



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
POSGRADO EN CIENCIAS BIOLÓGICAS
FACULTAD DE ESTUDIOS SUPERIORES IZTACALA
BIOLOGÍA EVOLUTIVA

DIVERSIDAD Y ESTRUCTURA GENÉTICA EN LOCI NEUTRALES DE
***CUCURBITA PEPO* SUBSP. *PEPO* EN MÉXICO**

TESIS

QUE PARA OPTAR POR EL GRADO DE:
MAESTRA EN CIENCIAS BIOLÓGICAS

PRESENTA:

CARMINA MARTÍNEZ GONZÁLEZ

TUTOR PRINCIPAL DE TESIS: DR. RAFAEL LIRA SAADE
FACULTAD DE ESTUDIOS SUPERIORES IZTACALA, UNAM

COMITÉ TUTOR: DRA. ALEJANDRA CITLALI MORENO LETELIER
INSTITUTO DE BIOLOGÍA, UNAM
DR. ANTONIO GONZÁLEZ RODRÍGUEZ
INSTITUTO DE INVESTIGACIONES EN ECOSISTEMAS Y SUSTENTABILIDAD,
UNAM

CD. MX.

OCTUBRE, 2019



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

OFICIO CPCB/1044/2019

Asunto: Oficio de Jurado para Examen de Grado.

M. en C. Ivonne Ramírez Wence
Directora General de Administración Escolar, UNAM
Presente

Me permito informar a usted que en la reunión ordinaria del Comité Académico del Posgrado en Ciencias Biológicas, celebrada el día 26 de agosto de 2019, se aprobó el siguiente jurado para el examen de grado de **MAESTRA EN CIENCIAS BIOLÓGICAS** de la alumna **MARTINEZ GONZALEZ CARMINA**, con número de cuenta 308168736, con la tesis titulada "**Caracterización de la diversidad y estructura genética en loci neutrales de *Cucurbita pepo* subsp. *pepo* en México**". realizada bajo la dirección del **DR. RAFAEL LIRA SAADE**:

Presidente: DR. DANIEL IGNACIO PIÑERO DALMAU
Vocal: DR. LUIS ENRIQUE EGUIARTE FRUNS
Secretario: DRA. ALEJANDRA MORENO LETELIER
Suplente: DRA. GABRIELA CASTELLANOS MORALES
Suplente: DRA. ERIKA AGUIRRE PLANTER

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

ATENTAMENTE
"POR MI RAZA HABLARA EL ESPIRITU"
Cd. Universitaria, Cd. Mx., a 27 de septiembre de 2019.

DR. ADOLFO GERARDO NAVARRO SIGÜENZA
COORDINADOR DEL PROGRAMA



c.c.p. Expediente del (la) interesado (a)

Agradecimientos institucionales

Agradezco al Posgrado en Ciencias Biológicas, al Instituto de Ecología y a la Facultad de Estudios Superiores Iztacala de la Universidad Nacional Autónoma de México por brindarme la oportunidad y el apoyo para realizar mis estudios de maestría.

Al Consejo Nacional de Ciencia y Tecnología (CONACyT) por concederme la beca apoyo a los estudiantes de maestría, ya que sin ella no habría sido posible llevar a cabo mi proyecto de tesis.

El financiamiento de este proyecto es gracias a CONABIO con el proyecto KE004 "Diversidad genética de las especies de *Cucurbita* en México e hibridación entre plantas genéticamente modificadas y especies silvestres de *Cucurbita*", el proyecto PE001 "Diversidad genética de las especies de *Cucurbita* en México. Fase II. Genómica evolutiva y de poblaciones, recursos genéticos y domesticación", el proyecto CONACyT de Investigación Científica Básica 2011.167826 "Genómica de poblaciones: estudios en el maíz silvestre, el teosinte (*Zea mays* ssp. *parviglumis* y *Z. mays* ssp. *mexicana*)", y el proyecto CONACyT Problemas Nacionales 247730.

Al Dr. Rafael Lira Saade por aceptarme como su alumna, y orientarme en el camino de la investigación. A mi comité tutor integrado por la Dra. Alejandra Citlali Moreno Letelier y el Dr. Antonio González Rodríguez que, con su apoyo y comentarios contribuyeron a mejorar el proyecto.

