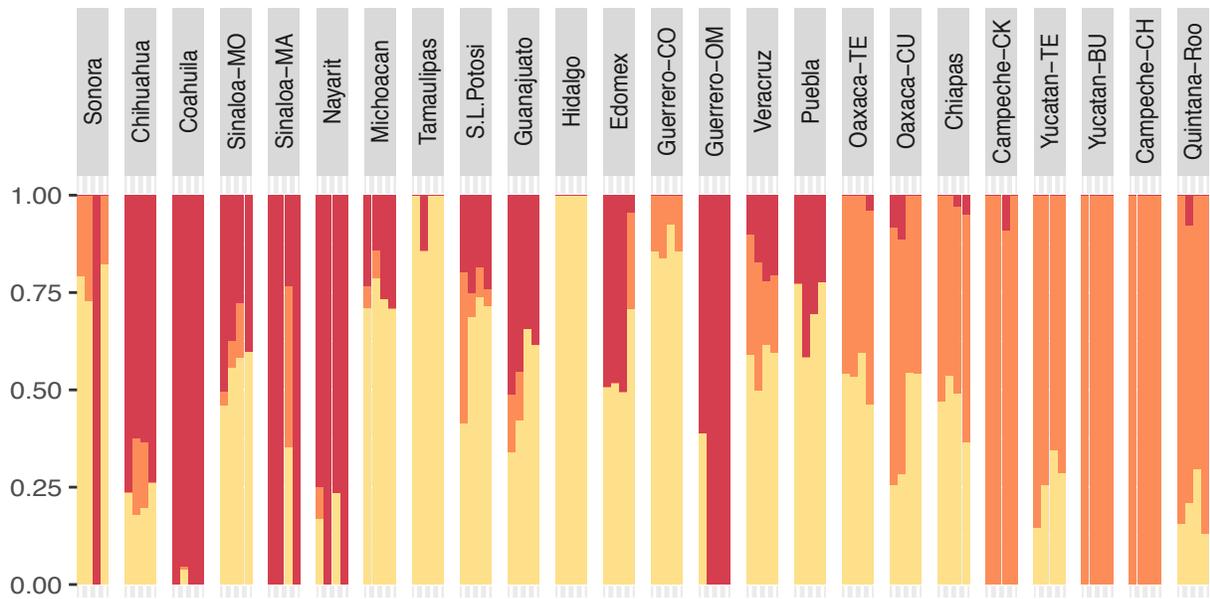


Figura 2. Análisis de Admixture de 24 localidades de *Cucurbita moschata* usando 1,797 SNPs neutrales. De acuerdo con una prueba de *Cross-validation* se identificó que el número óptimo de grupos genéticos es tres ($K = 3$).

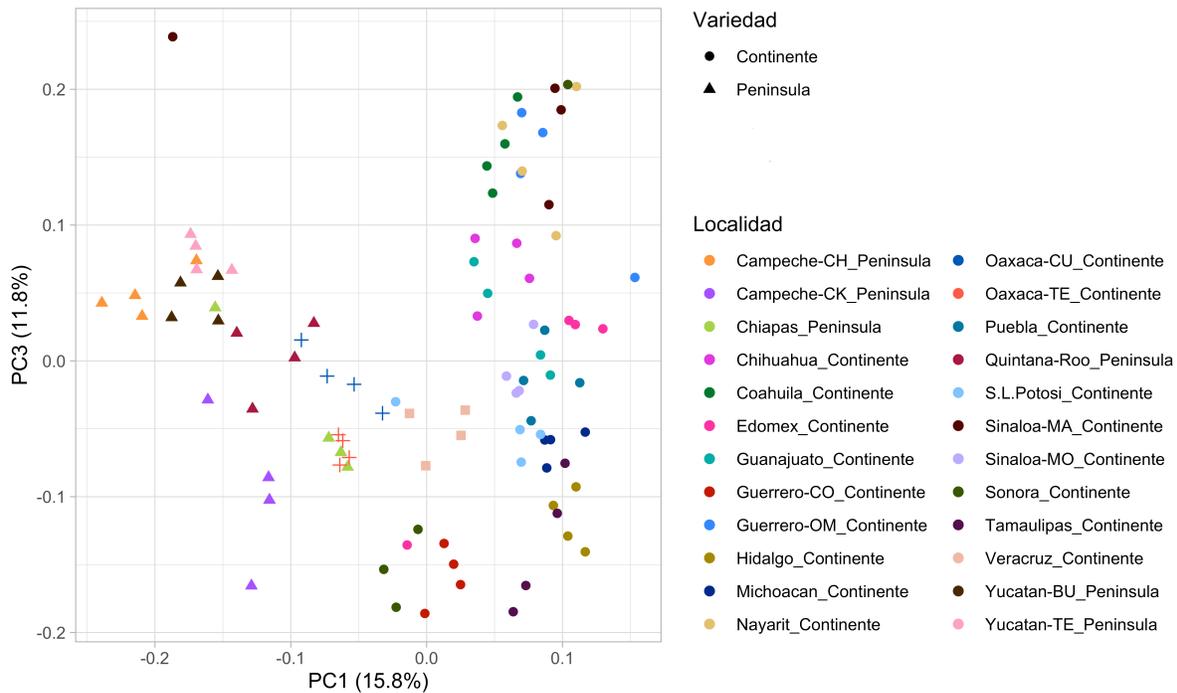


Figura 3. Análisis de componentes principales (PCA) para 24 localidades de *C. moschata* usando 1,797 SNPs neutrales. Los individuos de cada localidad fueron discriminados de acuerdo con las variedades descritas por Hernández-Rosales et al. (2020) usando loci de microsatélites nucleares: *a*) Variedades de la Península de Yucatán (Campeche-CK y CH, Yucatán-TE y BU y Quintana-Roo) y *b*) Variedades del Continente (el resto de las localidades de México). Las localidades de Oaxaca (TE y CU) se distinguen con un +; los individuos de Veracruz resaltan con una forma de □.

La prueba de Bayescan identificó 295 loci con señales de selección, de los 9,720 SNPs analizados (Figura 4). De acuerdo con la interpretación de la escala de Jeffrey (1961), 138 SNPs tienen una señal fuerte de selección, 88 SNPs muy fuerte y 69 SNPs son loci candidatos decisivos (Figura 4). De los 69 loci candidatos decisivos, 18 sugieren selección balanceadora y 51 selección direccional. Los puntos que se ubican antes del límite del criterio antes mencionado, representan loci cuyas frecuencias se pueden explicar mediante procesos demográficos, es decir, son loci sin señales de selección. Finalmente se detectaron 14 loci con valores de $\log(PO)$ por arriba de tres.

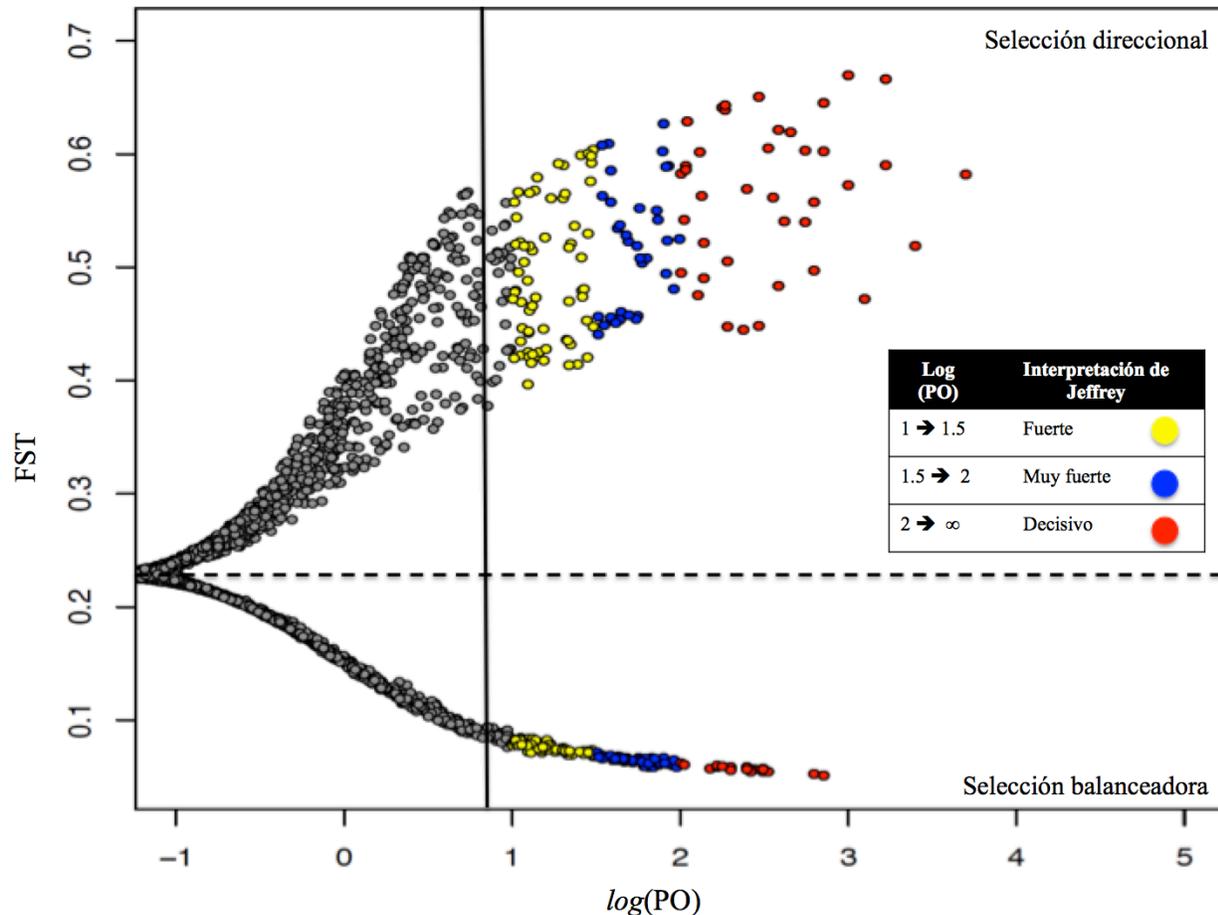


Figura 4. Prueba *Outliers* de 9,720 SNPs implementado el programa BayeScan para identificar loci con señales de selección de *C. moschata*. La línea paralela al eje Y representa el criterio de decisión para considerar loci candidatos, se consideró el 10% como falsos positivos. Los puntos por encima de la línea punteada se consideran polimorfismos bajo selección direccional y local mientras que los que se encuentran por debajo se consideran bajo selección balanceadora.

Discusión.

La preparación de bibliotecas con la metodología tGBS que implementamos permitió obtener fragmentos secuenciados más grandes y gracias a los cebadores degenerados, hay un elevado número de lecturas disponibles para la identificación de SNPs. La plataforma *Ion Torrent* genera secuencias con baja-intermedia calidad PHRED por lo que la metodología implementada debió ser más laxa en los criterios para los filtros de calidad y así evitar perder demasiada información. Con el *pipeline* implementado se obtuvieron ~1,700 SNPs bien identificados y sin errores de secuenciación.

La diversidad genética de las variedades locales de México parece ser baja ($H_E = 0.1324$) respecto al estudio de Lee et al. (2021). En este trabajo analizaron 2,071 SNPs que se obtuvieron

a partir de la plataforma *HiSeq 2000* (Illumina) de 610 accesiones provenientes de 42 países (África, América, Asia y Europa). Ellos reportaron una heterocigosis promedio para la especie de 0.281. Incluso considerando el grupo donde se asignaron las accesiones de América (que incluye alrededor de 117 muestras de México; $H_E = 0.315$) nuestros resultados señalan que *C. moschata* en México presenta baja diversidad genética. El mayor contraste entre ambos estudios se debe a que en el trabajo de Lee et al. (2021) las muestras de México (~110 accesiones) se analizan en conjunto con las accesiones del América desde el norte hasta el sur (ya que conforman un grupo genético bien definido), mientras que nuestro análisis se realizó bajo un enfoque a nivel local. También debe considerarse que el número de SNPs de nuestro estudio es distinto que el de Lee et al. (2021). No obstante, nuestros resultados indican que los patrones de diversidad genética entre las especies domesticadas del género *Cucurbita* se mantiene independientemente del tipo de marcador molecular. Usando la misma metodología que se implementó para obtener los SNPs de *C. moschata* (tGBS), la especie con la mayor diversidad genética es *C. pepo* ($H_E = 0.1627$ [2,061 SNPs]; Martínez-González et al., 2021), *C. moschata* es la especie que presenta los valores intermedios y *C. argyrosperma* presenta la menor diversidad genética ($H_E = 0.0940$ [2,861 SNPs]; Barrera-Redondo et al., 2021).

Contrario a los patrones de estructura genética obtenidos con loci de núcleo y mitocondria reportados por Hernández-Rosales et al (2020), los SNPs no reflejan estructura genética asociada a elevación o formas de frutos. Esto puede deberse en gran medida al tamaño de muestra, ya que con el objetivo de identificar principalmente genes con señales de selección se prefirió analizar la especie a lo largo de su distribución natural en México, así que por cada localidad se analizaron pocos individuos (solo cuatro por muestra) los cuales tal vez no representan los niveles de diversidad genética a nivel poblacional, así como los patrones de estructura genética. Posiblemente los SNPs no detectan los patrones genéticos previamente reportados ya que, a diferencia de los microsatélites, al obtener los estadísticos básicos de diversidad los polimorfismos de un solo par de bases no son tan sensibles al número de individuos muestreados (Aguirre-Liguori et al., 2020). Finalmente, para hacer comparaciones adecuadas entre estudios y corroborar si la elevación genera los patrones genéticos anteriormente reportados, así como contrarrestar el efecto del tamaño de muestra y verificar los niveles de diversidad es necesario incrementar el número de poblaciones más que el número de individuos o marcadores genéticos

(Aguirre-Liguori et al., 2020). Sin embargo, no debe descartarse la posibilidad de implementar análisis más adecuado que se ajusten a nuestro muestreo.

Ninguna localidad analizada de *C. moschata* presentó valores significativos de deficiencia o exceso de heterocigos (F_{IS}). Las tres especies de calabaza domesticada mencionadas anteriormente exhiben deficiencia de heterocigos, siendo *C. moschata* la especie con el valor de endogamia más alto ($F_{IS} = 0.2423$), seguido de *C. pepo* ($F_{IS} = 0.1648$; Martínez, 2019) mientras que *C. argyrosperma* tiene un valor más cercano al equilibrio ($F_{IS} = 0.0340$; Barrera-Redondo et al., 2020). De acuerdo a simulaciones y datos empíricos de microsatélites y SNPs, la F_{IS} es sensible a tamaños de muestra (Aguirre-Liguori et al., 2020). Las variedades locales o razas criollas de calabazas son comúnmente cultivadas a baja escala lo que favorece la polinización cruzada entre individuos cercanamente emparentados, esto promueve que los niveles de endogamia aumenten y por lo tanto decrezca la diversidad genética (Casas y Parra, 2016).

En términos de la diferenciación y estructura genética, las localidades analizadas de *C. moschata* si muestran el mismo patrón previamente descrito con loci de microsatélites nucleares asociado al Istmo de Tehuantepec que funge como barrera al flujo genético (Hernández-Rosales, et al., 2020). Las muestras de la Península de Yucatán están genéticamente bien diferenciadas del resto de las muestras de todo el país, para determinar si la estructura genética se correlaciona con el ambiente y/o responde a los criterios de selección y de manejo por parte de los agricultores es necesario desarrollar estudios etnobotánicos y ecológicos.

Finalmente, aunque se detectaron loci con señales de selección, se debe considerar que la prueba que implementa BayeScan es susceptible a arrojar falsos positivos, es por ello que necesario hacer pruebas de correlación con datos ambientales como la que implementa Bayescenv (de Villemereuil y Gaggiotti, 2015). Una vez identificados los loci con señales de selección, es importante realizar experimentos que permitan validar dichos genes.

Capítulo 4: ESTUDIO ETNOBOTÁNICO DE ESPECIES DOMESTICADAS DEL GÉNERO *Cucurbita* EN PERÚ.

Artículo de investigación: Barrera-Redondo, J., H. S. Hernández-Rosales, V. Cañedo-Torres, K. Aréstegui-Alegría, J. Torres-Guevara, F. Parra, I. Torres-García, y A. Casas. 2020. Landrace diversity and local selection criteria of domesticated squashes and gourds (*Cucurbita*) in the Central Andean mountain range of Peru: Tomayquichua, Huánuco. *Botanical Sciences*. 98: 101-116.

Perú carecen de registros adecuado de las variedades locales de las calabazas domesticadas y de los criterios de selección de los agricultores locales quienes modelan la diversidad genética de estos recursos. Por lo anterior, se realizó un estudio etnobotánico con la finalidad de sumar conocimiento sobre la diversidad biológica y el manejo de las calabazas y zapallos en la región de Tomayquichua, Huánuco, Perú. Este trabajo responde a dos preguntas: *i*) ¿Cuáles son las variedades locales de *Cucurbita* que reconocen, promueven y conservan los agricultores locales? y *ii*) ¿Qué mecanismos promueven la diversidad de estas variedades? Se realizaron entrevistas semiestructuradas y se foto-documentaron las variedades de calabaza que se cultivan en la región. Encontramos que *C. ficifolia* (conocida en México como chilacayote) es una calabaza ampliamente cultivada y apreciada, su reservorio genético está depositado en la gran diversidad de variedades locales de esta región y no dudamos que sea así para el resto del país. Debido a la gran diversidad de formas que se encuentran en Perú, este trabajo sugiere que el origen de *C. ficifolia* podría ser alguna región Andina del Perú, sin embargo, es necesario llevar a cabo estudios moleculares para esclarecer esta hipótesis.

VARIETADES LOCALES Y CRITERIOS DE SELECCIÓN DE ESPECIES DOMESTICADAS DEL GÉNERO *CUCURBITA* (CUCURBITACEAE) EN LOS ANDES CENTRALES DEL PERÚ: TOMAYQUICHUA, HUÁNUCO

LANDRACE DIVERSITY AND LOCAL SELECTION CRITERIA OF DOMESTICATED SQUASHES AND GOURDS (*CUCURBITA*) IN THE CENTRAL ANDEAN MOUNTAIN RANGE OF PERU: TOMAYQUICHUA, HUÁNUCO

JOSUÉ BARRERA-REDONDO¹, HELENA S. HERNÁNDEZ-ROSALES¹, VERÓNICA CAÑEDO-TORRES², KATIA ARÉSTEGUI-ALEGRÍA³, JUAN TORRES-GUEVARA⁴, FABIOLA PARRA⁴, IGNACIO TORRES-GARCÍA⁵, ALEJANDRO CASAS^{6*}

¹Laboratorio de Evolución Molecular y Experimental, Departamento de Ecología Evolutiva, Instituto de Ecología, Universidad Nacional Autónoma de México, México

²Dirección General de Diversidad Biológica, Ministerio del Ambiente, Lima Perú.

³Basoinsa S.L. Sucursal Perú, Lima, Perú.

⁴Centro de Investigación en Zonas Áridas, Universidad Nacional Agraria La Molina, Lima, Perú

⁵Escuela Nacional de Estudios Superiores, Universidad Nacional Autónoma de México, Campus Morelia, Michoacán, México.

⁶Instituto de Investigaciones en Ecosistemas y Sustentabilidad, Universidad Nacional Autónoma de México, Campus Morelia, Morelia, Michoacán, México.

*Autor para correspondencia: acasas@cieco.unam.mx

Resumen

Antecedentes: Es importante entender las prácticas agrícolas y los criterios de selección que moldean la diversidad de variedades locales en plantas domesticadas. La región andina comparte con la región mesoamericana y Norteamérica los centros de origen de las cinco especies domesticadas de *Cucurbita*. Sin embargo, los estudios sobre diversidad biológica de estas especies son escasos en Perú en comparación con México.

Preguntas: ¿Cuáles son las variedades locales de *Cucurbita* que reconocen, promueven y conservan los agricultores en el Distrito de Tomayquichua? ¿Qué mecanismos promueven la diversidad de estas variedades?

Especies estudiadas: *Cucurbita ficifolia* Bouché, *Cucurbita moschata* Duchesne, *Cucurbita maxima* Duchesne.

Sitio y año de estudio: Tomayquichua, Huánuco, Perú (junio de 2018 y 2019).

Métodos: Se realizaron entrevistas semiestructuradas a 29 familias de agricultores y comerciantes para evaluar la diversidad de variedades locales y las prácticas de manejo de las especies del género *Cucurbita*.

Resultados: *C. ficifolia* es la cucurbitácea más diversa en Tomayquichua (13 variedades) seguido de *C. moschata* (10 variedades), las cuales se siembran en chacras y traspatios. Se reportan siete variedades de *C. maxima*, con el uso predominante de una variedad comercial en cultivos intensivos.

Conclusión: Las prácticas agrícolas tradicionales parecen mantener la diversidad local de *C. ficifolia* y *C. moschata*. Las variedades locales de *C. maxima* en Tomayquichua parecen ser desplazadas progresivamente por una variedad mejorada destinada a aumentar la producción mediante prácticas agrícolas intensivas. La alta diversidad biológica de *C. ficifolia* en Perú sugiere un origen andino, pero se requieren estudios genéticos adicionales para aclararlo.