Agradecimientos a título personal

Al jurado de examen de grado integrado por el Dr. Daniel Ignacio Piñero Dalmau, el Dr. Luis Enrique Eguiarte Fruns, la Dra. Alejandra Moreno Letelier, Dra. Gabriela Castellanos Morales y la Dra. Erika Aguirre Planter.

Al Banco de Germoplasma del Campo Experimental Bajío del Instituto Nacional de Investigación Forestal, Agrícola y Pecuarias en Celaya Guanajuato, por aportar gran parte de las accesiones analizadas en este trabajo.

Al Laboratorio de Evolución Molecular y Experimental, así como al Dr. Luis Enrique Eguiarte Fruns, por proporcionarme un espacio para trabajar, así como acceso a las computadoras y el resto de la infraestructura que me permitió realizar todos los análisis y escrito de la tesis.

Un agradecimiento especial a la Dra. Gabriela Castellanos Morales, quien a pesar de no formar parte de mi comité tutorial me proporcionó los datos para realizar esta tesis, así como su apoyo y asesoramiento durante la elaboración del anteproyecto, todos los análisis y escrito de la tesis.

Por su ayuda en la realización de este trabajo: Josué Barrera Redondo, Gabriela Castellanos Morales, Canek Rivera Arroyo, Gerardo Soria Ortiz, Yocelyn Gutiérrez, Manuel García Ulloa, Jonás Aguirre Liguori, Helena Hernández Rosales, Guillermo Sánchez de la Vega, Paulina Hernández, Karen Ruiz Mondragón, Leslie Mariel Paredes Torres, Manuel Rosas, Enrique Scheinvar, Valeria Souza, Laura Espinosa Asuar y Erika Aguirre Palmer.

A mi abuelita, Guadalupe Luna, por enseñarme las cosas más importantes de la vida, por cuidarme, criarme y acompañarme durante toda mi vida. Sé que siempre creíste en mí, que me veías crecer y avanzar hacia mis metas. Te amo mucho abuelita, y aunque ya no estés aquí conmigo, siempre te recuerdo y sigues siendo todo un ejemplo de vida para mí.

A mis padres Irma González Luna y Juan Manuel Martínez Villegas, por brindarme su apoyo incondicional, guiarme y acompañarme en todas las decisiones de mi vida. A mis hermanos Alejandro y Aranza, porque siempre me motivaron a

seguir adelante, no rendirme, y a demostrarme a mí misma y a los demás que sin importar la adversidad, yo sería capaz de superarla.

A todas las personas que creyeron en mí y me ayudaron a superar al monstruo y seguir adelante, que estuvieron conmigo en el momento más difícil de mi vida, me dijeron que no me rindiera y continuara con mi sueño de dedicarme a hacer ciencia. Algunos incluso me dieron la oportunidad de trabajar en su laboratorio. No pondré nombres porque podría comprometer a algunas personas, pero creo que todos y cada uno de ustedes saben cuánto los aprecio, y cuán agradecida estoy que existan en mi vida. A todos ustedes, en verdad, gracias.

Índice

Resumen.....	1
Abstract.....	2
1. Introducción.....	3
1.1 Métodos de secuenciación masiva	6
1.2 Preparación de librerías.....	8
1.3 Relevancia del proyecto.....	9
2. Especie de estudio.....	12
2.1 Importancia del género	12
2.2 Generalidades taxonómicas.....	12
2.3 <i>Cucurbita pepo</i>	13
2.3.1 Características taxonómicas de las variedades silvestres.....	13
2.3.2 Características taxonómicas de las variedades cultivadas.....	14
2.4 Historia de domesticación de <i>Cucurbita</i>	16
2.4.1 Descubrimientos arqueológicos	16
2.4.2 Síndrome de domesticación	18
2.5 Trabajos previos de diversidad y estructura genética en <i>C. pepo</i>	19
3. Objetivos.....	23
3.1 Objetivo general.....	23
3.2 Objetivos particulares.....	23
4. Hipótesis.....	24
5. Método	25
5.1 Obtención de muestras.....	25
5.2 Extracción del DNA	26
5.3 Procesamiento de datos	27
5.4 Análisis de datos.....	28
5.4.1 Diversidad genética.....	28
5.4.2 Estructura genética	29
5.4.3 Flujo génico	30
6. Resultados	31
6.1 Calidad de datos.....	31