Palabras clave: Centro de diversidad, criterios de selección, *Cucurbita* spp., variedades locales.

Abstract

Background: It is important to understand the agricultural practices and criteria of human selection that shape the diversity of landraces in domesticated plants. The Andean region shares with Mesoamerica and North America the areas of origin of the five domesticated species of *Cucurbita*. However, studies on the biological diversity of these species have been scarcer in Peru than in Mexico.

Questions: Which landrace varieties of *Cucurbita* are recognized, promoted and conserved by the farmers in the District of Tomayquichua? Which mechanisms promote the diversity of these varieties?

Studied species: *Cucurbita ficifolia* Bouché, *Cucurbita moschata* Duchesne, *Cucurbita maxima* Duchesne.

Study site, dates: Tomayquichua, Huánuco, Peru (June of 2018 and 2019).

Methods: We used semi-structured interviews to 29 families of farmers and merchants to assess the diversity of landrace varieties and agricultural management practices of *Cucurbita* species.

Results: *C. ficifolia* was the most diverse cucurbit in Tomayquichua (13 varieties) followed by *C. moschata* (10 varieties), both which are grown in chacras and backyards. We report seven varieties of *C. maxima*, with the predominant use of a commercial variety through intensive agricultural practices.

Conclusion: The traditional agricultural practices seem to maintain the landrace diversities of *C. ficifolia* and *C. moschata*. The landraces of *C. maxima* in Tomayquichua seem to be displaced progressively by an enhanced variety destined to increase yield using intensive agricultural practices. The high biological diversity of *C. ficifolia* in Peru suggests an Andean origin for this species, but future studies with genetic data are needed.

Key words: Center of diversity, *Cucurbita* spp., landrace diversity, management.

La domesticación es un proceso evolutivo en el que diferentes fuerzas dirigidas por humanos y otras causas naturales (selección, deriva génica, flujo génico, sistemas reproductivos) operan en contextos antropogénicos e influyen en la generación de divergencias morfo-fisiológicas y genéticas entre los organismos domesticados y sus parientes silvestres (Casas *et al.* 2007, Pickersgill 2018). A su vez, se trata de un proceso de evolución mutualista en el cual un organismo asume control sobre la reproducción y cuidado de otro organismo, de manera que se pueda aprovechar algún recurso de interés (Zeder 2015).

La región andina, particularmente la región que comprende a Ecuador y Perú es considerada uno de los centros de origen de plantas domesticadas más importantes del planeta (Vavilov 1926, Harlan 1992, Diamond 2002) y comparte con la región mesoamericana y Norteamérica los centros de origen y diversificación de las cinco especies domesticadas de *Cucurbita* (Cucurbitaceae; Sanjur *et al.* 2002, Zheng *et al.* 2013).

El género *Cucurbita* es nativo del continente americano y es mayormente reconocido por sus especies domesticadas, comúnmente llamadas calabazas y zapallos. Comprende 20 taxa herbáceos, de 12 a 15 especies de las cuales cinco han sido domesticadas (Paris 2016, Lira *et al.* 2016). Las calabazas cultivadas son parte importante de los sistemas de policultivo tradicionales (e.g. *milpa* en México, *chacras* en Perú), las cuales mitigan la erosión del suelo y mantienen la humedad gracias a la amplitud de sus hojas que logran una cobertura significativa de los suelos, además de reducir el establecimiento y crecimiento de plantas arvenses (Lozada-Aranda *et al.* 2017). Adicionalmente, los cultivos de calabazas representan parte central de la alimentación en muchas regiones de América y el mundo (Eguiarte *et al.* 2018), con una producción de más de 24 millones de toneladas de alimento anualmente a nivel mundial (Paris 2016). Dada la importancia económica y evolutiva del género, se han realizado esfuerzos por comprender la domesticación independiente de sus especies y documentar la diversidad de las variedades locales (Paris 2016, Lira *et al.* 2016, Eguiarte *et al.* 2018).

Un aspecto importante para entender los procesos evolutivos que han modelado la variación de las especies domesticadas a lo largo del tiempo, consiste en documentar y analizar la diversidad morfológica y genética de las variedades locales a lo largo de su distribución natural (Casas *et al.* 1999, Gepts, 2014). En el caso del género *Cucurbita*, se han realizado esfuerzos para documentar su variación morfológica y genética en México (e.g. Andres 1990, Cerón-González *et al.* 2010, Sánchez-de la Vega *et al.* 2018). Sin embargo, los trabajos realizados en la región andina han sido relativamente escasos, documentando principalmente el norte de Perú (Delgado-Paredes *et al.* 2014). El estudio en ambas regiones es clave para analizar

temas aún controvertidos sobre el origen y difusión de estos cultivos (Lira *et al.* 2016). Las áreas montañosas tropicales de la región andina presentan escenarios ambientales marcadamente distintos en los gradientes altitudinales, los cuales constituyen contextos bioculturales excepcionalmente informativos para analizar los procesos de variación asociados a la domesticación y aquellos influidos por fuerzas evolutivas naturales (Velásquez-Milla *et al.* 2011).

Este estudio tiene como objetivo enriquecer el conocimiento sobre la diversidad biológica y el manejo de las calabazas y zapallos cultivados en el distrito de Tomayquichua, localizado en el departamento de Huánuco, provincia de Ambo, en la Sierra Central del Perú. Con base en esto, el estudio aspira a contribuir con información sobre las culturas andinas y sus interacciones con las cucurbitáceas. La información documentada parte del conocimiento y criterios de selección de los agricultores locales, quienes distinguen y conservan las distintas variedades locales de calabazas que se aprecian en esta región andina. Existen antecedentes que revelan una mayor diversidad biológica de plantas cultivadas en sistemas tradicionales comparados con los sistemas intensivos en distintas especies en todo el planeta, debido principalmente al papel de los agricultores en promover y conservar la diversidad de variedades locales (Brush 1986, Louette *et al.* 1997, Bellon & Berthaud 2004, Jarvis *et al.* 2008, Veteläinen *et al.* 2009). Debido a lo anterior, en este estudio se esperaba encontrar una mayor diversidad biológica de calabazas y zapallos que se cultivan bajo sistemas de producción agrícola tradicionales con respecto a los sistemas de cultivo intensivo.

Materiales y métodos

Área de estudio. El estudio se realizó en el distrito de Tomayquichua, dentro del Departamento de Huánuco, provincia de Ambo, Perú (Figura 1). La mayor parte de los datos colectados provienen de la Microcuenca de Warmiragra, entre 10° 02' 00" - 10° 05' 00" latitud sur y 76° 08' 00" - 76° 13' 00" longitud oeste (CCTA 2011). El área de estudio abarca los poblados de Armatanga y Chinchobamba en la zona media de la cuenca, perteneciente a la región natural denominada "quechua" que va entre 2,300 a 3,500 m, y el poblado de Tomayquichua en la zona baja de la cuenca, perteneciente a la región natural denominada "yunga" entre 1,000 a 2,300 m (Pulgar-Vidal 1979). Se realizó una colecta adicional de datos en el mercado local de la ciudad de Huánuco, ubicada en el centro del Distrito. De acuerdo con un censo realizado en 2007, se estima que el Distrito de Tomayquichua está conformado por alrededor de 3,500 habitantes (INEI 2007). La lengua local en Tomayquichua es el quechua y la segunda lengua es el español, siendo casi todos los

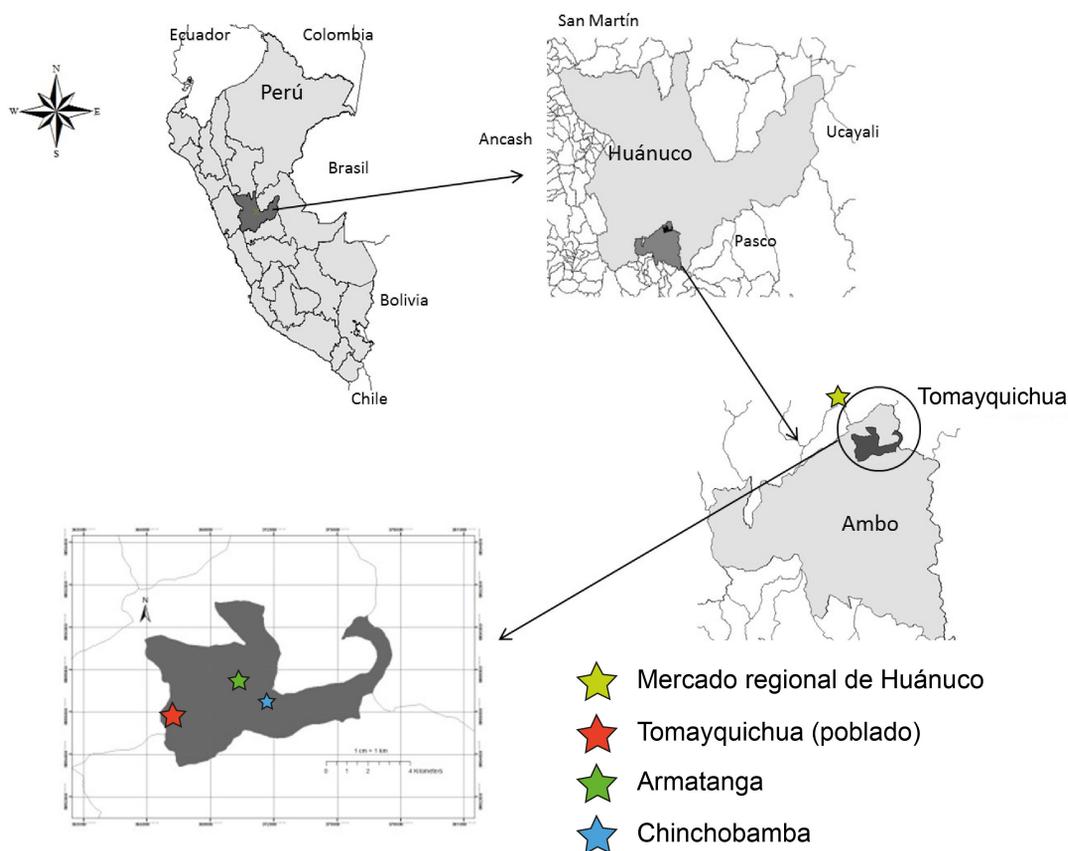


Figura 1. Mapa del sitio de estudio. Localización de Tomayquichua, el mercado de la ciudad de Huánuco y las comunidades estudiadas en la sierra central de los Andes peruanos (estrellas).

pobladores bilingües. En la zona baja de la cuenca se encuentra la capital del distrito (poblado de Tomayquichua), donde se realizan diversas actividades socioeconómicas primarias, secundarias y terciarias, mientras que en la zona media de la cuenca se realizan actividades agrícolas de supervivencia basadas en el autoconsumo, complementadas con intercambios comerciales y actos de reciprocidad entre los habitantes (CCTA 2011). Las principales plantas alimenticias que se siembran en la zona media de la cuenca son la papa (*Solanum tuberosum* L.), el maíz (*Zea mays* L.), el frijol (*Phaseolus vulgaris* L.), la caña de azúcar (*Saccharum officinarum* L.) y la alfalfa (*Medicago sativa* L.), mientras que la ganadería es una actividad marginal en la región (CCTA 2011). Los agroecosistemas en Tomayquichua pueden categorizarse en tres grupos: la chacra, el traspatio y el cultivo intensivo (Figura 2). La chacra, proveniente del quechua *chakra*, hace referencia a una parcela agrícola tradicional de dimensiones relativamente pequeñas dedicada a la producción de alimentos (Laimé-Ajacopa *et al.* 2007) y representa un sistema tradicional de producción agrícola en la región andina, el cual combina diferentes tipos de cultivos para hacer más eficiente el uso de recursos (*i.e.* agua, nutrientes,

luz solar) en el área destinada a la siembra (Torres *et al.* 2015). Los traspacios pueden considerarse como sistemas tradicionales de producción agrícola donde interactúa una elevada diversidad de animales y plantas al lado de edificaciones humanas, y cuya producción suele estar vinculada al autoconsumo (Mariaca-Méndez 2017). Por otro lado, el cultivo intensivo se refiere al uso de herramientas modernas que permiten la producción agrícola a escala industrial y favorecen el uso de variedades mejoradas para aumentar la producción (Casas & Parra 2016).

Especies estudiadas. Se caracterizaron tres especies domesticadas del género *Cucurbita* (Tabla 1): *Cucurbita ficifolia* Bouché, comúnmente conocida en Perú como “calabaza”, *C. moschata* Duchesne, comúnmente conocida como “shupe” o “loche”, y *C. maxima* Duchesne, conocida en Perú como “zapallo” (Lira *et al.* 2016). Tanto *C. maxima* como *C. moschata* son cultivos de gran importancia económica, la primera se siembra en regiones templadas a lo largo del planeta y la segunda se siembra en regiones tropicales, subtropicales e incluso semiáridas de distintos países (Paris 2016). Ambas especies se cultivan principalmente con fines de consumo alimenticio (*e.g.* fruto

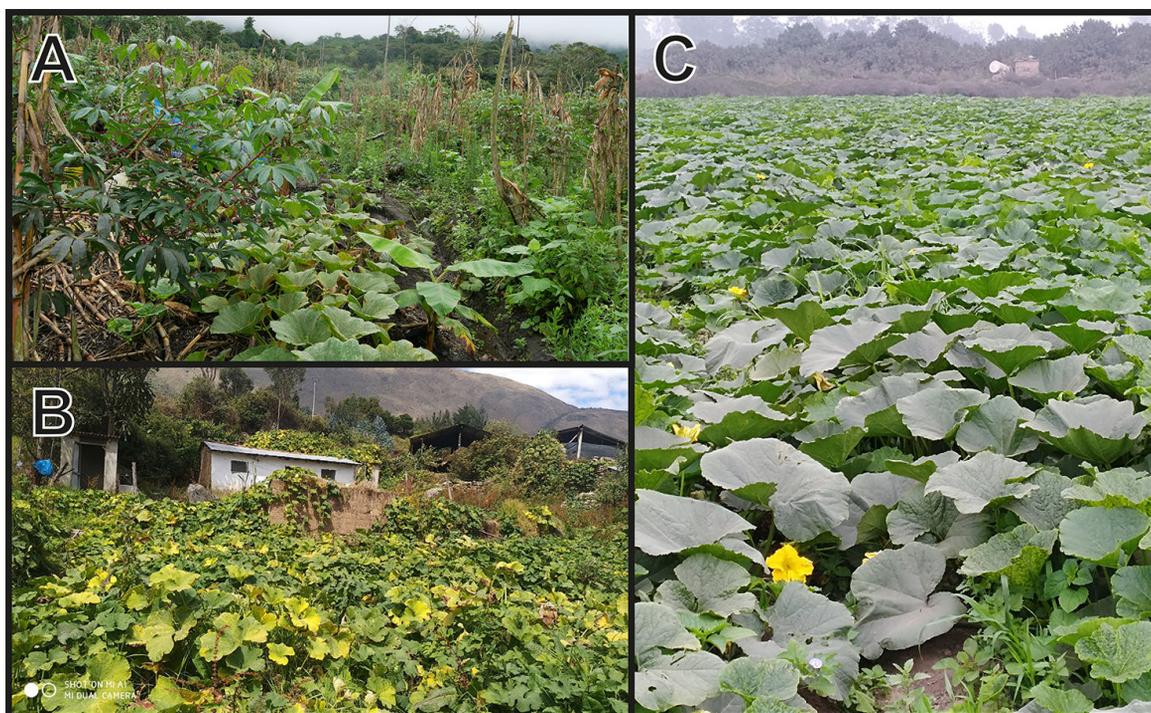


Figura 2. Sistemas de producción agrícola en Tomayquichua. A) chacra, B) traspatio, C) cultivo intensivo. Fotografía de la chacra proporcionada por el ministerio del ambiente del Perú y Aquagroinforest SAC.

maduro, fruto inmaduro, flores, tallos jóvenes, semillas), aunque también se pueden procesar para extraer sus saponinas en la fabricación de jabones. Las semillas pueden ser utilizadas como vermífugos o se les pueden extraer aceites comestibles e incluso los frutos se pueden utilizar con fines ornamentales (Casas *et al.* 1994, Díaz-Obregón *et al.* 2004, Marie-Magdeleine *et al.* 2009, Kim *et al.* 2012, Lira *et al.* 2016, Paris 2016). En cambio, *C. ficifolia* presenta un impacto económico menor, la cual se siembra de manera tradicional en regiones montañosas de México, Centroamérica y Sudamérica, principalmente para la elaboración de platillos y postres (Lira *et al.* 2016, Paris 2016). Los registros arqueológicos inequívocos más antiguos de *C. ficifolia* y *C. maxima* han sido encontrados en Perú y datan de aproximadamente 5,000 y 4,000 años respectivamente, mientras que el registro inequívoco más antiguo de *C. moschata* ha sido encontrado en México y data de aproximadamente 6,900 años de antigüedad (Whitaker 1981). Se ha propuesto que el centro de origen de *C. maxima* se encuentra en Sudamérica, específicamente en la región andina, dado que en esta zona se encuentra tanto su registro arqueológico más antiguo como su pariente silvestre (Whitaker 1981, Nee 1990). En cambio, el centro de origen de *C. moschata* es elusivo, del cual se ha hipotetizado un origen mesoamericano considerando los registros arqueológicos y su relación estrecha con la especie

mesoamericana *C. argyrosperma* (Whitaker 1981, Nee 1990, Lira *et al.* 2016), aunque la presencia de fitolitos y la abundancia de variedades locales de esta especie en Ecuador y Colombia también sugieren un posible centro de origen al norte de Sudamérica (Lira *et al.* 2016). Igualmente se desconoce el centro de origen de *C. ficifolia*, presentando evidencias lingüísticas que apuntan a un origen mesoamericano y evidencias arqueológicas que apuntan a un origen sudamericano (Whitaker 1981, Andres 1990, Lira *et al.* 2016).

Colecta de datos. Realizamos dos visitas al distrito de Tomayquichua del 27 al 29 de junio de 2018 y del 26 al 30 de junio de 2019. Se hizo una búsqueda aleatoria de parcelas, traspacios o áreas de cultivo intensivo que contuvieran especies del género *Cucurbita*, en cuyo caso se entrevistaron a los dueños del cultivo. Se llevaron a cabo entrevistas semiestructuradas a 27 familias de agricultores y 2 familias de comerciantes, de las cuales 15 corresponden a familias de la cuenca media de Warmiragra (poblados de Armatanga y Chinchobamba), 12 familias en la cuenca baja de Warmiragra (poblado de Tomayquichua) y dos familias de comerciantes en el mercado regional de la ciudad de Huánuco, el cual es un mercado establecido, con presencia de vendedores itinerantes en sus calles aledañas los días sábado y domingo. Al mercado, localizado a una altitud de

Tabla 1. Características morfológicas de las especies domesticadas del género *Cucurbita* cultivadas en el distrito de Tomayquichua, Huánuco, Perú (Modificado de [Eguiarte et al. 2018](#)).

	<i>C. FICIFOLIA</i>	<i>C. MAXIMA</i>	<i>C. MOSCHATA</i>
TALLO	Vigorosos, ligeramente angulosos.	Engrosados, redondeados o muy ligeramente angulosos.	Angulosos cuando verdes y aparecen surcos al secarse.
HOJA	Láminas con 3-5 lóbulos (redondeados), el central es más grande que los laterales, márgenes denticulados.	Tamaño variable. Ligeramente o nada lobada. Márgenes enteros a denticulados, ambas superficies pubescentes.	Grandes (> 20 cm). Con márgenes aserrados. Ligeramente lobadas (ovados o triangulares, ápice obtuso). El haz con manchas blancas en la intersección de las venas.
FLORES (MASCULINAS Y FEMENINAS)	No aromáticas. Color amarillo claro. Lóbulos agudos, márgenes enteros ondulados y doblándose hacia adentro. Ovario ovoide a elíptico.	Aromáticas. Color amarillo brillante. Lóbulos anchos, reflexos, márgenes crispados. Ovario de muy diversas formas, pubescente.	No aromáticas. Color amarillo-anaranjado. Lóbulos triangulares, agudos a acuminados, márgenes enteros ondulados y doblados hacia adentro. Ovario diverso.
FORMA DEL FRUTO	Globo, ovado o elíptico. Cáscara rígida, dura, gruesa, durable, sin costillas, densa y poco arrugada.	Diversas formas (turbaniformes). Algunos muy grandes (100 kg.). Lisos o con suaves costillas redondeada. Cáscara rígida o suave.	Formas y tamaños diversos conservando la forma del ovario. Lisos o comúnmente con costillas. Cáscara engrosada, rígida y durable o suave.
COLOR DEL FRUTO	Tres patrones: a) verde claro u oscuro, con o sin franjas longitudinales blancas. b) manchada de blanco o verde. c) blanca o crema.	Incluyen el azul-grisáceo, verde parduzco, rosado y rojo o rojizo-anaranjado.	Patrones de color muy variable que van del verde oscuro a claro. Liso o con manchas crema. Cubierta cerosa blanquecina.
PULPA	Notoriamente fibrosa. Blanca. Sabor dulce.	Ligeramente fibrosa. Diferentes tonalidades de amarillo a anaranjado. Sabor ligeramente suave.	No fibrosa de consistencia suave. Amarilla en distintas tonalidades a anaranjada, a veces verdosa. Sabor dulce.
SEMILLAS	Generalmente negras, aunque existen en baja frecuencia de color crema y pardo. Márgenes poco diferenciados del mismo color que la semilla.	Gruesas, elípticas infladas o tumescentes. Bronceadas (color canela) blancas a pardo claras. Lisas o muy poco estriadas. Márgenes redondeados diferenciados del color del centro.	Elípticas u ovado-elípticas. Blancas o pardo claras del centro y márgenes bien diferenciados en color y textura. Ápice obtuso a truncado.
DISTRIBUCIÓN NATURAL	Desde México hasta Argentina y Chile.	Sur de América (Argentina, Bolivia, Uruguay).	Mesoamérica hasta Sur de América.

1,800 m, acuden agricultores y comerciantes intermediarios de las regiones montañosas (comunidades con territorios arriba de los 4,500 m) y de tierras bajas (alrededor de 600 m) de áreas tropicales húmedas ([Figura 1](#)). En cada entrevista se consideró al núcleo familiar como unidad de información, donde un miembro de la familia fungía como el entrevistado principal y se permitía la intervención espontánea del resto de los miembros para complementar y corroborar la información recabada. De los entrevistados principales, 13 fueron hombres y 11 fueron mujeres, además de 5 entrevistas donde participaron múltiples miembros de la familia de manera equitativa. Las entrevistas nos permitieron registrar la siguiente información: i) Nombres

de las variedades conocidas por el entrevistado y características diagnósticas de la variedad; ii) sistema de siembra de las variedades (traspatio, chacra o intensivo); iii) usos de las variedades (*e.g.* alimentación, medicinal) y destino del producto (*e.g.* autoconsumo, alimentación del ganado, venta); iv) periodos de siembra y cosecha; v) origen de las semillas (heredadas, intercambio de semillas con vecinos o compra de semillas); vi) plantas asociadas al cultivo de las variedades; vii) criterios de selección de las variedades (consumo preferencial y/o preferencia de una variedad sobre otra, de acuerdo con sus usos). Adicionalmente se realizó una foto-documentación de las variedades registradas en el estudio.

Variedades locales de *Cucurbita* spp. en Tomayquichua, Perú.

Tabla 2. Variedades de *Cucurbita* spp. documentadas en el distrito de Tomayquichua, Huánuco, Perú. Se reportan las especies encontradas en la región, los usos que se les da a los cultivos, las formas que pueden adquirir los frutos, las variedades descritas por los entrevistados y las características de cada variedad.

Especie	Usos	Tipo de siembra	Formas	Variedades	Características de la variedad	
<i>Cucurbita ficifolia</i> Bouché	Principalmente autoconsumo. Frutos inmaduros consumidos en sopas y platillos “segundos”. Frutos maduros consumidos en postres. Frutos sobrantes como alimento para animales de traspatio. Semillas con uso medicinal (desparasitante).	Traspatio y chacra	Redonda (esférica)	<i>Verde</i>	Fruto maduro de color verde con franjas blancas. Semilla negra y pulpa blanca.	
				<i>Cushta / conchubano / muru</i>	Fruto maduro de color blanco con franjas verdes. Semilla negra y pulpa blanca.	
				<i>blanca</i>	Fruto maduro de color blanco. Semilla negra y pulpa blanca.	
				<i>Limón</i>	Fruto maduro de color verde brillante. Semilla negra y pulpa blanca.	
				Ovoide	<i>Limón</i>	Fruto maduro de color verde brillante. Semilla negra y pulpa blanca.
					<i>Blanca</i>	Fruto maduro de color blanco. Semilla negra y pulpa blanca.
			<i>Puru</i> (alargado)	<i>Naranja</i>	Fruto maduro de color naranja. Semilla negra y pulpa blanca.	
				<i>Costa</i>	Fruto maduro de color marmoleado blanco con verde. Semilla negra y pulpa blanca.	
				<i>Blanco / matish</i>	Fruto maduro color blanco crema. Semilla negra, pulpa blanca.	
				<i>Limón</i>	Fruto maduro de color verde brillante. Semilla negra y pulpa blanca.	
			<i>Tullu</i>	<i>Verde</i>	Fruto redondo, pequeño, de color verde oscuro con franjas blancas y exocarpo engrosado. Semilla negra y pulpa blanca. Cocción rápida en el horno.	
				<i>Limón</i>	Fruto redondo, pequeño, de color verde brillante con exocarpo engrosado. Semilla negra y pulpa blanca. Cocción rápida en el horno.	
<i>Zapallo calabaza</i>		<i>Redonda</i>	Fruto verde con franjas color crema. Tamaño excepcionalmente grande. Semillas negras, pulpa blanca.			

Especie	Usos	Tipo de siembra	Formas	Variedades	Características de la variedad	
<i>Cucurbita moschata</i> Duchesne	Autoconsumo. Fruto inmaduro para sopas y platillos “segundos”. Frutos maduros para postres.	Traspatio y chacra	<i>Redonda</i> (semiesférica)	<i>S hirka / plomo</i>	Fruto blanco jaspeado con verde, semilla blanca. Pulpa amarilla.	
				<i>Shupe</i>	Fruto anaranjado con franjas verdes, semilla blanca y pulpa anaranjada.	
				<i>Blanca</i>	Fruto blanco (color crema) con semilla blanca, y pulpa amarilla.	
				<i>Verde</i>	Fruto verde con franjas color crema, semilla blanca, y pulpa anaranjada	
				<i>Coyotita</i> (piriforme)	<i>Shupe cushta</i>	Fruto color anaranjado con franjas verdes. Semillas blancas-doradas con borde definido de color dorado. Pulpa anaranjada.
					<i>Verde cushta</i>	Fruto verde con franjas blancas. Semillas blancas-doradas con borde definido de color dorado. Pulpa anaranjada.
				<i>Alargada</i>	<i>Shirka / plomo</i>	Fruto blanco jaspeado con verde, semilla blanca y pulpa amarilla.
					<i>Shupe</i>	Fruto anaranjado con franjas verdes, semilla blanca, pulpa anaranjada.
					<i>Blanca</i>	Fruto blanco (color crema), semilla blanca, pulpa amarilla.
					<i>Verde</i>	Fruto verde con franjas blancas. Semillas blancas-doradas con borde definido de color dorado. Pulpa naranja.
<i>Cucurbita maxima</i> Duchesne	Venta para consumo humano.	Intensiva	Redonda	<i>Macre</i> (mejorada)	Fruto maduro color verde con pulpa amarilla y semilla café.	
	Autoconsumo familiar	Chacra	Redonda	<i>Verde</i>	Fruto maduro color verde con pulpa amarilla y semilla café.	
				<i>Pinto</i>	Fruto verde, con pintas claras y de color rojizo. Pulpa amarilla y semilla café.	
				<i>Blanca</i>	Fruto blanco (color crema), pulpa amarilla y semilla café.	
				<i>Plomo</i>	Fruto color plateado, con franjas verdes, pulpa amarilla y semilla café.	
				<i>Alargada</i>	<i>Verde</i>	Fruto maduro color verde con pulpa amarilla y semilla café.
					<i>Plomo</i>	Fruto color plateado, con franjas verdes, pulpa amarilla y semilla café.

Resultados

Variedades locales de calabazas y zapallos (Cucurbita spp.). Se registraron 13 variedades distintas de *C. ficifolia* (Tabla 2, Figura 3), todas distribuidas en la cuenca media de Warmiragra (hasta los 3,400 m). Cada variedad está caracterizada por uno de los seis colores que adquiere el fruto maduro, mientras que las demás características de la especie son compartidas (semilla negra, fruto maduro con pulpa blanca, zarcillos con tres a cinco ramificaciones; Tabla 1). Además, se reportan tres formas distintas para el fruto maduro de *C. ficifolia* (redonda, ovalada y alargada o “puru”; Figura 4) y tres categorías de tamaño, una intermedia que comprende la mayor parte de estas calabazas, así como una variedad pequeña o “tullu” y una variedad de tamaño grande denominada “zapallo calabaza”.

Se registraron 10 variedades de *C. moschata* a partir de la combinación de tres formas distintas (esférica o “redonda”, piriforme o “coyotita” y alargada) y cuatro colores en el exocarpo maduro (Tabla 2, Figura 5). También se identificaron dos colores distintos de pulpa en el fruto maduro (amarilla y anaranjada) y dos colores de semilla (blanca y blanca-dorada con borde dorado definido). Se reporta una variedad denominada “shirka” en la cuenca media de Warmiragra. Los frutos de esta variedad son similares en forma y tamaño a los frutos de *C. ficifolia*

mientras que sus semillas son bastante más parecidas en color a las de *C. moschata* (Tabla 2); estas características sugieren posibles eventos de hibridación entre estas dos especies (*moschata* × *ficifolia*). También se reporta una variedad “shupe” en la zona baja de la Cuenca de Warmiragra, la cual se caracteriza por su color anaranjado con franjas verdes y presenta tres formas distintas (redonda, “coyotita” o “cushta” y alargada; Figura 5B). A diferencia de las otras especies, se encontraron solo siete variedades de *C. maxima*, conformadas a partir de dos formas distintas y cuatro patrones de color en el fruto maduro (Tabla 2). La variedad con mayor distribución en la región se le denomina “macre”, una variedad mejorada que se caracteriza por sus frutos grandes (> 50 kg) (Figura 6).

Caracterización del agroecosistema de producción. El periodo de siembra de *C. ficifolia* es variable y abarca los meses de septiembre a diciembre. Por otro lado, el periodo de cosecha de los frutos maduros abarca los meses de junio y julio. El cultivo de *C. ficifolia* se encuentra tanto en traspatios como en chacras, acompañado usualmente por maíz, frijol y haba (*Vicia faba* L.). No obstante, también se reportó su siembra en menor medida junto con papa y tarwi (*Lupinus mutabilis* Sweet). En este último caso, las semillas de *C. ficifolia* se siembran en almácigos y los individuos se trasplantan a la chacra cuando son juveniles, lo cual evita su

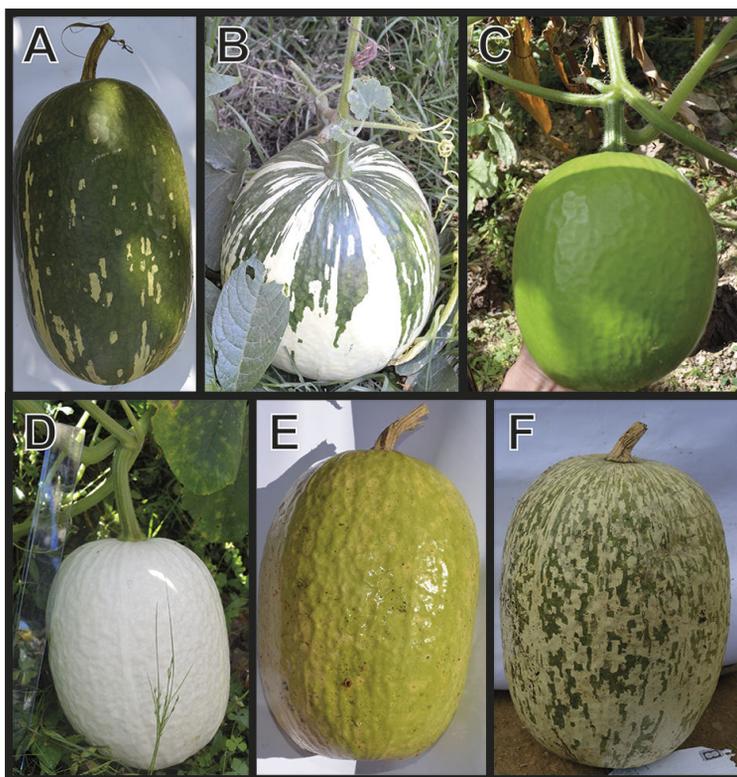


Figura 3. Variedades de *Cucurbita ficifolia* reportadas en el distrito de Tomayquichua, Huánuco, Perú. A) Variedad verde, B) variedad cushta/conchubano, C) variedad limón, D) variedad blanca, E) variedad naranja F) y variedad costa.

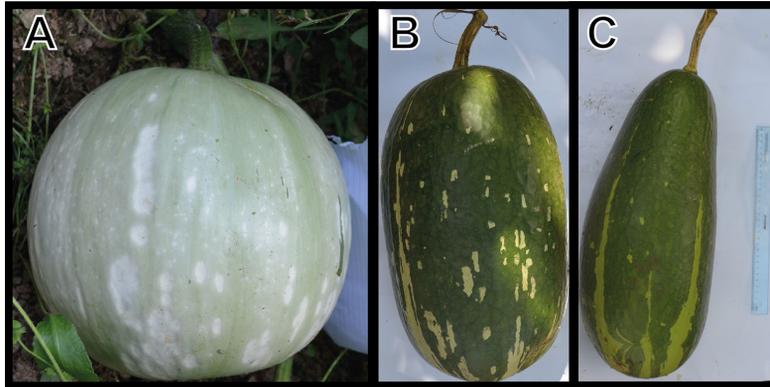


Figura 4. Formas del fruto de *Cucurbita ficifolia* reportadas en el distrito de Tomayquichua, Huánuco, Perú. A) Forma redonda (esférica), B) forma ovoide C) y forma “puru” (alargada).

desplazamiento por otras variedades. El cultivo de esta especie es muy común en las comunidades estudiadas. Aunque la superficie en que se siembra varía año con año, en 2019 se registró su cultivo en Armatanga en cerca de 80 predios, mientras que en Chinchobamba en aproximadamente 40 predios, en promedio con una superficie de 210 m² por predio. En estos predios se cultivan calabazas en manchones de aproximadamente 30 m². Además, se cultivan chacras de maíz, que constituyen aproximadamente 35 % de las chacras de estos poblados; en éstas la asociación de calabaza y maíz se practica en aproximadamente 40 % (se registraron 17 chacras de maíz asociado con calabaza en 2019, aunque es variable entre los años), cada una de ellas de aproximadamente 0.75 hectáreas. En estas comunidades se practican diversos cultivos dentro de las chacras como la papa, el tarwi, el maíz, las arvejas (*Pisum sativum* L.), el ajonjolí (*Sesamum indicum* L.) y el rocoto (*Capsicum pubescens* Ruiz & Pav.); de manera que entre los traspatios y las chacras de cultivos en asociación con calabaza sumaron aproximadamente 17 hectáreas cultivadas principalmente con *C. ficifolia* y en menor medida con *C. moschata* en 2019, aunque la superficie de siembra destinada a los traspatios es relativamente menor al de las chacras. Además de la siembra activa de *C. ficifolia* por parte de los agricultores, las semillas de algunos frutos que se abandonan y descomponen en los traspatios germinan y sus plantas restablecen las matas de calabazas, las cuales se pueden encontrar extendidas por el piso, las paredes de las casas, las cercas o estructuras de soporte construidas *ad hoc* para las calabazas. El periodo de siembra de *C. moschata* se lleva a cabo en el mes octubre y los frutos maduros se cosechan a partir del mes de junio. Su siembra se realiza tanto en traspatios como en chacras, de manera similar a *C. ficifolia*, y se suele acompañar de maíz, frijol, haba y papa. En estos predios es notoria la predominancia de *C. ficifolia*, aunque la proporción de superficie ocupada por cada especie no fue evaluada en este estudio. El cultivo de *C. maxima* comprende un periodo de siembra entre febrero y marzo,

mientras que el periodo de cosecha de frutos maduros abarca los meses de junio y julio. El sistema de producción de zapallo que predomina en la región es intensivo, en monocultivos o en asociación con maíz, recibiendo riego y aplicación de agroquímicos. Sin embargo, también se encuentran en chacras bajo sistemas de policultivo, principalmente asociadas a maíz, en los cuales es posible identificar seis variedades locales cuyas semillas se conservan por las familias de agricultores generación tras generación. La superficie bajo manejo intensivo comprende el valle aluvial del río Huallaga, el cual comprende cerca de 3,000 hectáreas en el territorio de Tomayquichua, pero solamente entre el 5 y 8 % de la superficie es cultivada con zapallo, aproximadamente 30 hectáreas en chacras en asociación con maíz y alrededor de 200 hectáreas como monocultivo intensivo.

Manejo y selección de las variedades locales. La producción de *C. ficifolia* está destinada al autoconsumo y su principal uso es alimenticio. Los frutos inmaduros se consumen en guisos y sopas, mientras que la pulpa de los frutos maduros se utiliza para la preparación de postres como la mazamorra (postre elaborado a base de calabaza o maíz, el cual contiene leche, azúcar y especias). Los agricultores reconocen los frutos maduros al observar un cambio en la coloración del exocarpo y un aumento en el grosor y rigidez de éste, lo cual ocurre un tiempo indeterminado previo al periodo de cosecha. Previo a este cambio, los frutos poseen una coloración verde uniforme y son considerados como inmaduros. En caso de haber una producción abundante de frutos, se asigna una porción a la alimentación de animales de traspatio como cerdos (*Sus scrofa* ssp. *domestica* L.) y perros (*Canis lupus* ssp. *familiaris* L.). En esta región, las semillas se consumen para tratar enfermedades gastrointestinales, principalmente como vermífugos. Las semillas han sido mantenidas a lo largo de dos o más generaciones de cada familia entrevistada. Las semillas de varios frutos y variedades se dejan secar al sol,

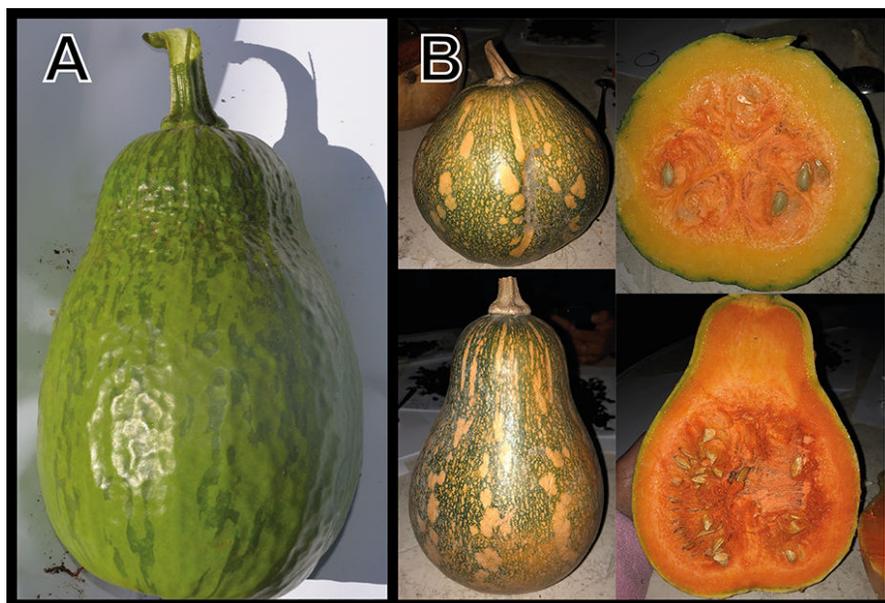


Figura 5. Variedades y formas de *Cucurbita moschata* reportadas en el distrito de Tomayquichua, Huánuco, Perú. A) Variedad *shirka*. B) Variedad *shupe* con las formas redonda (arriba) y “*coyotita*” (abajo).

usualmente se mezclan, se almacenan alejadas de la humedad y se guardan por un periodo entre 9 a 10 meses para su siembra en la siguiente temporada de cultivo. Algunas familias manifestaron que las semillas se pueden almacenar hasta dos temporadas (24 meses) antes de volver a sembrarlas.

Las entrevistas efectuadas indican criterios de selección similares en todas las variedades de tamaño intermedio de *C. ficifolia*, es decir, que no hay preferencia por cultivar mayoritariamente una u otra de estas variedades. Nuestros resultados indican que todas las variedades de tamaño intermedio son igualmente apreciadas por los agricultores locales debido a que su uso alimenticio, sus características organolépticas y sus características de cocción son similares, independientemente de su forma o su color. Sin embargo, se encontró una preferencia sobre la variedad denominada “*tullu*”, un término que hace alusión a frutos pequeños que son de rápida cocción, en los cuales el fruto entero se cuece en horno. También se hizo referencia a que las calabazas “*tullu*” poseen una corteza dura y gruesa, un atributo igualmente apreciado que permite colocar los frutos directamente al horno para su cocción. Aunque algunas personas entrevistadas desconocían atributos particulares de esta variedad, la mayor parte de las familias entrevistadas identificaron por consenso las características referidas. También se identificó una variedad denominada “*zapallo calabaza*”, la cual es de mayor tamaño que el resto de las variedades, llegando a pesar hasta 20 kg. Esta variedad posee atributos similares de sabor y textura a las variedades de tamaño intermedio, pero se favorece por su tamaño excepcionalmente grande.



Figura 6. Variedad mejorada *macre* de *Cucurbita maxima* reportada en el distrito de Tomayquichua, Huánuco, Perú.