6.2 Diversidad genética	31
6.3 Estructura genética.....	33
6.4 Flujo génico.....	38
7. Discusión	40
7.1 Calidad de datos	40
7.2 Diversidad genética	41
7.3 Estructura, diferenciación y flujo genético	44
7.4 Relevancia de los recursos genéticos y su conservación	48
8. Conclusiones	51
9. Perspectivas	52
Bibliografía.....	54

Resumen

Entre los recursos fitogenéticos más importantes del mundo se encuentran las calabazas (*Cucurbita* spp.). En particular *C. pepo*, que cuenta con una historia de domesticación de al menos 10,000 años, siendo considerada la calabaza cultivada con mayor importancia económica a nivel mundial. En México, *C. pepo* ssp. *pepo* se ha cultivado extensivamente, tanto en el sistema tradicional de cultivo (milpa) como en cultivos más tecnificados, y como consecuencia de esto se ha propiciado un constante flujo genético entre las diferentes regiones de México y entre variedades de *C. pepo* tanto comerciales como silvestres. El objetivo de este trabajo fue caracterizar los niveles de diversidad y estructura genética en loci neutrales de las diferentes variedades de *C. pepo* y cultivos de *C. pepo* subsp. *pepo* en México utilizando una aproximación genómica. Para ello se analizarán datos de tGBS para 14 poblaciones mexicanas (n = 69 plantas); también se analizaron variedades comerciales de EUA de *C. pepo* (*Black Beauty Zuchinni*, *Cocozelle*, *Delicata Honey Boat*, *Straightneck* y *Spaghetti*; n = 12), *C. pepo* subsp. *fraterna* (n = 8) y *C. pepo* subsp. *ovifera* var. *ovifera* (n = 6). Se encontró que con datos genómicos (12,881 SNPs) la estructura genética de las diferentes subespecies y variedades de *C. pepo* difiere de lo previamente reportado: si bien se mantienen tres grupos genéticos, estos no se separan según la clasificación taxonómica, sino que se definen en función de su cercanía evolutiva a *C. pepo* subsp. *fraterna* o a *C. pepo* subsp. *ovifera*, o a la probabilidad de ser variedades comerciales. También encontramos que las variedades de *C. pepo* subsp. *pepo* cultivadas en México están siendo desplazadas por variedades comerciales mejoradas, especialmente en las regiones del Bajío y norte de México. Las variedades comerciales suelen tener mayor resistencia a ciertas plagas, además de representar un mejor precio en el mercado, pero su popularidad ponen en riesgo la diversidad genética de las variedades tradicionales, afectando también el futuro de este valioso recurso fitogenético.

Abstract

Among the most important plant genetic resources in the world are the pumpkins (*Cucurbita* spp.). In particular, *C. pepo* has a history of domestication of at least 10,000 years and it is considered the pumpkin with greater economic importance worldwide. In Mexico, *C. pepo* ssp. *pepo* is cultivated extensively, both in the traditional cultivation system (milpa), as well as in more technical crops, and as a consequence there has been gene flow between different regions of Mexico and with other varieties of *C. pepo*, both commercial and wild. The objective of this thesis was to characterize the levels of diversity and genetic structure in neutral loci of the different varieties of *C. pepo* and crops of *C. pepo* subsp. *pepo* in Mexico using a genomic approach. For this, tGBS data for 14 Mexican populations (n = 69 plants) were analyzed, including also commercial US varieties from *C. pepo* (Black Beauty Zuchinni, Cocozelle, Delicata Honey Boat, Straightneck and *Spaghetti*; n = 12), *C. pepo* subsp. *fraterna* (n = 8) and *C. pepo* subsp. *ovifera* var. *ovifera* (n = 6). We found 12,881 SNPs that the genetic structure of the different subspecies and varieties of *C. pepo* differs with genomic data from what was previously reported using other genetic markers. Although there are three genetic groups, these are not separated according to the taxonomic classification, but they are defined according to their proximity to *C. pepo* subsp. *fraterna*, or to *C. pepo* subsp. *ovifera*, or the probability of being commercial varieties. We also found that traditional varieties of *C. pepo* subsp. *pepo* grown in Mexico are being displaced by improved commercial varieties, especially in the Bajío and northern regions of Mexico. Commercial varieties tend to have greater resistance to various pests, in addition to having a higher market price, but the genetic diversity of traditional varieties may be at risk, also affecting the future of this valuable plant genetic resource.