La producción de *C. moschata* también está destinada al autoconsumo y en muy baja medida para su venta. Los frutos inmaduros se consumen en platillos principales y sopas, mientras que la pulpa de los frutos maduros se utiliza para elaborar mazamorra. Las semillas pueden o no consumirse junto con la pulpa del fruto al preparar la mazamorra. La variedad “*coyotita*” de *C. moschata* presenta una protuberancia que la vuelve más deseable sobre la forma redonda, ya que dicha protuberancia no contiene semillas y se conforma exclusivamente de pulpa, la cual es apreciada al momento de consumirse. La variedad denominada “*cushta*” también es apreciada por estas razones, la cual presenta una forma de guitarra o piriforme. Al igual que *C. ficifolia*, las semillas de *C. moschata* son

conservadas por las familias a través de dos o más generaciones, con eventuales intercambios con vecinos.

La producción intensiva de la variedad “*macre*” de *C. maxima* se destina para su venta en los mercados, principalmente en el mercado de Lima. Los agricultores compran las semillas cada temporada de cultivo en las tiendas de productos agrícolas, regionalmente denominadas “agros”. Los productores de zapallo “*macre*” mencionan que nunca utilizan las semillas obtenidas de los frutos cosechados porque el rendimiento de la producción total es menor. Para el resto de las variedades de esta especie, cultivadas en chacras tradicionales, los agricultores reportaron guardar las semillas generación tras generación y eventualmente intercambiarlas entre familiares y amigos. Generalmente se siembra a finales de agosto o principios de septiembre y los sistemas intensivos pueden producir alrededor de 20 toneladas por hectárea. Los agricultores atribuyen como el principal problema para la producción de *C. maxima* a un virus que se manifiesta en las hojas.

Discusión

Manejo de especies del género Cucurbita en Tomayquichua. Al igual que lo reportado en este estudio, los agricultores en Cuba reconocen la temporada de cosecha de frutos maduros de *Cucurbita* al observar un cambio en los patrones de coloración del exocarpo y un aumento en su rigidez, el cual coincide con la desecación del follaje ([Biblioteca ACTAF 2007](#)). Adicionalmente, se reconoce al periodo de cosecha de los frutos inmaduros a la mitad del tiempo o tres cuartas partes del tiempo previo a alcanzar el estado de madurez ([Biblioteca ACTAF 2007](#)). Los tiempos de cosecha de frutos maduros de *Cucurbita* reportados en este estudio difieren en algunos casos con lo observado en otros estudios. Por ejemplo, la temporada de cosecha de frutos maduros de *C. moschata* en Tomayquichua ocurre alrededor de los nueve meses posterior a la temporada de siembra, mientras que otros estudios realizados en Cuba reportan tiempos de cosecha de *C. moschata* entre los tres y cinco meses posteriores al periodo de siembra ([Biblioteca ACTAF 2007](#)). Por otro lado, los tiempos de cosecha reportados para *C. maxima* en este estudio ocurre alrededor de los tres a cinco meses, de manera coincidente con los periodos de cosecha de *C. moschata* reportados en Cuba ([Biblioteca ACTAF 2007](#)), pero distinto a los periodos de cosecha reportados para *C. maxima* en Nueva Zelanda, los cuales rondan de los 21 a los 40 días posteriores al periodo de siembra, ocurriendo una mayor incidencia de plagas cuando la cosecha era tardía ([Hawthorne 1990](#)). A pesar de que los trabajos realizados en Nueva Zelanda no reportan diferencias en el patrón sugerido de cosecha a lo largo del año, estas diferencias en tiempos de cosecha podrían deberse a las diferencias climáticas entre Nueva Zelanda y Tomayquichua, las cuales pueden influir en la

susceptibilidad de las cosechas contra plagas por hongos ([Hawthorne 1990](#)).

Impacto del sistema de siembra sobre las variedades de Cucurbita. Tanto *C. ficifolia* como *C. moschata* se siembran bajo sistemas agrícolas tradicionales como los traspatios o huertas y las chacras ([Torres et al. 2015](#), [Mariaca-Méndez 2017](#)), lo cual parece repercutir positivamente en su diversidad, particularmente debido al cuidado y mantenimiento de las variedades, así como el prestigio de los productores asociado a la conservación de diversidad, de la misma forma en que ha sido reportado para otros taxa cultivados ([Brush 1986](#), [Louette et al. 1997](#), [Bellon & Berthaud 2004](#), [Jarvis et al. 2008](#), [Veteläinen et al. 2009](#), [Velásquez-Milla et al. 2011](#)). Dado que ambos taxa se cosechan con fines de autoconsumo, no se registraron sistemas de siembra intensiva, lo que también favorece el mantenimiento de las variedades locales ([Brush 1986](#), [Louette et al. 1997](#), [Bellon & Berthaud 2004](#)). Adicionalmente, los sistemas tradicionales suelen involucrar policultivos, los cuales propician el flujo genético entre distintas variedades de *Cucurbita* spp., fomentan la diversidad de los cultivos y dan pie a posibles eventos de hibridación, los cuales ocurren de forma natural dentro del género *Cucurbita* ([Whitaker & Bemis 1964](#), [Sánchez-de la Vega et al. 2018](#)), y que posiblemente hayan ocurrido para generar a la variedad “*shirka*”. Tanto *C. ficifolia* como *C. moschata* se siembran desde hace dos o más generaciones, dando seguimiento a los linajes de calabazas, por lo que se puede hablar de una práctica agrícola intergeneracional que da lugar a procesos de domesticación ([Zeder 2015](#)). Por otra parte, los agricultores de *C. máxima* combinan métodos semi-intensivos que permiten la conservación de variedades locales, así como sistemas intensivos, cuya producción se destina a la venta en el mercado y en los que las semillas deben adquirirse asimismo en el mercado. Aunque las variedades se mantienen, predomina el cultivo de la variedad “*macre*”. En comparación con otras especies, *C. maxima* presenta baja diversidad de variedades locales. Esto contrasta con la noción de que los centros de domesticación albergan la mayor cantidad de variación y formas endémicas para una especie ([Harlan 1992](#)), ya que el centro de origen de *C. maxima* se encuentra posiblemente en la región andina ([Whitaker 1981](#), [Nee 1990](#)). En cambio *C. moschata*, cuyo centro de origen se encuentra posiblemente en la región mesoamericana o al norte de Sudamérica ([Lira et al. 2016](#)) muestra una diversidad de variedades locales mayor a la de *C. maxima* tanto en Tomayquichua (cuatro patrones de color de fruto, tres formas del fruto, dos colores de pulpa y dos colores de semilla) como en el norte de Perú (de cinco a trece patrones de color de fruto, cuatro categorías de tamaño, cinco formas del fruto y seis colores de semilla;

[Delgado-Paredes et al. 2014](#)). Además, los patrones de diversidad morfológica al norte de Perú reportados por [Delgado-Paredes et al. \(2014\)](#) revelan una diversidad ligeramente mayor de *C. maxima* a la reportada en este estudio, a pesar de no haber realizado un muestreo sistemático para esta especie como de *C. ficifolia* y *C. moschata*, al presentar cuatro patrones de color y variación tanto en el tamaño del fruto como en el color y tamaño de las semillas, mientras que en Tomayquichua observamos cuatro patrones de color de fruto maduro y dos formas distintas en el fruto sin variedad en el color de la semilla. Por lo tanto, sospechamos que este patrón de diversidad puede deberse al desplazamiento de las variedades locales de *C. maxima* por la variedad mejorada, lo cual ha sido reportado como un factor que promueve la erosión genética en diferentes especies de plantas domesticadas ([Brush 1986](#), [Louette et al. 1997](#), [Bellon & Berthaud 2004](#), [Jarvis et al. 2008](#), [Veteläinen et al. 2009](#), [Velásquez-Milla et al. 2011](#)). Es importante realizar subsecuentes muestreos para verificar si el patrón de diversidad morfológica asociada a los sistemas agrícolas tradicionales reportados en este estudio también se observa en otras zonas de la región andina.

Usos y criterios de selección. Las tres especies estudiadas son aprovechadas principalmente por su fruto, tanto maduro como inmaduro, lo cual podría favorecer la selección de frutos con mayor pulpa cuando se consumen maduros y la selección de frutos de fácil cocción cuando se consumen inmaduros. El aprovechamiento de la pulpa es evidente en relación con el criterio de selección de las formas “*coyotita*” y “*shupe*” sobre la forma “*redonda*” en *C. moschata*, las cuales presentan una mayor proporción de pulpa por tamaño del fruto. Este aprovechamiento también es evidente en la variedad “*zapallo calabaza*” de *C. ficifolia*, la cual produce frutos de tamaño excepcionalmente grandes. El caso del fruto inmaduro es evidenciado por la preferencia de frutos de las variedades “*tullu*”, en las cuales se encontró un consenso sobre el tamaño pequeño y el exocarpo resistente como atributos de la variedad. El uso de semillas de calabaza con fines medicinales reportado en este estudio para *C. ficifolia* ha sido estudiado previamente para *C. maxima*, cuyas semillas son usadas en zonas rurales para eliminar parásitos y las cuales han demostrado tener propiedades antiparasitarias en estudios preclínicos ([Diaz-Obregón et al. 2004](#)). El efecto antiparasitario de las semillas de *Cucurbita* ha sido reportado en otros estudios realizados con *C. moschata* ([Casas et al. 1994](#), [Marie-Magdeleine et al. 2009](#)) y *C. pepo* ([Casas et al. 1994](#), [Grzybek et al. 2016](#)), lo cual sugiere que las semillas de *C. ficifolia* podrían generar resultados prometedores en pruebas antiparasitarias.

Diversidad de variedades en Cucurbita ficifolia y centro de origen. *Cucurbita ficifolia* se ha descrito como la especie domesticada del género *Cucurbita* con menor diversidad morfológica y genética en México en comparación con las demás especies ([Andres 1990](#), [Cerón-González et al. 2010](#)). Nosotros reportamos un patrón distinto en Tomayquichua, donde *C. ficifolia* es la especie más diversa con 13 variedades como resultado de la combinación de seis patrones distintos de colores del fruto, tres formas distintas y tres categorías de tamaño con criterios de selección distintos. Nuestros resultados concuerdan con estudios recientes que sugieren una gran diversidad morfológica en Perú para esta especie, al describir de tres a cuatro colores distintos del fruto, seis formas distintas y cuatro categorías de tamaño para la región norte de Perú ([Delgado-Paredes et al. 2014](#)), aunque estudios recientes han encontrado una diversidad aparentemente similar en Nayarit (México), sin reportar el número exacto de variedades locales ([Ruelas-Hernández et al. 2015](#)). La diversidad morfológica observada en este estudio podría ser el producto de plasticidad fenotípica ([Sultan 2000](#)) o el producto de la diversidad genética y la historia evolutiva de *C. ficifolia*. A pesar de que carecemos de datos genéticos para las poblaciones de *C. ficifolia* distribuidas en Tomayquichua, otros trabajos han encontrado cierta correlación positiva entre la variación genética y la variación morfológica dentro de las poblaciones domesticadas de *C. moschata*, *C. pepo* y *C. argyrosperma* en México utilizando isoenzimas ([Wilson 1989](#)), al igual que correlaciones entre la variación genética en regiones codificantes y el uso de variedades de *C. maxima* en España utilizando polimorfismos amplificados relacionados a secuencias (SRAPs por sus siglas en inglés) ([Ferriol et al. 2004](#)). Estudios recientes a nivel genómico en *C. pepo* han encontrado 48 variantes a lo largo de 15 de los 20 cromosomas de la especie que pueden explicar del 1.5 al 62.9 % de la varianza fenotípica observada en distintos caracteres de importancia agronómica, lo que nuevamente sugiere una correlación entre la variación genética y morfológica en el género ([Montero-Pau et al. 2017](#)). Adicionalmente, un estudio con polimorfismos de longitud de fragmentos amplificados (AFLPs por sus siglas en inglés) en poblaciones mexicanas de *C. ficifolia* identificó agrupaciones relacionadas al color del fruto, uno de los caracteres analizados en el presente estudio, y al color de la semilla, sugiriendo que los fenotipos observados en las accesiones están ligados a sus genotipos, formando linajes evolutivos discretos ([Moya-Hernández et al. 2018](#)). Finalmente, la coexistencia observada de distintas variedades de *C. ficifolia* en un solo espacio con características ambientales similares (*i.e.* en una misma chacra o traspatio) sugiere que la variación morfológica observada en *C. ficifolia* no es moldeada

solamente por el ambiente, sino que es heredable y es explicada mayoritariamente por un componente genético y no por un componente epigenético o de plasticidad fenotípica. La domesticación de *C. ficifolia* ha sido pobremente estudiada, y su centro de origen es aún incierto, en parte porque se desconoce de la existencia de algún pariente silvestre cercano (Lira *et al.* 2016). Se ha planteado que los centros de domesticación suelen albergar la mayor cantidad de variación y formas endémicas (Harlan 1992). Dada la alta variación biológica observada en Perú (Delgado-Paredes *et al.* 2014), es posible que *C. ficifolia* haya sido domesticada en esta región andina. Las variedades locales observadas de esta especie se han mantenido gracias a preferencia por los agricultores locales en Tomayquichua (Brush 1986, Jarvis *et al.* 2008, Velásquez-Milla *et al.* 2011). Los registros arqueológicos más antiguos de *C. ficifolia* tienen una antigüedad de 5,000 años y se han encontrado en Huaca Prieta, al noroeste de Perú (Whitaker 1981) lo cual apoya un centro andino de domesticación para la especie. Sin embargo, las aparentes similitudes en diversidad morfológica tanto en Perú como en México (Ruelas-Hernández *et al.* 2015) hacen que los estudios de diversidad morfológica sean insuficientes para proponer a la región andina como centro de domesticación de *C. ficifolia*. Futuros estudios filogeográficos en el norte y sur del continente americano serán indispensables para dilucidar la historia evolutiva y el centro de origen de *C. ficifolia*, tomando como punto de partida el registro arqueológico y la diversidad morfológica observada en ambas regiones. Se concluye que las prácticas agrícolas tradicionales son particularmente activas en el mantenimiento de la diversidad biológica observada tanto en *C. ficifolia* como en *C. moschata*. Las variedades locales de *C. maxima* en la región parecen ser desplazadas progresivamente por una variedad mejorada destinada a aumentar la producción mediante prácticas agrícolas intensivas. La alta diversidad morfológica de *C. ficifolia* observada en Perú, en conjunto con la evidencia arqueológica, abre la posibilidad de un origen andino para esta última especie, pero se requieren posteriores estudios con datos genéticos para poner a prueba esta hipótesis.

Agradecimientos

Los autores expresan su agradecimiento a las comunidades rurales del distrito de Tomayquichua por brindarnos su tiempo para realizar las entrevistas. Un agradecimiento especial a Don Marino Ticlavilca y su familia por alojarnos en su hogar y recibirnos con calidez y afecto. Se agradece al Ministerio del Ambiente del Perú (MINAM) y a AQUAGROINFREST SAC por proporcionarnos la fotografía de la chacra utilizada en la Figura 2. Asimismo, se agradece a la Universidad Nacional Agraria la Molina (UNALM) y a la Universidad Nacional

Autónoma de México (UNAM) por organizar el Curso Internacional “Domesticación, manejo y conservación *in situ* de recursos genéticos”, a través del cual se incentivó la realización de este trabajo. También agradecemos al Instituto Nacional de Innovación Agraria (INIA), Perú, al Programa de Apoyo a la Investigación e Innovación Tecnológica (PAPIIT, UNAM IN206217), al CONACyT (proyecto A1-S-14306) y a la CONABIO (proyecto RG023) por el apoyo financiero para la logística en la realización del curso. Se agradece al MINAM por las facilidades otorgadas a Verónica Cañedo-Torres, así como al Programa de Apoyo a los Estudios de Posgrado (PAEP) de la UNAM por el apoyo económico para transporte y viáticos otorgados a Josué Barrera-Redondo y Helena S. Hernández-Rosales. Josué Barrera-Redondo es estudiante del Programa de Doctorado en Ciencias Biomédicas de la UNAM y recibió la beca 583146 del CONACyT. Helena S. Hernández-Rosales es estudiante del Programa de Doctorado en Ciencias Biomédicas de la UNAM y recibió la beca 385367 del CONACyT. Se agradece al proyecto CONABIO KE004 por cubrir los gastos de publicación de este manuscrito.

Literatura citada

- Andres TC. 1990. Biosystematics, theories on the origin and breeding potential of *Cucurbita ficifolia*. In: Bates DM, Robinson RW, Jeffrey C. eds. *Biology and Utilization of the Cucurbitaceae*. Ithaca, New York: Comstock, Cornell University Press. pp.102-119. ISBN-13: 978-0801416705; DOI: <https://doi.org/10.1126/science.248.4956.751>
- Bellon MR, Berthaud J. 2004. Transgenic maize and the evolution of landrace diversity in Mexico: the importance of farmers' behavior. *Plant Physiology* **134**: 883-888. <https://doi.org/10.1104/pp.103.038331>
- Biblioteca ACTAF. 2007. Instructivo técnico del cultivo de la calabaza. Cuba: Asociación Cubana de Técnicos Agrícolas y Forestales, Instituto de Investigaciones de Viandas Tropicales.
- Brush SB. 1986. Genetic diversity and conservation in traditional farming systems. *Journal of Ethnobiology* **6**: 151-167.
- Casas A, Caballero J, Valiente-Banuet A, Soriano JA, Dávila P. 1999. Morphological variation and the process of domestication of *Stenocercus stellatus* (Cactaceae) in Central Mexico. *American Journal of Botany* **86**: 522-533. <https://doi.org/10.2307/2656813>
- Casas A, Otero-Arnaiz A, Pérez-Negrón E, Valiente-Banuet A. 2007. *In situ* management and domestication of plants in Mesoamerica. *Annals of Botany* **100**: 1101-1115. DOI: <https://doi.org/10.1093/aob/mcm126>

- Casas A, Parra F. 2016. El manejo de recursos naturales y ecosistemas: la sustentabilidad en el manejo de recursos genéticos. In: Casas A, Torres-Guevara J, Parra F, eds. *Domesticación en el Continente Americano Volumen 1. Manejo de Biodiversidad y Evolución Dirigida por las Culturas del Nuevo Mundo*. Morelia, México: Editorial Morevalladolid, pp. 25-50. ISBN: 978-612-4147-59-3
- Casas A, Viveros JL, Caballero J. 1994. Etnobotánica mixteca. Sociedad, cultura y recursos naturales en la Montaña de Guerrero. México: Secretaría de Educación Pública (SEP) / Instituto Nacional Indigenista (INI)/ Consejo Nacional para la Cultura y las Artes (CONACULTA), ISBN: 968-29-6421-0
- CCTA[Coordinadora de Ciencia y Tecnología en los Andes]. 2011. Microcuenca de Warmiragra. http://www.ccta.org.pe/informacion/fichas/c_cuencas/cr_warmiragra.html (acceso el 29 de julio del 2018).
- Cerón-González L, Legaria-Solano JP, Villanueva-Verduzco C, Sahagún-Castellanos J. 2010. Diversidad genética en cuatro especies mexicanas de calabaza (*Cucurbita* spp.). *Revista Fitotecnia Mexicana* **33**: 189-196.
- Delgado-Paredes GE, Rojas-Idrogo C, Sencie-Tarazona A, Vásquez-Núñez L. 2014. Caracterización de frutos y semillas de algunas cucurbitáceas en el norte del Perú. *Revista Fitotecnia Mexicana* **37**: 7-20.
- Diamond J. 2002. Evolution, consequences and future of plant and animal domestication. *Nature* **418**: 700-707. DOI: <https://doi.org/10.1038/nature01019>
- Díaz-Obregón D, Lloja-Lozano L, Carbajal-Zúñiga V. 2004. Estudios preclínicos de *Cucurbita maxima* (semilla de zapallo) un antiparasitario intestinal tradicional en zonas urbano rurales. *Revista de Gastroenterología del Perú* **24**: 323-327.
- Eguiarte LE, Hernández-Rosales HS, Barrera-Redondo J, Castellanos-Morales G, Paredes-Torres LM, Sánchez-de la Vega G, Ruiz-Mondragón KY, Vázquez-Lobo A, Montes-Hernández S, Aguirre-Planter E, Souza V, Lira R. 2018. Domesticación, diversidad y recursos genéticos y genómicos de México: El caso de las calabazas. *TIP Revista Especializada en Ciencias Químico-Biológicas* **21**: 85-101. DOI: <https://doi.org/10.22201/fesz.23958723e.2018.0.159>
- Ferriol M, Picó B, Nuez F. 2004. Morphological and molecular diversity of a collection of *Cucurbita maxima* landraces. *Journal of the American Society for Horticultural Science* **129**: 60-69. DOI: <https://doi.org/10.21273/JASHS.129.1.0060>
- Gepts P. 2014. The contribution of genetic and genomic approaches to plant domestication studies. *Current Opinion in Plant Biology* **18**: 51-59. DOI: <https://doi.org/10.1016/j.pbi.2014.02.001>
- Grzybek M, Kukula-Koch W, Strachecka A, Jaworska A, Phiri AM, Paleolog J, Tomczuk K. 2016. Evaluation of Anthelmintic Activity and Composition of Pumpkin (*Cucurbita pepo* L.) Seed Extracts-*In Vitro* and *in Vivo* Studies. *International Journal of Molecular Sciences* **17**: 1456. DOI: <https://doi.org/10.3390/ijms17091456>
- Harlan JR. 1992. *Crops & Man*. Wisconsin: Crop Science Society of America, ISBN-13: 978-0891181071
- Hawthorne BT. 1990. Age of fruit at harvest influences incidence of fungal storage rots on fruit of *Cucurbita maxima* D. hybrid 'Delica'. *New Zealand Journal of Crop and Horticultural Science* **18**: 141-145. DOI: <https://doi.org/10.1080/01140671.1990.10428085>
- INEI [Instituto Nacional de Estadística e Informática]. 2007. Censos nacionales 2007: XI de población y VI de vivienda. <https://web.archive.org/web/20110725205514/http://desa.inei.gob.pe/censos2007/tabulados/> (acceso el 16 de julio del 2019).
- Jarvis DI, Brown AHD, Cuong PH, Collado-Panduro L, Latourniere-Moreno L, Gaywali S, Tanto T, Sawadogo M, Mar I, Sadiki M, Hue NTN, Arias-Reyes L, Balma D, Bajrachary J, Castillo F, Rijal D, Belqadi L, Rana R, Saidi S, Ouedraogo J, Zangre R, Rhrib K, Chavez JL, Schoen D, Sthapit BR, De Santis P, Fadda C, Hodgkin T. 2008. A global perspective of the richness and evenness of traditional crop variety diversity maintained by farming communities. *Proceedings of the National Academy of Sciences of the United States of America* **108**: 5326-5331. DOI: <https://doi.org/10.1073/pnas.0800607105>
- Kim MY, Kim EJ, Kim YN, Choi C, Lee BH. 2012. Comparison of the chemical compositions and nutritive values of various pumpkin (Cucurbitaceae) species and parts. *Nutrition Research and Practice* **6**: 21-27. DOI: <https://doi.org/10.4162/nrp.2012.6.1.21>
- Laime-Ajacopa T, Cazazola E, Layme-Pairumani F, Plaza-Martínez P. 2007. Diccionario Bilingüe Iskay simipi yuyayk'ancha "Quechua - Castellano, Castellano - Quechua". La Paz, Bolivia: AGRUCO-Obras generales.
- Lira R, Eguiarte L, Montes S, Zizumbo-Villarreal D, Colunga-GarcíaMarín P, Quesada M. 2016. *Homo sapiens-Cucurbita* interaction in Mesoamerica: Domestication, Dissemination and Diversification. In: Lira R, Casas A, Blancas J, eds. *Ethnobotany of Mexico*. New York: Springer-Verlag, 389-402. ISBN: 978-1-4614-6669-7

- Louette D, Charrier A, Berthaud J. 1997. *In situ* conservation of maize in Mexico: genetic diversity and maize seed management in a traditional community. *Economic Botany* **51**: 20-38. DOI: <https://doi.org/10.1007/BF02910401>
- Lozada-Aranda M, Rojas-Barrera I, Mastretta-Yanes A, Ponce-Mendoza A, Burgeff C, Orjuela-R MA, Oliveros O. 2017. Las milpas de México. *Oikos* **17**: 10-12.
- Mariaca-Méndez R. 2017. Avances en el estudio de la fauna de los traspatios familiares en el sureste de México. In: Casas A, Torres-Guevara J, Parra F, eds. *Domesticación en el Continente Americano Volumen 2. Investigación para el manejo sustentable de recursos genéticos en el nuevo mundo*. México: Editorial Morevalladolid, Morelia. pp. 475-508. ISBN: 978-607-02-9334-4
- Marie-Magdeleine C, Hoste H, Mahieu M, Varo H, Archimede H. 2009. *In vitro* effects of *Cucurbita moschata* seed extracts on *Haemonchus contortus*. *Veterinary Parasitology* **161**: 99-105. DOI: <https://doi.org/10.1016/j.vetpar.2008.12.008>
- Montero-Pau J, Blanca J, Esteras C, Martínez-Pérez EM, Gómez P, Monforte AJ, Cañizares J, Picó B. 2017. An SNP-based saturated genetic map and QTL analysis of fruit-related traits in zucchini using genotyping-by-sequencing. *BMC Genomics* **18**: 94. DOI: <https://doi.org/10.1186/s12864-016-3439-y>
- Moya-Hernández A, Bosquez-Molina E, Serrato-Díaz A, Blancas-Flores G, Alarcón-Aguilar FJ. 2018. Analysis of genetic diversity of *Cucurbita ficifolia* Bouché from different regions of Mexico, using AFLP markers and study of its hypoglycemic effect in mice. *South African Journal of Botany* **116**: 110-115. DOI: <https://doi.org/10.1016/j.sajb.2018.02.409>
- Nee M. 1990. The domestication of *Cucurbita* (Cucurbitaceae). *Economic Botany* **44**: 56-68. DOI: <https://doi.org/10.1007/BF02860475>
- Paris HS. 2016. Genetic Resources of Pumpkins and Squash, *Cucurbita* spp. In: Grumet R, Katzir N, Garcia-Mas J. eds. *Plant Genetics and Genomics of Cucurbitaceae*. Cham, Switzerland: Springer Nature, Gewerbestrasse 11, 6330 111-154. DOI: <https://doi.org/10.1007/978-3-319-49332-9>; ISBN-13: 978-3319493305
- Pickersgill B. 2018. Parallel vs. convergent evolution in domestication and diversification of crops in the Americas. *Frontiers in Ecology and Evolution* **6**: 56. DOI: <https://doi.org/10.3389/fevo.2018.00056>
- Pulgar-Vidal J. 1979. Geografía del Perú: Las Ocho Regiones Naturales del Perú. Lima, Perú: Universo SA, 145-161. ISBN 9972-58-007-5
- Ruelas-Hernández PG, Aguilar-Castillo JA, García-Paredes JD, Valvidia-Bernal R, López-Guzmán GG. 2015. Diversidad morfológica de especies cultivadas de calabaza (*Cucurbita* spp.) en el estado de Nayarit. *Revista Mexicana de Ciencias Agrícolas* **6**: 1845-1856.
- Sánchez-de la Vega G, Castellanos-Morales G, Gámez N, Hernández-Rosales HS, Vásquez-Lobo A, Aguirre-Planter E, Jaramillo-Correa JP, Montes-Hernández S, Lira-Saade R, Eguiarte LE. 2018. Genetic resources in the “calabaza pipiana” squash (*Cucurbita argyrosperma*) in Mexico: Genetic diversity, genetic differentiation and distribution models. *Frontiers in Plant Science* **9**: 1-18. DOI: <https://doi.org/10.3389/fpls.2018.00400>
- Sanjurj OI, Piperno DR, Andres TC, Wessel-Beaver L. 2002. Phylogenetic relationships among domesticated and wild species of *Cucurbita* (Cucurbitaceae) inferred from a mitochondrial gene: Implications for crop plant evolution and areas of origin. *Proceedings of the National Academy of Sciences of the United States of America* **99**: 535-540. DOI: <https://doi.org/10.1073/pnas.012577299>
- Sultan SE. 2000. Phenotypic plasticity for plant development, function and life history. *Trends in Plant Science* **5**: 537-542. DOI: [https://doi.org/10.1016/S1360-1385\(00\)01797-0](https://doi.org/10.1016/S1360-1385(00)01797-0)
- Torres B, Maza OJ, Aguirre P, Hinojosa L, Günter S. 2015. The contribution of traditional agroforestry to climate change adaptation in the ecuadorian amazon: the chakra system. In: Leal-Filho W, eds. *Handbook of Climate Change Adaptation*. Berlin, Heidelberg: Springer, ISBN: 978-3-642-38669-5
- Vavilov NI. 1926. *Origin and geography of cultivated plants*. Cambridge, UK: Cambridge University Press. ISBN: 0521404274, 9780521404273.
- Velásquez-Milla D, Casas A, Torres-Guevara J, Cruz A. 2011. Ecological and socio-cultural factors influencing *in situ* conservation of crop diversity by traditional Andean households in Peru. *Journal of Ethnobiology and Ethnomedicine* **7**: 40. DOI: <https://doi.org/10.1186/1746-4269-7-40>
- Veläläinen M, Negri V, Maxted N. 2009. European landraces: on-farm conservation, management and use. Rome, Italy: Technical Bulletin No. 15. Biodiversity International, ISBN-13: 978-92-9043-805-2
- Whitaker TW. 1981. Archeological Cucurbits. *Economic Botany* **35**: 460-466. DOI: <https://doi.org/10.1007/BF02858596>

Whitaker TW, Bemis WP. 1964. Evolution in the genus *Cucurbita*. *Evolution* **18**: 553-559. DOI: <https://doi.org/10.2307/2406209>

Wilson HD. 1989. Discordant patterns of allozyme and morphological variation in mexican *Cucurbita*. *Systematic Botany* **14**: 612-623. DOI: <https://doi.org/10.2307/2419006>

Zeder MA. 2015. Core questions in domestication research.

Proceedings of the National Academy of Sciences of the United States of America **112**: 3191-3198. DOI: <https://doi.org/10.1073/pnas.1501711112>

Zheng YH, Alverson AJ, Wang QF, Palmer JD. 2013. Chloroplast phylogeny of *Cucurbita*: Evolution of the domesticated and wild species. *Journal of Systematics and Evolution* **51**: 326-334. DOI: <https://doi.org/10.1111/jse.12006>

Editor de sección: Andrea Martínez Ballesté

Contribuciones de los autores: JBR concibió el estudio, diseñó el estudio, realizó trabajo de campo y redactó el manuscrito. HSHR concibió el estudio, diseñó el estudio, realizó trabajo de campo y redactó el manuscrito. VCT realizó trabajo de campo. KAA realizó trabajo de campo. JTG realizó trabajo de campo y aportó información sobre los sistemas agrícolas regionales y el contexto del cultivo de calabazas y zapallos. FP realizó trabajo de campo. ITG realizó trabajo de campo. AC concibió el estudio, diseñó el estudio, realizó trabajo de campo y redactó el manuscrito. Todos los autores revisaron la versión final del manuscrito.

Discusión general.

Esta tesis es el primer estudio a gran escala que analiza con detalle de la diversidad genética y la historia evolutiva de *C. moschata*. Además, este trabajo es el primero que evalúa los patrones de diversidad, flujo genético y estructura genética de las tres principales especies cultivadas de calabazas en México (*C. pepo* spp. *pepo*, *C. moschata* y *C. argyrosperma* spp. *argyrosperma*). Finalmente, se llevo a cabo el primer estudio formal que documenta la diversidad de los cultivos locales de *C. ficifolia* en una región andina del Perú.

México es centro de diversificación y alta diversidad genética de *C. moschata*.

Mesoamérica se ha propuesto como el centro de origen y domesticación de *C. moschata* (Whitaker y Bird, 1949; Dillehay, 2007; OECD, 2012) y entre los resultados más destacados en esta tesis que apoyan esta hipótesis puedo mencionar los siguientes:

1. Alta diversidad en *C. moschata* en México. Se identificó a Veracruz como el estado con la mayor diversidad mitocondrial reportada hasta el momento ($H_d = 0.889$) y las poblaciones analizadas son relativamente cercanas geográficamente a las Cuevas de Ocampo, en Tehuacán, Puebla, donde se han encontrado los registros arqueobotánicos más antiguos de esta especie (Whitaker et al., 1957).
2. Se encontró que los haplotipos de *C. moschata* diversificaron en México hace ~13,000 años. Este fechamiento coincide con los tiempos de domesticación propuestos para distintas especies cultivadas (Meyer y Purugganan, 2013).

Sin embargo, en este trabajo se reconoce que para afirmar el origen Mesoamericano de *C. moschata* es aún necesario:

1. Incluir en los análisis variedades locales que abarquen toda la distribución de la especie, partiendo desde Centroamérica hasta Sudamérica.
2. Realizar esfuerzos de colecta, documentación y análisis formal para determinar la existencia de su pariente o ancestro silvestre, hasta el momento desconocido.
3. Implementar análisis morfológicos detallados y correlacionar los datos genómicos, así como con los datos morfológicos.
4. Utilizar métodos de inferencia Bayesiana para poner a prueba distintas hipótesis referentes al origen, la domesticación e la historia demográfica de esta especie.

El alcance de los resultados presentados en este trabajo si permiten señalar que México es un centro de diversificación de esta especie. Primero, se identificaron variedades asociadas a un gradiente altitudinal y variedades relacionadas con la geografía (Hernández-Rosales, et al., 2020). Con estos resultados se puede entender porque *C. moschata* es fácilmente cultivada casi en cualquier lugar en todo el mundo. Además, dada la alta diversidad genética a nivel mitocondrial (11 haplotipos, $H_d = 0.589$ y $\pi = 0.0023$) e intermedia a nivel nuclear (H_E 11 loci de microsatélites nucleares = 0.395; H_E 1,797 SNPs = 0.1324) se puede indicar que una parte importante del acervo de diversidad de *C. moschata* se ve representado y salvaguardo en las variedades locales de México.

Calabazas arriba y abajo: la divergencia de dos linajes de *C. moschata* explica su amplia distribución geográfico-altitudinal.

El estudio de esta especie a partir de marcadores mitocondriales nos permitió identificar con mayor detalle que existen dos linajes principales que pudieron haber divergido hace unos 13,500 años en México (ver Anexo 1). Asimismo, se identificó que estos linajes están asociados a un gradiente altitudinal. Con estos resultados, se propuso que en México se pueden diferenciar las variedades de calabaza de *tierras altas* (> 1,200 m) y las variedades de calabaza de *tierras bajas* (< 1,200 m; Hernández-Rosales, et al., 2020). *Cucurbita moschata*, igualmente a otras especies de plantas domesticadas como la cebada (*Hordeum vulgare*) en Etiopia (Tanto Hadado et al., 2010) o las variedades locales de maíz (*Zea mays*) en México (Van Heerwaarden et al., 2011), muestra patrones divergencia de linajes asociados a la elevación. Con los SNPs no se detectó este patrón, sin embargo, no se descarta que los loci con señales fuertes de selección podrían estar asociados a condiciones ambientales relacionadas a un gradiente altitudinal.

Con los resultados de este trabajo, se propone que esta especie pudiera estar presentando procesos de adaptación local a distintas condiciones ambientales asociadas a la altitud. Por lo anterior, se proporciona evidencia que permite entender la amplia distribución geográfico-altitudinal de esta especie y su facilidad de cultivo en distintas condiciones tropicales, así como la tolerancia a sequías y temperaturas extremas en todas partes del mundo (Lira et al., 1995).

Los estudios de identificación de genes candidatos con señales de selección permiten entender la adaptación local, la diferenciación ecológica de las especies y un mejor entendimiento de las bases genéticas a las respuestas al estrés ambiental causado por el cambio climático (Xu

et al., 2014; Aguirre-Liguori et al., 2019). Las variedades de *tierras altas* y *tierras bajas* de *C. moschata* deben ser estudiadas bajo este enfoque para comprender con mayor profundidad la historia evolutiva y procesos de adaptación de esta especie. Hoy en día la identificación de genes candidatos en la biología aplicada es fundamental para el mejoramiento de cultivos de manera más rápida y eficiente (Xu et al., 2014).

La geografía y las preferencias culturales modelan la estructura genética de *C. moschata* en México.

El trabajo de esta tesis también permitió identificar con loci de microsatélites nucleares que en México hay dos principales grupos genéticos de *C. moschata* que se diferencian posiblemente debido a que el Istmo de Tehuantepec funge como una barrera al flujo genético (Hernández-Rosales et al., 2020). Cabe señalar que esta estructura genética fue confirmada con el análisis de 1,797 SNPs de las mismas 24 localidades previamente estudiadas (Capítulo 1 y 3 respectivamente). Las variedades de calabazas de la Península de Yucatán, como su nombre lo indica, son calabazas que se distribuyen en toda la región de la Península de Yucatán, junto con el estado de Chiapas. Estas variedades conforman un grupo genético bien definido (Hernández-Rosales et al., 2020). A partir de los registros fotográficos identificamos, que se caracterizan por ser de tamaños pequeños y aunque diversos predominan aquellos de costillas redondeadas y aplanados y despliegan una amplia variación en los patrones coloración. Las calabazas de esta región destacan por ser preferentemente consumidas en su forma inmadura como verduras.

Por otro lado, las variedades de calabazas que se denominaron Continentales hacen referencia a todas aquellas calabazas que crecen fuera de la Península de Yucatán en cualquier otro lugar del territorio mexicano. Aunque predominan dos grupos genéticos, estas calabazas no presentan una estructura genética bien definida (Hernández-Rosales et al., 2020). Contrario a las variedades de la Península, estas calabazas se caracterizan por desplegar diversos tamaños en general mucho más grandes que las de la Península, formas redondas, o con cuello u ovaladas y patrones de coloración muy diversos. Ya que se consume principalmente la pulpa y las semillas, por lo común estas variedades de calabazas son cosechadas hasta que los frutos estén totalmente maduros.

Como ya se indicó, la principal barrera geográfica que parece limitar el flujo genético entre estas dos variedades de calabaza es el Istmo de Tehuantepec. Esta región ha experimentado

diversos levantamientos y oscilaciones del nivel del mar (Lambeck y Chappell, 2001), lo cual probablemente está asociado con que el flujo genético de especies que se distribuyen en tierras bajas se vea disminuido y las poblaciones se diferencien entre cada lado del Istmo. Este mismo patrón de diferenciación genética asociados a esta barrera geográfica se ha reportado para otra especie domesticada del género *Cucurbita* (Sánchez-de la Vega, et al., 2018) y otras especies silvestres de plantas y animales (revisar Ornelas et al., 2013). Sin embargo, resulta indispensable integrar muestras de *C. moschata* provenientes del estado de Tabasco para aportar mayor evidencia de la barrera al flujo genético que se propone en este trabajo.

Debido que esta tesis no se analizó de manera formal los datos morfométricos y tampoco se documentaron los criterios de selección de los agricultores locales, la hipótesis planteada acerca de las formas y las preferencias culturales que diferencian las variedades de la Península de Yucatán de las variedades Continentales deberá ser corroborada con estudios morfológicos y etnobotánicos formales. No obstante, no se descarta llevar a cabo análisis tipo Bayesenv o Bayesenv considerando formas generales de los frutos y/o generando grupos de acuerdo a su asignación a un grupo genético identificada con el programa Admixture.

Las especies domesticadas del género *Cucurbita*: Juntas, pero no revueltas.

Se analizaron variedades locales de tres de las especies domesticadas más utilizadas y ampliamente cultivadas en México (i.e., *C. argyrosperma* spp. *argyrosperma* L. (calabaza pipiana), *C. moschata* Duchesne (calabaza de Castilla), *C. pepo* spp. *pepo* L (calabacita)) con cinco loci de microsatélites. Se observó que están genéticamente diferenciadas entre sí ($G_{ST} = 0.457-0.508$) y que el flujo genético es limitado entre ellas. Estos resultados son congruentes con lo reportado por Paris (2016) sobre que estas taxa se encuentran claramente aislados reproductivamente.

Se sabe bien de la capacidad de generar híbridos entre las especies del género *Cucurbita* tanto a partir de su cruce espontáneo como mediante experimentos de cruces artificiales (Whitaker and Bemis, 1962 y 1964; Decker-Walters et al., 1990; Cruz-Reyes et al., 2015). Considero que evaluar la intensidad y dirección del flujo genético entre especies es fundamental para fines de mejoramiento de cultivos y para la conservación de los recursos fitogenéticos.

El hecho de que las barreras reproductivas no sean tan estrictas entre las especies del género *Cucurbita* brinda la posibilidad de transferir rasgos que permitan el mejoramiento y

rendimiento de los cultivos. Dentro de las calabazas domesticadas un buen ejemplo es la transferencia de resistencia a hongos en los cultivos (Paris and Cohen, 2002). Cabe destacar, que estos experimentos se desarrollan mediante cruza artificiales y no se dan de manera espontánea dentro de las parcelas y cultivos.

Por otro lado, que las barreras naturales a la reproducción entre especies eviten en gran medida la generación de híbridos potencialmente fértiles pueden ayudar en la conservación de los reservorios genéticos representados en las variedades locales. También reduce la posible contaminación de los transgenes de los cultivos mejorados hacia las especies silvestres y hacia las mismas variedades locales.

Latinoamérica y los recursos genéticos de las calabazas: El caso del Perú.

A lo largo de esta tesis se ha insistido en la importancia de documentar las variedades locales de las especies domesticadas y conocer su diversidad genética. Los criterios para definir lo que se considera que tiene mayor valor es distinto para cada domesticador, así como sus necesidades inmediatas y sus preferencias culturales (Casas et al., 1999). También se debe recordar que las condiciones ambientales son distintas a lo largo de la distribución natural del género *Cucurbita* (Castellanos-Morales et al., 2018). Por lo anterior, no debe ser extraño que la diversificación de las especies domesticadas a lo largo del continente americano no sea igual.

La región andina particularmente la que corresponde al Perú es reconocida como centro de origen de *C. maxima* y se ha sugerido que también de *C. ficifolia* (Vavilov 1926; Harlan 1992, Diamond 2002). Junto con Mesoamérica, esta región comparte los centros de origen y diversificación de algunas especies domesticadas del género *Cucurbita* (Sanjur et al., 2002; Zheng et al. 2013). En esta tesis se presentan datos que enriquecen el conocimiento de la diversidad biológica de las calabazas en la región norte del Perú.

A partir de las entrevistas semiestructuradas llevadas a cabo en Perú, se encontró que las especies mayormente consumidas y cultivadas bajo sistemas agrícolas tradicionales en esta región son *C. ficifolia* y *C. moschata*, mientras que *C. máxima* es la única especie que se cultiva en sistemas de siembra semi-intensivos (Barrera-Redondo et al., 2020). Al igual que en México, los sistemas tradicionales involucran policultivos que promueven el flujo genético entre variedades.

Los registros arqueológicos más antiguos de *C. ficifolia* tienen una antigüedad de 5,000 años y se han encontrado en Huaca Prieta, al noroeste de Perú (Whitaker 1981). En conjunto con la caracterización de 13 variedades en *C. ficifolia*, los resultados presentados permiten sugerir un posible origen Andino de esta especie (Delgado-Paredes et al. 2014). Sin embargo, debido a la región tan pequeña que se analizó, no se descarta que la variación observada sea producto de plasticidad fenotípica (Sultan, 2000). Para corroborar y entender mejor la historia evolutiva de esta especie, se propone llevar a cabo estudios genéticos y filogeográficos que analicen variedades del norte y sur del continente, así como un estudio más detallado de la diversidad biológica del resto de las especies de *Cucurbita* en todo Perú.

Lo importante no es que marcador genético usamos, sino la pregunta evolutiva.

Los patrones generales de estructura genética de las 24 localidades analizadas de *C. moschata* obtenidos con 11 loci de microsatélites nucleares (Hernández-Rosales et al., 2020) y con 1,797 SNPs son congruentes entre sí. Con estos resultados, se puede asegurar que cuando los marcadores son selectivamente neutros, sin importar si se analizan unos cuantos o miles de ellos, se pueden hacer estudios evolutivos confiables.

Los microsatélites continúan siendo una excelente opción para desarrollar estudios de genética de poblaciones e historia demográfica. Sin embargo, deben tomarse en cuenta las siguientes recomendaciones: (1) elegir loci variables y que no presenten altas tasas de alelos nulos; (2) de preferencia que hayan sido diseñados y que hayan sido probados para la especie que se desea analizar; (3) se conozca la región dentro del genoma en la que se localizan y que se asegure que no se encuentre ligados y, (4) asegurar su naturaleza neutra, es decir, que estas regiones no se encuentren bajo presiones de selección. En términos de poder de cómputo, es mucho más sencillo analizarlos porque, además de las paqueterías de *R*, existen varios programas con interfaz gráfica que facilitan su análisis.

En cambio, cuando se pretende entender el papel de la selección y su interacción con las otras fuerzas evolutivas, la implementación de SNPs es más adecuada. Estos marcadores se encuentran distribuidos de manera relativamente homogénea a lo largo del genoma (Escalante et al., 2014), así que se pueden identificar fácilmente cuando la *selección natural purificadora o direccional* está actuando porque comúnmente se distinguen bajos niveles de variación genética. Por otro lado, cuando zonas del genoma presentan altos niveles de variación se puede entender que

la *selección balanceadora* está favoreciendo a los individuos heterocigos para un locus (Eguiarte et al., 2013).

Esta tesis presenta un primer análisis para conocer si dentro de nuestros datos obtenidos mediante tecnologías NGS existen señales que indiquen la presencia de loci con señales de selección. Los resultados presentados no son concluyentes y se pretende como una meta futura próxima el realizar pruebas de correlación con datos ambientales como la que implementan programas como Bayescenv (de Villemereuil y Gaggiotti, 2015). Los resultados que se obtengan podrán eventualmente ser utilizados con fines de mejoramiento genético y conservación de los recursos que representa esta especie.

Finalmente, es de suma importancia recordar que se debe tener un sólido marco teórico de la genética de poblaciones para realizar estudios con marcadores genéticos dentro de especies domesticadas o no. Los resultados deben ser, independientemente el marcador que se analice, interpretados de manera correcta asumiendo los supuestos que subyacen cada fuerza evolutiva.

Los recursos genéticos de *C. moschata*.

En México las variedades locales de *C. moschata* son extensamente consumidas. Pese a que su uso como hortaliza es poco generalizado, pero la pulpa de sus frutos maduros, que se caracteriza por ser suave y de sabor dulce, la que es mayormente apreciada para la elaboración de múltiples platillos dulces y salados. En esta región del mundo la cosecha de los frutos maduros coincide con las festividades del Día de Muertos. Uno de los platillos dulces típicos que se colocan en las ofrendas es la “calabaza en tacha”, por lo que no es de extrañar que este postre se prepare con dicha especie. El aprecio cultural de las calabazas en México es invaluable.

Las variedades que se cultivan en nuestro país representan un importante recurso genético debido a la amplia variación (morfológica y genética) y adaptaciones asociadas a un gradiente altitudinal que despliegan. La documentación de las variedades reportadas en este trabajo, así como la identificación de áreas de alta diversidad genética permitirán implementar planes de manejo y conservación de estos recursos fitogenéticos adecuados.

La investigación aquí desarrollada proporciona un mejor entendimiento de la historia evolutiva de *Cucurbita moschata* en México y abre la posibilidad de poner a prueba las hipótesis propuestas en cada uno de los capítulos presentados.

Conclusiones generales.

1. El estudio de la diversidad genética, los patrones generales de la genética de poblaciones y la filogeografía permiten obtener información para contribuir en el desarrollo de planes de conservación y manejo sostenible de los recursos genéticos de las calabazas. El alcance de los resultados de esta tesis permitirá crear propuestas de conservación partiendo del conocimiento de los procesos evolutivos que han modelado la historia de estas especies en México.
2. En México se encontraron dos linajes mitocondriales de *C. moschata*, asociados a un gradiente de elevación, los cuales divergieron hace ~13,000 años. Este evento de divergencia podría estar relacionado a la adaptación a diferentes condiciones ambientales de esta especie lo que permite entender la amplia distribución de sus variedades locales.
3. En Veracruz existen localidades cercanas a la región en donde se han encontrado los restos arqueobotánicos más antiguos de *C. moschata*. Aunado a esto, es la localidad con la mayor diversidad genética a nivel mitocondrial para esta especie. Estos resultados aportan evidencia a la propuesta de que Mesoamérica es el o uno de los centros de domesticación de la calabaza de Castilla.
4. México es uno de los principales centros de diversificación de *C. moschata* en vista de su variación genética a nivel mitocondrial y nuclear. La diversidad de los recursos genéticos de *C. moschata* en esta región se ve representada en dos principales categorías: *i*) las variedades de tierras altas y las variedades de tierras bajas y *ii*) variedades de la Península de Yucatán y variedades de la región continental de México. La riqueza de formas únicas que despliega esta especie en México (principalmente en la Península de Yucatán) y hace prioritario el crear bancos de germoplasma para conservar y resguardar la diversidad de esta especie.
5. La calabacita (*C. pepo* spp. *pepo*), la calabaza de castilla (*C. moschata*) y la calabaza pipiana (*C. argyrosperma* spp. *argyrosperma*) son las tres especies domesticadas del género *Cucurbita* mayormente apreciadas en México por su versatilidad como alimento y recurso genético. En los agroecosistemas tradicionales de México las variedades locales de *C. pepo* spp. *pepo*, *C. moschata* y *C. argyrosperma* spp. *argyrosperma* se cultivan juntas, aunque la hibridación

espontánea es posible existen barreras naturales que impiden altos niveles de introgresión entre especies. El poco flujo genético detectado entre estas especies es una ventaja que permite que se mantengan como grupos genéticos bien definidos y ayuda a que se conserven los acervos genéticos de cada una de ellas.

6. El zapallo (*C. maxima*), la calabaza (*C. ficifolia*) y el loche (*C. moschata*) son las tres especies domesticadas de *Cucurbita* mayormente cultivadas y consumidas en la región andina del distrito de Tomayquichua en Perú. *Cucurbita ficifolia* es la especie más diversa en su morfología en esta región, sin descartar que esta diversidad pudiera ser un reflejo de plasticidad fenotípica, este resultado apoya la hipótesis del origen andino de esta especie.

7. La metodología implementada para la identificación de polimorfismo de una sola base de *C. moschata* resultó adecuada. Estos marcadores permitirán realizar análisis más sofisticados para poner a prueba las hipótesis sobre el origen, domesticación y diversificación de esta especie. Además, dado que esta calabaza se cultiva en ambientes contrastantes, con los SNPs será posible identificar genes candidatos a estar bajo selección.

8. Con el análisis de SNPs neutrales a lo largo de todo el genoma se apoyaron los patrones generales de diversidad genética previamente descritos en esta tesis en relación con las variedades locales de *C. moschata* se mantienen. La calabaza de castilla en México es en términos genéticos la segunda especie más diversa después de *C. pepo* spp. *pepo*. Aparentemente a nivel global, México no es una región de alta diversidad genética para esta especie, sin embargo, este patrón debe tomarse con reserva hasta realizar análisis más formales.

9. Las calabazas en México y América Latina son un importante recurso genético, su conservación y uso sostenible debe priorizarse ya que representan un valor alimenticio y cultural de gran relevancia en el país.

PERSPECTIVAS.

i. Para tener un mejor entendimiento de la historia evolutiva de *C. moschata* se recomienda en el futuro incluir variedades locales del resto de la distribución natural de toda la especie, partiendo desde Centroamérica hasta el sur del continente en los análisis de filogeografía y de fechamiento de tiempos de divergencia. También se insiste en llevar a cabo un mayor esfuerzo en la búsqueda de un posible ancestro silvestre que permitiría definir con mayor claridad el centro de origen y domesticación de esta especie. También se sugiere poner a prueba distintos escenarios de la domesticación de esta especie considerando a sus especies cercanas del Grupo *Argyrosperma*. Tampoco debe descartarse un posible origen híbrido.

ii. Para complementar el estudio de la diversidad de las calabazas cultivadas en México y en América Latina, se sugiere realizar estudios formales de carácter etnobotánico con énfasis en las preferencias culturales y criterios de selección por parte de los agricultores locales. Esta información permitirá reconocer con mayor precisión la distribución de estos recursos genéticos en nuestro país y en el resto de América Latina para generar programas de conservación y uso sostenible más formales.

iii. Considerando las eventuales consecuencias del cambio climático, los programas de mejoramiento de cultivos tomarán relevancia inmediata. Sin perder la prioridad de salvaguardar los reservorios genéticos de cada especie, es necesario reconocer los genes con potencial para el mejoramiento de cultivos a partir de la identificación de genes candidatos con señales de selección. Se recomienda que se realicen ensayos de hibridación intra e interspecíficos a través de experimentos controlados en jardines comunes empleando marcadores genéticos bien identificados.

iv. En el caso de las calabazas del Perú y el resto del continente, se recomienda realizar esfuerzos por documentar la diversidad genética de las variedades locales que crecen en cada región. Particularmente la información de *C. ficifolia* es aún limitada, se recomienda que los esfuerzos se concentren en esta especie dado su posible origen Andino y la gran diversidad morfológica que se ha documentado recientemente en esta región.

Anexo I: Fechamiento de tiempos de divergencia entre los haplotipos de ADN mitocondrial de *Cucurbita moschata*.

Teoría neutra, reloj molecular y teoría de coalescencia.

A partir de los primeros estudios realizados para describir la diversidad genética utilizando marcadores moleculares (p. ej., aloenzimas) se encontró que existía una gran cantidad de variación genética dentro de las poblaciones. Las controversias sobre el origen y el mantenimiento de dicha variación no se hicieron esperar, y entre 1960–1970 se desarrollaron dos hipótesis principales para explicar este hecho. Por un lado, el **modelo balanceado** atribuye que solo la *selección balanceadora* es la encargada de mantener la gran mayoría de las variantes dentro de las poblaciones. Mientras que **la teoría neutra** atribuye que la mayoría de esas variantes moleculares son neutras con respecto a la selección y por ello la variación no se asocia a una adecuación diferencial (Hedrick, 2005).

Motoo Kimura fue quien desarrolló la teoría neutra de la evolución entre 1950–1955, pero no publicó este trabajo hasta 1968. Su teoría tiene dos aspectos principales: *i*) la selección natural juega un papel menor en la determinación del mantenimiento de las variantes moleculares –en otras palabras, las mutaciones son selectivamente neutras porque tienen la misma adecuación relativa unas de las otras; y *ii*) la mayoría de las variantes moleculares son el resultado de la interacción entre la mutación– que genera nuevas variantes– y la deriva genética –que permite eventualmente su fijación o pérdida.

El campo de estudio de la evolución molecular implica el análisis de secuencias de ADN, ARN y proteínas con el objetivo de dilucidar los procesos y eventos históricos que originan y mantienen las diferencias entre estas secuencias a lo largo del tiempo (Hamilton, 2009). La teoría neutra de Kimura dicta las bases del modelo nulo de la evolución a nivel molecular, la cual postula que los genes y las proteínas evolucionan a velocidades o tasas relativamente constantes debido a la deriva genética (Dayan et al., 1989; Futuyma, 2005).

El **reloj molecular** es un fenómeno que permite datar la divergencia entre especies o linajes aprovechando que la sustitución de variantes moleculares se da a tasas constantes. La teoría neutra predice que las tasas de sustitución son igual a las tasas de mutación de un *locus* dado (Hedrick, 2005). Entonces, la divergencia entre dos copias de un gen será igual al número de mutaciones que se han producido a lo largo de las dos ramas de esas copias en la filogenia. Así, para los genes que están acumulando diferencias por deriva, el número de diferencias entre

dos especies es proporcional, en términos generales, al tiempo transcurrido desde su ancestro más reciente, por lo tanto, debería el reloj molecular es la relación lineal entre el tiempo y la divergencia, aunque debemos mencionar que hay heterogeneidades en las tasas de sustitución (Futuyma y Kirkpatrick, 2017). De cualquier manera, el reloj molecular es una herramienta poderosa que permite estimar el tiempo en que las poblaciones y las especies divergieron, así como la diversificación de familias génicas y el origen de la variación (Kumar, 2003).

La **teoría de coalescencia** considera que usualmente es más fácil modelar una genealogía hacia atrás en el tiempo (Nordborg, 2001). Imaginemos una población de tamaño constante en donde se forman constantemente por mutación nuevos alelos de los cuales unos se perderán y otros se fijarán solo por deriva genética. Bajo neutralidad, todas esas nuevas variantes coalescen en un único alelo ancestral, porque todas esas mutaciones ocurren en un único punto en el tiempo. Esta teoría permite entonces visualizar las conexiones filogenéticas entre genotipos o haplotipos de individuos dentro y entre especies, mediante el análisis de rutas verticales de descendencia hasta que derivan en un ancestro común (Vázquez-Domínguez et al., 2009).

Origen de las especies domesticadas.

Muchos estudios evolutivos de domesticación tienen como parte de sus objetivos dilucidar **cuándo** se originaron las especies domesticadas. Ayuda mucho cuando se tiene identificado el pariente silvestre y la ubicación de las áreas geográficas que aportaron los primeros organismos domesticados. Sin embargo, no es el caso para todas las especies y muchas veces solo se cuenta con los patrones actuales de distribución de la variación genética.

Las primeras técnicas para determinar la antigüedad de restos arqueobotánicos de plantas o con síndromes tempranos de domesticación implicaban asignar las fechas equivalentes a los objetos asociados y localizados dentro de un estrato. Sin embargo, estos fechamientos fueron duramente criticados debido a sus amplios márgenes de error. Posteriormente, el desarrollo de técnicas como la espectrometría de masas ha sido fundamental para fechar de manera más precisa los procesos de domesticación (Parra y Casas, 2007).

La reconstrucción de filogenias moleculares permite identificar las relaciones de parentesco entre las especies domesticadas y sus parientes silvestres, así como entre los linajes dentro de cada especie o taxa domesticado. Estas hipótesis evolutivas, bajo el marco teórico de la

filogeografía, han permitido completar la interpretación de los estudios arqueológicos enfocados en rastrear el origen de la agricultura para estudiar los procesos de domesticación (Zeder, 2006).

Es así que hoy sabemos que las especies cultivadas han acompañado a la humanidad desde hace ya más de 12,000 años (Meyer y Purugganan, 2013). La asociación más antigua entre humanos y plantas cultivadas aparentemente fue en el Medio Oriente y el Creciente Fértil (lo que hoy corresponde a la región de Líbano, Israel, Palestina, Siria, Irak, el sudeste de Turquía y el noroeste de Jordania). Posteriormente, hace unos 10,000 años se domesticaron plantas en diferentes partes del mundo, como China, Mesoamérica, los Andes, Oceanía. Y más recientemente se domesticaron otras especies en África y el norte de América (6,000–8,000 años) (Meyer et al., 2012)

Objetivo general.

Determinar los tiempos de divergencia de los haplotipos mitocondriales de *C. moschata* y analizarlos bajo un marco conceptual de evolución bajo domesticación.

Métodos.

Se determinaron los tiempos de divergencia entre los haplotipos de *C. moschata* con las secuencias amplificadas con los cebadores *trnL* (UAA)5'-*trnF* (GAA) (Taberlet et al., 1991) e identificadas como haplotipos mitocondriales (Genbank: MN381966–MN381976) (Hernández-Rosales et al., 2020). También se calculó el tiempo de divergencia de esta especie respecto a sus taxa hermanos, para ello se tomaron cuatro haplotipos mitocondriales de *C. argyrosperma* spp. *sororia* y uno de *C. argyrosperma* spp. *argyrosperma* (Genbank: MT040496–MT040508; Tapia, 2020). Como grupo externo se utilizaron tres secuencias de tres subespecies de *C. pepo* (Genbank: MN411624; N411621 y N411622; Ruiz Mondragón, 2017). Todas las secuencias se alinearon manualmente (1,029 pb).

Primero se calculó el modelo de sustitución que mejor se adaptaba al grupo de secuencias utilizando el programa jModelTest v 2.1.6 (Darriba et al., 2012). Para datar adecuadamente los tiempos de divergencia se utilizó el programa BEAST v.1.8.0 (Drummond et al., 2012). El árbol filogenético se calibró usando una distribución normal y se utilizó un punto calculado mediante reloj molecular basado en el tiempo de divergencia entre *C. argyrosperma* y *C. moschata* (1.06 Ma [0.29,1.9]; Castellanos-Morales et al., 2018). Se realizaron varias corridas de prueba previas

para ajustar y afinar los parámetros para el análisis final. Finalmente, se realizaron dos corridas independientes de 30 millones de generaciones cada una y se descartó el 10% de *burn-in*. Se implementó un modelo de reloj molecular relajado con distribución LogNormal y un modelo de tamaño poblacional constante.

Con el programa Tracer v 1.7 (Rambaut et al., 2018), se examinó que la convergencia de los valores fuera estable para cada análisis (probabilidad posterior para la corrida 1 = -1811.873 y probabilidad posterior para la corrida 2 = -1118.838). Se consideró que cada corrida alcanzaba la estacionariedad si el tamaño efectivo de la muestra, *effective sample size* (ESS), para cada parámetro (incluyendo el reloj molecular) se encontraba en valores superiores a 505. Con el programa LogCombiner v 1.8.0 se unieron los resultados de ambas corridas y eliminando el 10% de los árboles como *burn-in* (1,000 árboles iniciales de cada corrida). Finalmente se obtuvo un árbol de máxima credibilidad con las fechas de divergencia, *maximum clade credibility* (MCC), con el programa TreeAnnotator v 1.7 El árbol final se visualizó con el programa FigTree v 1.4. 3 (<http://tree.bio.ed.ac.uk/software/figtree/>).

Resultados

Las relaciones filogenéticas entre *C. moschata*, los taxa de *C. argyrosperma* y los taxa de *C. pepo* se mantienen congruentes con los reportados en la literatura (Kates et al., 2017; Castellanos-Morales et al., 2018; Figura 1).

El análisis de reloj molecular según BEAST reveló que *C. moschata* divergió de *C. argyrosperma* hace 68,000 años (0.29–1.9), como se muestra en la Figura 1. Se identificó que *C. moschata* tiene dos principales linajes que habrían divergido en México hace ~13,500 años (0–0.26; HPD de 95%). Los linajes corresponden a dos grandes grupos de haplotipos que se diferencian de acuerdo con los dos haplotipos (A y B) más comunes encontrados en 24 poblaciones de México (ver Hernández-Rosales et al., 2020). Los haplotipos derivados dentro de cada uno de estos linajes divergieron más recientemente entre 8,000–5,100 años (Figura 1).

Node	Post. Prob.	Node age (Ma.)	95 HDP
I	0.99	0.0135	0.0 – 0.260
II	0.42	0.0088	0.0 – 0.168
III	0.88	0.0051	0.0 – 0.107
IV	0.99	0.0050	0.0 – 0.104
V	0.41	0.0027	0.0 – 0.063
VI	0.97	0.0024	0.0 – 0.054
VII	0.88	0.0013	0.0 – 0.046
VIII	0.82	0.0010	0.0 – 0.034
IX	0.99	0.0009	0.0 – 0.024
X	0.49	0.0008	0.0 – 0.028
XI	0.31	0.0006	0.0 – 0.021
XII	0.54	0.0019	0.0 – 0.044
XIII	0.99	0.0044	0.0 – 0.095
XIV	0.99	0.0158	0.0 – 0.308
XV	0.95	0.0015	0.0 – 0.058
XVI	0.99	0.0170	0.0 – 0.347

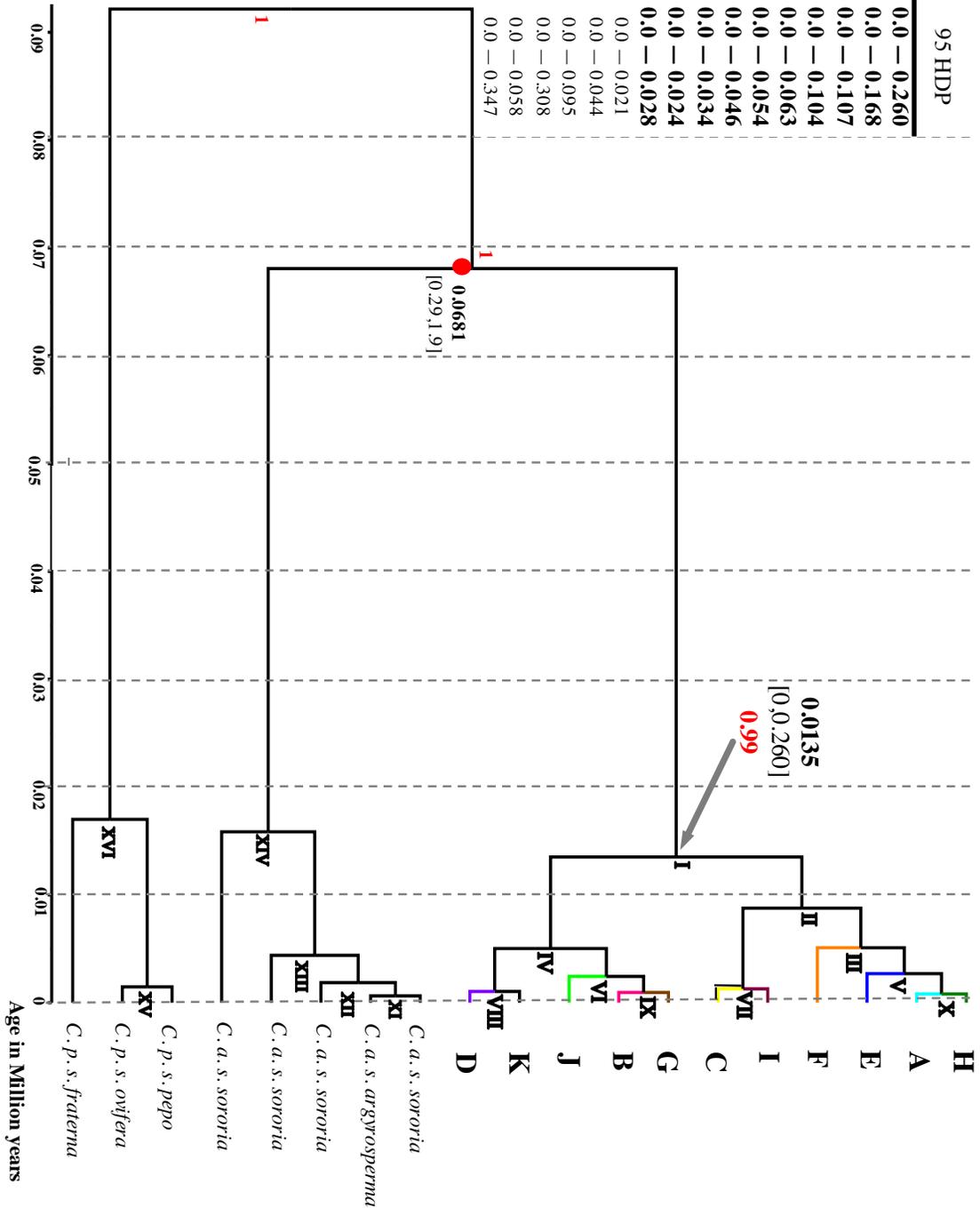


Figura 1. Cronograma de haplotipos de ADN mitocondrial de *Cucurbita moschata* (A–K) construido con BEAST. Los números romanos señalan cada nodo. La edad de cada nodo la probabilidad posterior y el valor de 95% HDP en la tabla en la parte superior izquierda de la figura. El círculo rojo representa el punto de calibración.

Discusión y conclusiones.

El análisis de tiempos de divergencia para los haplotipos de *C. moschata* proporciona una evidencia más para sugerir que su origen, domesticación y diversificación se llevó a cabo en Mesoamérica dentro de la región que corresponde a México al menos como un evento.

El tiempo de divergencia entre ambos haplogrupos o linajes de *C. moschata* (~13,000 años) es más antiguo que las fechas de domesticación para las calabazas determinado a partir de datos arqueológicos (~10,000 años; Whitaker and Cutler, 1971; Smith, 1997). Los datos arqueológicos más antiguos registrados para esta especie hasta este momento se han encontrado en México (Whitaker et al., 1957). Se debe tomar en cuenta que las especies comenzaron a divergir genéticamente antes de la domesticación y por ello los restos arqueológicos encontrados debieron haber sido más comunes. Además, la región donde se encontraron estos restos arqueobotánicos se encuentra muy cercana a la región de mayor diversidad genética a nivel mitocondrial reportado por Hernández-Rosales et al. (2020). Conjuntamente los haplotipos determinados como los más ancestrales (C y K) también se encuentran en Veracruz.

Las regiones de alta diversidad genética usualmente concuerdan con el centro de domesticación de las especies cultivadas, y el caso de *C. moschata* es muy parecido a lo que se ha reportado para *Campsicum* sp., en el que la región de alta diversidad de los chiles domesticados coincide con el centro de su domesticación (Kraft et al., 2014).

Aunque aún no se ha identificado al pariente silvestre directamente relacionado a la domesticación de esta especie, la distribución del taxón silvestre más estrechamente relacionado, *C. argyrosperma* spp. *sororia*, solo se distribuye en Mesoamérica (Sánchez-de la Vega, 2018). Al mismo tiempo, Tapia (2020) identificó que los haplotipos de *C. argyrosperma* spp. *sororia* filogenéticamente más relacionados a *C. moschata* se distribuyen en la vertiente del Pacífico entre los estados de Guerrero, Chiapas y en el Golfo de México en Veracruz. Otro grupo de evidencias que apoya su relación con *C. argyrosperma* spp. *sororia* es que *C. moschata* es capaz de producir híbridos altamente fértiles con esta especie (Merrick, 1990).

Todas las evidencias anteriores señalan que el ancestro de *C. moschata* pudo haber tenido un origen en Mesoamérica y es posible que *C. argyrosperma* spp. *sororia* sea el pariente silvestre directo.

La divergencia de los haplogrupos de *C. moschata* en México relacionados a un gradiente altitudinal es una evidencia para entender su amplia distribución en distintos ambientes. La

región de Veracruz (México) representa un posible centro de domesticación de esta especie. México es centro un centro de alta diversidad genética.

Perspectivas.

Para tener una mejor comprensión respecto al proceso de domesticación de *C. moschata* se debe considerar los siguientes elementos que mejorarían el fechamiento de los haplotipos:

- Integrar poblaciones que representen toda la distribución de las variedades tradicionales y razas criollas de la especie, con mayor énfasis en las regiones de Centroamérica y Sudamérica.
- Analizar la filogeografía de esta *C. moschata* con un mayor número de marcadores de mitocondria y con los SNPs disponibles para las especies del Grupo *Argyrosperma*.
- Integrar en los análisis de tiempos de divergencia los haplotipos reportados por Tapia (2020).
- Calibrar el reloj molecular con datos fósiles. Sin embargo, reconocemos que los registros fósiles son mucho más antiguos y no podrían ajustarse adecuadamente a los tiempos de domesticación (~12,000 años).

Literatura citada.

- Acevedo, F., H. Ocampo, S. Lorenzo, A. Sol, O. García, A. López, M. Canales, et al. 2009. La bioseguridad en México y los organismos genéticamente modificados: cómo enfrentar un nuevo desafío. II: 319-353.
- Aguirre-Dugua, X., L. E. Eguiarte, A. González-Rodríguez, y A. Casas. 2012. Round and large: Morphological and genetic consequences of artificial selection on the gourd tree *Crescentia cujete* by the Maya of the Yucatan Peninsula, Mexico. *Annals of Botany*. 109: 1297-1306.
- Aguirre-Liguori, J. A., B. S. Gaut, J. P. Jaramillo-Correa, M. I. Tenailon, S. Montes-Hernández, F. García-Oliva, S. J. Hearne, y L. E. Eguiarte. 2019. Divergence with gene flow is driven by local adaptation to temperature and soil phosphorus concentration in teosinte subspecies (*Zea mays parviglumis* and *Zea mays mexicana*). *Molecular Ecology*. 28: 2814-2830.
- Aguirre-Liguori, J. A., J. A. Luna-Sánchez, J. Gasca-Pineda, y L. E. Eguiarte. 2020. Evaluation of the Minimum Sampling Design for Population Genomic and Microsatellite Studies: An Analysis Based on Wild Maize. *Frontiers in Genetics*. 11: 1-16.
- Aguirre-Liguori, J. A., M. I. Tenailon, A. Vázquez-Lobo, B. S. Gaut, J. P. Jaramillo-Correa, S. Montes-Hernandez, V. Souza, y L. E. Eguiarte. 2017. Connecting genomic patterns of local adaptation and niche suitability in teosintes. *Molecular Ecology*. 26: 4226-4240.
- Albany, N., J. Vilchez, A. Nava, M. González, y C. Castro de Rincón. 1998. El análisis de conglomerado para complementar el estudio de patrones electroforéticos en *Psidium* spp. *Revista de la Facultad de Agronomía de la Universidad del Zulia*. 15: 142-152.
- de Albuquerque, H. Y. G., E. J. de Oliveira, A. C. Brito, L. R. B. de Andrade, C. D. do Carmo, C. V. Morgante, E. A. Vieira, et al. 2019. Identification of duplicates in Cassava Germplasm Banks based on single-nucleotide polymorphisms (SNPs). *Scientia Agricola*. 76: 328-336.
- Alexander, D. H., J. Novembre, y K. Lange. 2009. Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*. 19: 1655-1664.
- Allem, A. C. 2000. The terms genetic resource, biological resource, and biodiversity examined. *Environmentalist*. 20: 335-341.
- Altieri, M. A. 1995. Agroecology: The science of sustainable agriculture. 2nd Edition, Boulder, CO, Westview Press.
- Avise, J. C. 2009. Phylogeography: Retrospect and prospect. *Journal of Biogeography*. 36: 3-15.
- Azevedo-Meleiro, C. H., y D. B. Rodríguez-Amaya. 2007. Qualitative and quantitative

- differences in carotenoid composition among *Cucurbita moschata*, *Cucurbita maxima*, and *Cucurbita pepo*. *Journal of Agricultural and Food Chemistry*. 55: 4027-4033.
- Balvino-Olvera, F. J., K. F. Sánchez-Gómez, J. A. Lobo, G. Avila-Sakar, R. Cruz-Reyes, G. Sánchez-Montoya, Y. Herrerías-Diego, et al. 2017. Latitudinal structured populations of the Mexican wild squash *Cucurbita argyrosperma* spp. *sororia* revealed by microsatellite markers. *Crop and Pasture Sciences*. 68: 850-858.
- Bandelt, H. J., P. Forster, y A. Röhl. 1999. Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*. 16: 37-48.
- Barrera-Redondo, J., E. Ibarra-Laclette, A. Vázquez-Lobo, Y. T. Gutiérrez-Guerrero, G. Sánchez de la Vega, D. Piñero, S. Montes-Hernández, et al. 2019. The Genome of *Cucurbita argyrosperma* (Silver-Seed Gourd) Reveals Faster Rates of Protein-Coding Gene and Long Noncoding RNA Turnover and Neofunctionalization within *Cucurbita*. *Molecular Plant*. 12: 506-520.
- Barrera-Redondo, J., H. S. Hernández-Rosales, V. Cañedo-Torres, K. Aréstegui-Alegría, J. Torres-Guevara, F. Parra, I. Torres-García, y A. Casas. 2020. Landrace diversity and local selection criteria of domesticated squashes and gourds (*Cucurbita*) in the Central Andean mountain range of Peru: Tomayquichua, Huánuco. *Botanical Sciences*. 98: 101-116.
- Barrera-Redondo, J., G. S. Vega, J. A. Aguirre-Liguori, G. Castellanos-Morales, Y. T. Gutiérrez-Guerrero, X. Aguirre-dugua, E. Aguirre-Planter, et al. 2021. The domestication of *Cucurbita argyrosperma* as revealed by the genome of its wild relative. *Horticulture Research* 8:109.
- Bisognin, D. A. 2002. Origin and evolution of cultivated cucurbits. *Ciência Rural*. 32: 715-723.
- Bolger, A. M., M. Lohse, y B. Usadel. 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*. 30: 2114-2120.
- Bukasov, S. M. 1981. Las Plantas Cultivadas de México, Guatemala y Colombia. Turrialba: CATIE-GTZ (Project).
- Bunting, H. y B. Pickersgill, 1996. What is a plant genetic resource? *The Biologist*. 43: 227-230.
- Castellanos-Morales, G., L. M. Paredes-Torres, N. Gámez, H. S. Hernández-Rosales, G. Sánchez-de la Vega, J. Barrera-Redondo, E. Aguirre-Planter, et al. 2018. Historical biogeography and phylogeny of *Cucurbita*: Insights from ancestral area reconstruction and niche evolution. *Molecular Phylogenetics and Evolution*. 128: 38-54.
- Castellanos-Morales, G., K. Y. Ruiz-Mondragón, H. S. Hernández-Rosales, G. Sánchez-de la

- Vega, E. Aguirre-Planter, S. Montes-Hernández, R. Lira-Saade, y L. E. Eguiarte. 2019. Tracing back the origin of pumpkins (*Cucurbita pepo* ssp. *pepo* L.) in Mexico. *Proceedings of the Royal Society B*. 286: 20191440
- Catchen, J., P. A. Hohenlohe, S. Bassham, A. Amores, y W. A. Cresko. 2013. Stacks: An analysis tool set for population genomics. *Molecular Ecology*. 22: 3124-3140.
- Ceccarelli, S. 2009. Evolution, plant breeding and biodiversity. *Journal of Agriculture and Environment for International Development*. 103: 131-145.
- Clement, M., D. Posada, y K. A. Crandall. 2000. TCS: A computer program to estimate gene genealogies. *Molecular Ecology*. 9: 1657-1659.
- Darriba, D., G. L. Taboada, R. Doallo, y D. Posada. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*. 9: 772.
- Dayan, T., D. Simberloff, E. Tchernov, y Y. Yom-Tov. 1989. Interand intraspecific character displacement in mustelids. *Ecology*. 70: 1526-1539.
- Decker, D. S. 1988. Origin(s), evolution, and systematics of *Cucurbita pepo* (Cucurbitaceae). *Economic Botany*. 42: 4-15.
- Decker-Walters, D. S., T. W. Walters, U. Posluszny y P. G. Kevan. 1990. Genealogy and gene flow among anual domesticated species of *Cucurbita*. *Canadian Journal of Botany*. 68: 782-789.
- Dillehay, T. D., J. Rossen, T. C. Andres, y D. E. Williams. 2007. Preceramic adoption of peanut, squash, and cotton in Northern Peru. *Science*. 316: 1890-1893.
- Doebley J. 1989. Isozymic Evidence and the Evolution of Crop Plants. En: D. E. Soltis, P. S. Soltis, y T. R. Dudley (eds.). *Isozymes in Plant Biology*. Springer, Dordrecht. 165-191.
- Doebley, J. F., B. S. Gaut, y B. D. Smith. 2006. The Molecular Genetics of Crop Domestication. *Cell*. 127: 1309-1321.
- Doyle, J. J., y J. L. Doyle. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin*. 19: 11-15.
- Drozdowska, A., y T. Hermanowski. 2016. Predictors of generic substitution: The role of psychological, sociodemographic, and contextual factors. *Research in Social and Administrative Pharmacy*. 12: 119-129.
- Drummond, A. J., M. A. Suchard, D. Xie, y A. Rambaut. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*. 29: 1969-1973.

- Eguiarte, L. E., H. S. Hernández-Rosales, J. Barrera-Redondo, G. Castellanos-Morales, L. M. Paredes-Torres, G. Sánchez-de la Vega, K. Y. Ruiz-Mondragón, et al. 2018. Domesticación, diversidad y recursos genéticos y genómicos de México: El caso de las calabazas. *TIP Revista Especializada en Ciencias Químico-Biológicas*. 21: 85-101.
- Eguiarte, L. E., J. A. Aguirre-Liguori, L. Jardón-Barbolla, E. Aguirre-Planter, y V. Souza. 2013. Genómica de poblaciones: nada en evolución va a tener sentido si no es a la luz de la genómica, y nada en genómica tendrá sentido si no es a la luz de la evolución. *TIP Revista Especializada en Ciencias Químico-Biológicas*. 16: 42-56.
- Escalante, A. E., L. Jardón-Barbolla, S. Ramírez-Barahona, y L. E. Eguiarte. 2014. The study of biodiversity in the era of massive sequencing. *Revista Mexicana de Biodiversidad*. 85: 1249-1264.
- FAO. Food and Agriculture of the United Nations. 2019. FAOSTAT. Crops. Pumpkins, squash and gourds. Last Update: January 18, 2019. Consultado el 08 de octubre de 2020. URL: <http://www.fao.org/faostat/en/#data/QC/>.
- Ferriol, M., y B. Pico. 2008. Pumpkin and winter squash. En: J. Prohens y F. Nuez (eds.). *Handbook of Plant Breeding, vegetables I*. Springer, New York, USA.
- Filov, A. I. 1966. Ekologija i klasifikacija tykuy. *Bjulleten Glavnogo Botaniceskogo Sada*. 63: 33-41.
- Foll, M., and O. Gaggiotti. 2008. A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics*. 180: 977-993.
- Formisano, G., C. Roig, C. Esteras, M. R. Ercolano, F. Nuez, A. J. Monforte, y M. B. Picó. 2012. Genetic diversity of Spanish *Cucurbita pepo* landraces: An unexploited resource for summer squash breeding. *Genetic Resources and Crop Evolution*. 59: 1169-1184.
- Futuyma, D. J. 2005. *Evolution*. Sinauer Associates, Massachusetts.
- Gaut, B. S., C. M. Díez, y P. L. Morrell. 2015. Genomics and the Contrasting Dynamics of Annual and Perennial Domestication. *Trends in Genetics*. 31: 709-719.
- Gepts, P. 2003. Crop Domestication as a Long-Term Selection Experiment. En: *Plant Breeding Reviews*. J. Janick (ed.). Wiley Online Library. 1-44.
- Gliessmann, S. R. 1983. Allelopathic interactions in crop-weed mixtures: applications for weed management. *Journal Chemical Ecology*. 9: 991-999.
- Gliessmann, S. R. 1990. The ecology and management of traditional farming systems. En: M.A.

- Altieri y S. B. Hecht (eds.). *Agroecology, and Small Farm Development*. BocaRaton, Florida, CRC Press.
- Goldman, A. 2004. *The complete squash: a passionate grower's guide to pumpkins, squash, and gourds*. Artisan, N.Y., USA.
- Goldstein, D. B., y D. D. Pollock. 1994. Least squares estimation of molecular distance - noise abatement in phylogenetic reconstruction. *Theoretical Population Biology*. 45: 219-226.
- Gong, L., H. S. Paris, G. Stift, M. Pachner, J. Vollmann, y T. Lelley. 2013. Genetic relationships and evolution in *Cucurbita* as viewed with simple sequence repeat polymorphisms: the centrality of *C. okeechobeensis*. *Genetic Resources and Crop Evolution*. 60: 1531-1546.
- Gong, L., H. S. Paris, M. H. Nee, G. Stift, M. Pachner, J. Vollmann, y T. Lelley. 2012. Genetic relationships and evolution in *Cucurbita pepo* (pumpkin, squash, gourd) as revealed by simple sequence repeat polymorphisms. *Theoretical and Applied Genetics*. 124: 875-891.
- Goodwin, T. W. 1980. *The biochemistry of the carotenoids*, vol. 1. New York: Chapman and Hall.
- Grover, A., y P. C. Sharma. 2016. Development and use of molecular markers: *Past and present*. *Critical Reviews in Biotechnology*. 36: 290-302.
- Guerra-García, A., y D. Piñero. 2017. Current approaches and methods in plant domestication studies. *Botanical Sciences*. 95: 345-362.
- Guillén, S., A. Casas, T. Terrazas, E. Vega, y A. Martínez-Palacios. 2013. Differential survival and growth of wild and cultivated seedlings of columnar cacti: Consequences of domestication. *American Journal of Botany*. 100: 2364-2379.
- Hamilton, M. 2009. *Populations genetics*. Wiley-Blackwell. Malasia. 68-73, 105-149 y 235-281.
- Hammer, K. 1984. Das Domestikationssyndrom. *Die Kulturpflanze*. 32: 11-34.
- Hawkes, J. G., 1983. *The diversity of crop plants*. Harvard University Press, Cambridge.
- Hedrick, P. 2005. *Genetics of populations*. Jones and Bartlett Publishers. Sudbury, 53 Massachusetts.
- Hernández-Rosales, H. S., G. Castellanos-Morales, G. Sánchez-de la Vega, E. Aguirre-Planter, S. Montes-Hernández, R. Lira-Saade, y L. E. Eguiarte. 2020. Phylogeographic and population genetic analyses of *Cucurbita moschata* reveal divergence of two mitochondrial lineages linked to an elevational gradient. *American Journal of Botany*. 107: 510-525.
- Hernández Vargas, P. 2017. Documentación sobre la existencia de *Cucurbita máxima* Duch ex

- Lam. (Cucurbitaceae) en México. Tesis de Licenciatura, Universidad Nacional Autónoma de México, Facultad de Estudios Superiores Cuautitlán. Estado de México.
- Hernández-Xolocotzi, E. 1988. La agricultura tradicional en México. *Comercio Exterior*. 38: 673-678.
- Hickerson, M. J., B. C. Carstens, J. Cavender-Bares, K. A. Crandall, C. H. Graham, J. B. Johnson, L. Rissler, et al. 2010. Phylogeography's past, present, and future: 10 years after Avise, 2000. *Molecular Phylogenetics and Evolution*. 54: 291-301.
- Hillman, G. C., y M. S. Davies. 1990. Measured domestication rates in wild wheats and barley under primitive cultivation, and their archaeological implications. *Journal of World Prehistory*. 4: 157-222.
- Hoffmann, S., C. Otto, G. Doose, A. Tanzer, D. Langenberger, S. Christ, M. Kunz, et al. 2014. A multi-split mapping algorithm for circular RNA, splicing, trans-splicing and fusion detection. *Genome Biology*. 15: R34.
- Hubby, J. L., y R. C. Lewontin. 1966. A molecular approach to the study of genic heterozygosity in natural populations. I. The number of alleles at different loci in *Drosophila pseudoobscura*. *Genetics*. 54: 577-594.
- Jaenicke-Després, V., E. S. Buckler, B. D. Smith, M. T. P. Gilbert, A. Cooper, J. Doebley, y S. Pääbo. 2003. Early Allelic Selection in Maize as Revealed by Ancient DNA. *Science*. 302: 1206-1208.
- Jarret, R. L., I. J. Levy, T. L. Potter, S. C. Cermak, y L. C. Merrick. 2013. Seed oil content and fatty acid composition in a genebank collection of *Cucurbita moschata* Duchesne and *C. argyrosperma*. Huber. *Plant Genetic Resources: Characterisation and Utilization*. 11: 149-157.
- Jeffrey, H. 1961. The theory of probability. Oxford: Oxford University Press
- Jombart, T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*. 24: 1403-1405.
- Kantar, M. B., A. R. Nashoba, J. E. Anderson, B. K. Blackman, y L. H. Rieseberg. 2017. The Genetics and Genomics of Plant Domestication. *BioScience*. 67: 971-982.
- Kates, H. R., P. S. Soltis, y D. E. Soltis. 2017. Evolutionary and domestication history of *Cucurbita* (pumpkin and squash) species inferred from 44 nuclear loci. *Molecular Phylogenetics and Evolution*. 111: 98-109.

- Kimura, M. 1986. DNA and the neutral theory. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*.
- Kistler, L., L. A. Newsom, T. M. Ryan, A. C. Clarke, B. D. Smith, y G. H. Perry. 2015. Gourds and squashes (*Cucurbita* spp.) adapted to megafaunal extinction and ecological anachronism through domestication. *Proceedings of the National Academy of Sciences of the United States of America*. 112: 15107-15112.
- Kraft, K. H., C. H. Brown, G. P. Nabhan, E. Luedeling, J. de. J. Luna Ruiz, G. Coppens d'Eeckenbrugge, R. J. Hijmans, y P. Gepts. 2014. Multiple lines of evidence for the origin of domesticated chili pepper, *Capsicum annuum*, in Mexico. *Proceedings of the National Academy of Sciences of the United States of America*. 111: 6165-6170.
- Kumar, S. 2005. Molecular clocks: four decades of evolution. *Nature Reviews Genetics*. 6: 554-566.
- Lambeck, K., y J. Chappell. 2001. Sea level change through the last glacial cycle. *Science*. 292: 679-686.
- Larson, G., D. R. Piperno, R. G. Allaby, M. D. Purugganan, L. Andersson, M. Arroyo-Kalin, L. Barton, et al. 2014. Current perspectives and the future of domestication studies. *Proceedings of the National Academy of Sciences of the United States of America*. 111: 6139-6146.
- Lee, H. Y., S. Jang, C. R. Yu, B. C. Kang, J. H. Chin, y K. Song. 2021. Population structure and genetic diversity of *Cucurbita moschata* based on genome-wide high-quality snps. *Plants*. 10: 1-10.
- Lin, T., G. Zhu, J. Zhang, X. Xu, Q. Yu, Z. Zheng, Z. Zhang, et al. 2014. Genomic analyses provide insights into the history of tomato breeding. *Nature Genetics*. 46: 1220-1226
- Lira, R. 1985. Identidad taxonómica de las calabazas cultivadas (*Cucurbita* spp.) en la península de Yucatán. *Biótica*. 10: 301-307.
- Lira, R. 1988. Cucurbitaceae de la Península de Yucatán: taxonomía y etnobotánica. Tesis de Maestría, Instituto Nacional de Investigaciones sobre Recursos Bióticos, Xalapa, Veracruz, México. 221 pp.
- Lira, R., Eguiarte, L., Montes, S., Zizumbo-Villarreal, D., Marín, P. C. G., y Quesada, M. 2016. Homo sapiens-*Cucurbita* interaction in mesoamerica: domestication, dissemination, and diversification. En: R. Lira, A. Casas, y J. Blancas (eds.). *Ethnobotany of Mexico*. New

- York, NY: Springer. 389-401.
- Litt, M., y J. A. Luty. 1989. A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *American Journal of Human Genetics*. 44: 397-401.
- Lozada-Aranda, M., I. Rojas-Barrera, A. Mastretta-Yanes, A. Ponce-Mendoza, C. Burgeff, M. A. Orjuela, y G. Oliveros. 2017. Las milpas de México. *Oikos*. 17: 10-12.
- Lust, T. A., y H. S. Paris. 2016. Italian horticultural and culinary records of summer squash (*Cucurbita pepo*, Cucurbitaceae) and emergence of the zucchini in 19th-century Milan. *Annals of Botany*. 118: 53-69.
- Martínez-González C., Castellanos-Morales G., Barrera-Redondo J., Sánchez-de la Vega G., Hernández-Rosales H.S., Gasca-Pineda J., Aguirre-Planter E., Moreno-Letelier A., Escalante A.E., Montes-Hernández S., Lira-Saade R. and Eguiarte L. 2021. Recent and Historical Gene Flow in Cultivars, Landraces, and a Wild Taxon of *Cucurbita pepo* in Mexico. *Frontiers in Ecology and Evolution* 9:656051.
- Mastretta-Yanes, A., F. Acevedo-Gasman, C. Burgeff, M. Cano-Ramírez, D. Piñero, y J. Sarukhán. 2018. An Initiative for the Study and Use of Genetic Diversity of Domesticated Plants and Their Wild Relatives. *Frontiers in Plant Science* 9: 1-7.
- Mastretta-Yanes, A., M. R. Bellon, F. Acevedo, C. Burgeff, D. Piñero, y J. Sarukhán. 2019. A mexican program for the conservation and use of genetic diversity of domesticated plants and their wild relatives. *Revista Fitotecnia Mexicana*. 42: 321-334.
- Matsuoka, Y., Y. Vigouroux, M. M. Goodman, J. G. Sanchez, E. Buckler, y J. Doebley. 2002. A single domestication for maize shown by multilocus microsatellite genotyping. *Proceedings of the National Academy of Sciences of the United States of America*. 99: 6080-6084.
- Merrick, L. C. 1990. Systematics and evolution of a domesticated squash, *Cucurbita argyrosperma*, and its wild and weedy relatives. En: D. M. Bates, R. W. Robinson, y C. Jeffrey (eds.). *Biology and utilization of the Cucurbitaceae*. Cornell University Press, Ithaca, NY. 77-95.
- Merrick, L. C. 1991. Systematic, Evolution, and Ethnobotany of a Domesticated Squash, its Wild Relatives and Allied Species in the Genus *Cucurbita*. Ph.D. Thesis, Cornell University, Ithaca, New York. 315 pp.
- Merrick, L. C. Squashes, pumpkins and gourds. 1995. En: J. Smartt, y N. W. Simmonds (eds.).

- Evolution of crop plants. London: Longman Scientific and Technical. 97-105.
- Meyer, R. S., A. E. Duval, y H. R. Jensen. 2012. Patterns and processes in crop domestication: An historical review and quantitative analysis of 203 global food crops. *New Phytologist*. 196: 29-48.
- Meyer, R. S., y M. D. Purugganan. 2013. Evolution of crop species: genetics of domestication and diversification. *Nature Reviews Genetics*. 14: 840-852.
- Montero-Pau, J., J. Blanca, A. Bombarely, P. Ziarolo, C. Esteras, C. Martí-Gómez, M. Ferriol, et al. 2018. De novo assembly of the zucchini genome reveals a whole-genome duplication associated with the origin of the Cucurbita genus. *Plant Biotechnology Journal*. 16: 1161-1171.
- Montes-Hernández S.; y L. E. Eguiarte. 2002. Genetic structure and indirect estimates of gene flow in three taxa of *Cucurbita* (Cucurbitaceae) in western Mexico. *American Journal of Botany*. 89: 1156-1163.
- Moreno-Letelier, A., J. A. Aguirre-Liguori, D. Piñero, A. Vázquez-Lobo, y L. E. Eguiarte. 2020. The relevance of gene flow with wild relatives in understanding the domestication process. *Royal Society Open Science*. 7: 191545.
- Mueller, U. G., N. M. Gerardo, D. K. Aanen, D. L. Six, y T. R. Schultz. 2005. The evolution of agriculture in insects. *Annual Review of Ecology, Evolution, and Systematics*. 36: 563-595.
- Nee, M. 1990. The domestication of *Cucurbita* (Cucurbitaceae). *Economic Botany*. 44: 56-68.
- Nguyen, N. N., M. Kim, J. K. Jung, E. J. Shim, S. M. Chung, Y. Park, G. P. Lee, y S. C. Sim. 2020. Genome-wide SNP discovery and core marker sets for assessment of genetic variations in cultivated pumpkin (*Cucurbita* spp.). *Horticulture Research*. 7: 121.
- Nordborg, M. 2001. Coalescent theory. En: D. J. Balding, M. Bishop, y C. Cannings (eds.). *Handbook of statistical genetics*. John Wiley and Sons Inc., Chichester. 179-212.
- Ochoa-Zavala, M., P. Suárez-Montes, M. Chávez-Pesqueira, D. E. López-Cobos, C. Julia-Figueroa, A. David, H. Raquel, et al. 2016. Diferencias en morfología y visitantes florales entre *Lupinus mutabilis* y *Lupinus aff. ballianus* en la microcuenca de Warmiragra, Perú. En: *Domesticación en el continente americano. Manejo de biodiversidad y evolución dirigida por las culturas del Nuevo Mundo*. 463-479.
- Olsen, K. M., y J. F. Wendel. 2013. A Bountiful Harvest: Genomic Insights into Crop Domestication Phenotypes. *Annual Review of Plant Biology*. 64: 47-70.

- Organization for Economic Co-operation and Development (OECD). 2012. Consensus document on the biology of *Cucurbita* L. (squashes, pumpkins, zucchinis, and gourds). Series on harmonization of regulatory oversight in biotechnology. Paris, France.
- Ott, A., S. Liu, J. C. Schnable, C. T. Yeh, K. S. Wang, y P. S. Schnable. 2017. tGBS® genotyping-by-sequencing enables reliable genotyping of heterozygous loci. *Nucleic Acids Research*. 45: e178
- Paredes-Torres, L. 2016. Filogenia molecular del género *Cucurbita* (Cucurbitaceae) usando secuencias de cloroplasto. Tesis de Licenciatura, Universidad Nacional Autónoma de México, Facultad de Estudios Superiores Iztacala. Estado de México.
- Paris, H. S. 2016. Genetic Resources of Pumpkins and Squash, *Cucurbita* spp. En: R. Grumet R, N. Katzir, J. Garcia-Mas. (eds.). Genetics and Genomics of Cucurbitaceae. Plant Genetics and Genomics: Crops and Models, vol 20. Springer, Cham. 111-154.
- Paris, H. S., N. Yonash, V. Portnoy, N. Mozes-Daube, G. Tzuri, y N. Katzir. 2003. Assessment of genetic relationships in *Cucurbita pepo* (Cucurbitaceae) using DNA markers. *Theoretical and Applied Genetics*. 106: 971-978.
- Pickersgill, B. 2007. Domestication of plants in the Americas: Insights from Mendelian and molecular genetics. *Annals of Botany*. 100: 925-940.
- Piñero, D., Caballero-Mellado, J., Cabrera-Toledo, D., et al. 2008. La diversidad genética como instrumento para la conservación y el aprovechamiento de la biodiversidad: estudios en especies mexicanas. En: R. Dirzo, R. González, y I. J. March (eds.). Capital Natural de México, Mexico City: Comisión Nacional para el Conocimiento y Uso de la Biodiversidad. 437-494.
- Plummer, M., N. Best, K. Cowles, y K. Vines. 2006. CODA: convergence diagnosis and output analysis for MCMC. *R News*. 6: 7-11.
- Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, M. A. R. Ferreira, D. Bender, J. Maller, et al. 2007. PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*. 81: 559-575.
- Purugganan, M. D., y D. Q. Fuller. 2011. Archaeological data reveal slow rates of evolution during plant domestication. *Evolution*. 65: 171-183.
- Rambaut, A., A. J. Drummond, D. Xie, G. Baele, y M. A. Suchard. 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*. 67: 901-

904.

- Rambaut, A., M. Suchard, y A. Drummond. 2014. Tracer v 1.6 Available at. Rambaut, A., and A. Drummond. 2013. TreeAnnotator v 1.7. Available as part of the BEAST package at <http://beast.bio.ed.ac.uk>
- Ranere, A. J., D. R. Piperno, I. Holst, R. Dickau, y J. Iriarte. 2009. The cultural and chronological context of early Holocene maize and squash domestication in the central balsas river Valley, Mexico. *Proceedings of the National Academy of Sciences of the United States of America*. 106: 5014-5018.
- Rosenberg, N. A., y M. Nordborg. 2002. Genealogical trees, coalescent theory and the analysis of genetic polymorphisms. *Nature Reviews Genetics*. 3: 380-390.
- Ruíz-Mondragón, K. Y. 2017. Filogeografía y diversidad genética en *Cucurbita pepo* L. 1753 en México. Tesis de Licenciatura, Universidad Nacional Autónoma de México, Facultad de Ciencias. CDMX.
- Saisho, D., y M. D. Purugganan. 2007. Molecular phylogeography of domesticated barley traces expansion of agriculture in the old world. *Genetics*. 177: 1765-1776.
- Sánchez-de la Vega, G., G. Castellanos-Morales, N. Gámez, H. S. Hernández-Rosales, A. Vázquez-Lobo, E. Aguirre-Planter, J. P. Jaramillo-Correa, et al. 2018. Genetic Resources in the “Calabaza Pipiana” Squash (*Cucurbita argyrosperma*) in Mexico: Genetic Diversity, Genetic Differentiation and Distribution Models. *Frontiers in Plant Science*. 9: 400.
- Sanjur, O. I., D. R. Piperno, T. C. Andres, y L. Wessel-Beaver. 2002. Phylogenetic relationships among domesticated and wild species of *Cucurbita* (Cucurbitaceae) inferred from a mitochondrial gene: Implications for crop plant evolution and areas of origin. *Proceedings of the National Academy of Sciences of the United States of America*. 99: 535-540.
- Schaffer, A. A., C. D. Boyer, y H. S. Paris. 1986. Inheritance of rind lignification and warts in *Cucurbita pepo* L. and a role for phenylalanine ammonia lyase in their control. *Z Pflanzenzücht*. 96:1 47-53.
- Sinha, A. K., y P. K. Mishra. 2012. Agronomic evaluation of landraces of rice (*Oryza sativa*) of Bankura district of West Bengal. *Columban Journal of Life Science*. 13: 39-42.
- Smith, B. D. 1997. The Initial Domestication of *Cucurbita pepo* in the Americas 10,000 Years Ago. *Science*. 276: 932-934.
- Stetter, M. G., D. J. Gates, W. Mei, y J. Ross-Ibarra. 2017. How to make a domesticate. *Current*

- Biology*. 27: R896–R900.
- Sun, H., S. Wu, G. Zhang, C. Jiao, S. Guo, Y. Ren, J. Zhang, et al. 2017. Karyotype Stability and Unbiased Fractionation in the Paleo-Allotetraploid Cucurbita Genomes. *Molecular Plant*. 10: 1293-1306.
- Taberlet, P., L. Gielly, G. Pautou, y J. Bouvet. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology*. 17: 1105-1109.
- Tanto Hadado, T., D. Rau, E. Bitocchi, y R. Papa. 2010. Adaptation and diversity along an altitudinal gradient in Ethiopian barley (*Hordeum vulgare* L.) landraces revealed by molecular analysis. *BMC Plant Biology*. 10: 121.
- Tapia Aguirre, F. 2020. Filogeografía y diversidad genética de *Cucurbita argyrosperma* ssp. *sororia*. Tesis de Licenciatura, Universidad Nacional Autónoma de México, Facultad de Estudios Superiores Iztacala. Estado de México.
- Van De Wouw, M., C. Kik, T. Van Hintum, R. Van Treuren, y B. Visser. 2010. Genetic erosion in crops: Concept, research results and challenges. *Plant Genetic Resources: Characterisation and Utilization* 8: 1-15.
- Van Heerwaarden, J., J. Doebley, W. H. Briggs, J. C. Glaubitz, M. M. Goodman, J. D. J. S. Gonzalez, y J. Ross-Ibarra. 2011. Genetic signals of origin, spread, and introgression in a large sample of maize landraces. *Proceedings of the National Academy of Sciences of the United States of America*. 108: 1088-1092.
- Vázquez-Domínguez, E., S. Castañeda-Rico, T. Garrido-Garduño, y T. A. Gutiérrez-García. 2009. Avances metodológicos para el estudio conjunto de la información genética, genealógica y geográfica en análisis evolutivos y de distribución. *Revista Chilena de Historia Natural*. 82: 277-297.
- Velásquez-Milla, D., A. Casas, J. Torres-Guevara, y A. Cruz-Soriano. 2011. Ecological and socio-cultural factors influencing in situ conservation of crop diversity by traditional Andean households in Peru. *Journal of Ethnobiology and Ethnomedicine*. 7: 40.
- De Villemereuil, P., y O. E. Gaggiotti. 2015. A new FST-based method to uncover local adaptation using environmental variables. *Methods in Ecology and Evolution*. 6: 1248-1258.
- Whitaker, T. W. 1947. American Origin of the Cultivated Cucurbits. I. Evidence from the Herbs. II. *Survey of Old and Recent Botanical Evidence*. *Annals of the Missouri Botanical Garden*. 34: 101-111.

- Whitaker, T. W., y J. B. Bird. 1949. Identification and significance of the Cucurbit materials from Huaca Prieta, Peru. *American Museum Novitates*. Nueva York, USA.
- Whitaker, T. W., y H. C. Cutler. 1965. Cucurbits and cultures in the Americas. *Economic Botany* 19: 344-349.
- Whitaker, T. W., y H. C. Cutler. 1971. Pre-historic cucurbits from the Valley of Oaxaca. *Economic Botany*. 25: 123-127.
- Whitaker, T. W., H. C. Cutler, y R. S. Macneish. 1957. Cucurbit Materials from 3 Caves near Ocampo, Tamulipas. *American Antiquity*. 22: 352-358.
- Wien, H. C., S. C. Stapleton, D. N. Maynard, C. McClurg, R. Nyankanga, y D. Riggs. 2002. Regulation of female flower development in pumpkin (*Cucurbita* spp.) by temperature and light. En: D. N. Maynard (ed.). *Cucurbitaceae*. Alexandria: ASHS Press. 307-315.
- Wilson, H. D. 1989. Discordant patterns of alloenzyme and morphological variation in Mexican *Cucurbita*. *Systematic Botany*. 14: 612-623.
- Wilson, H. D., J. Doebley, y M. Duvall. 1992. Chloroplast DNA diversity among wild and cultivated members of *Cucurbita* (Cucurbitaceae). *Theoretical and Applied Genetics*. 84: 859-865.
- Xu, J., Y. Yuan, Y. Xu, G. Zhang, X. Guo, F. Wu, Q. Wang, et al. 2014. Identification of candidate genes for drought tolerance by whole-genome resequencing in maize. *BMC Plant Biology*. 14: 83.
- Yao, Z., F. M. You, A. N'Diaye, R. E. Knox, C. McCartney, C. W. Hiebert, C. Pozniak, y W. Xu. 2020. Evaluation of variant calling tools for large plant genome re-sequencing. *BMC Bioinformatics*. 21: 360.
- Zeder, M. A. 2006. Central questions in the domestication of plants and animals. *Evolutionary Anthropology* 15: 105-117.
- Zeder, M. A. 2015. Core questions in domestication research. *Proceedings of the National Academy of Sciences of the United States of America*. 112: 3191-3198.
- Zeder, M. A., E. Emshwiller, B. D. Smith, y D. G. Bradley. 2006. Documenting domestication: The intersection of genetics and archaeology. *Trends in Genetics*. 22: 139-155.
- Zohary, D. 2004. Unconscious selection and the evolution of domesticated Plants. *Economic Botany*. 58: 5-10.