1. Introducción

La domesticación no es un evento instantáneo en el que un organismo silvestre se transforma repentinamente en domesticado. Más bien es un largo proceso de selección consciente o inconsciente realizado por los humanos para hacer a los organismos más benéficos para un uso posterior. Incluso se ha sugerido que el concepto de domesticación se puede aplicar al manejo del mismo ambiente (Casas et al., 1996; Gepts y Papa, 2003; Kareiva et al., 2007; Smith y Zeder, 2013; Zeder, 2015). Podemos considerar a este proceso en términos de un mutualismo entre el domesticador y la especie domesticada, ya que ambas poblaciones asociadas con el tiempo se vuelven cada vez más interdependientes (Zeder et al., 2006).

Durante milenios y hasta el presente las comunidades rurales han realizado un uso de los recursos que tienen a su disposición. Estas prácticas normalmente se incluyen en una categoría general de aprovechamiento de los recursos denominado *forrajeo*, la cual implica la recolección de productos naturales dentro de su hábitat (González-Insuasti y Caballero, 2007; Casas et al., 2017). Esta práctica puede ser eventual o sistemática, circunstancial o planificada, aleatoria o siguiendo una estrategia, manual o involucrando herramientas, generalizada o selectiva. Y cada forma de forrajeo es clave para entender los factores que pueden llegar a detonar el manejo de los recursos (Casas et al., 2017).

En México, podemos identificar un amplio espectro de formas de manejo, los cuales presentan dos categorías generales, el manejo *in situ* y el manejo *ex situ*. La primera forma de manejo ocurre en el lugar donde se encuentran originalmente las plantas. Bajo este manejo, las personas pueden realizar una recolección, protección, tolerancia o incluso promoción del recurso deseado, de tal forma que con el tiempo las características de este recurso van cambiando. En la segunda, el manejo ocurre fuera del ámbito natural, en lugares artificiales como los huertos, campos de cultivo y bancos de semillas donde la gente lleva numerosos ejemplares de plantas (Blancas et al., 2010; Casas et al., 2017).

Fuera de estas formas de manejo, podemos decir que en general una vez que se inicia el proceso de domesticación, a través de la selección artificial se

incrementa la frecuencia de las características deseables de los individuos seleccionados. Estos fenotipos de domesticación son rasgos que se seleccionan durante la transformación inicial y el establecimiento de las nuevas variedades domesticadas a partir de su ancestro silvestre. La selección gradual de estos caracteres produce cambios morfológicos y fisiológicos importantes, por lo que a menudo las plantas de cultivo se vuelven bastante diferentes de sus progenitores silvestres y, en los casos más extremos de domesticación, las plantas no pueden ya sobrevivir sin la intervención humana (Gepts y Papa, 2003; Meyer y Purugganan, 2013; Zizumbo-Villarreal, 1993).

A este fenómeno se le conoce como síndrome de domesticación, donde los rasgos que se seleccionaron o se siguen seleccionando porque confieren adaptación a un ambiente cultivado, a menudo incluyen la pérdida de latencia, aumento en el tamaño de las semillas, cambios en la arquitectura de los brotes reproductivos, hábitos de crecimiento más compacto y menos ramificado, mayor producción o aumento de tamaño en frutos, y rasgos adicionales que complacen al consumidor a través de la novedad por ejemplo, pigmentación, mejor sabor, menor toxicidad o usos adicionales (Gepts, 2014; Gepts y Papa, 2003; Meyer y Purugganan, 2013; Zeder, 2006; Zeder et al., 2006).

El proceso de domesticación consta de cuatro etapas: 1) manejo, donde existe una administración *in situ* de los recursos por parte de los humanos, y se lleva a cabo una recolección de poblaciones de plantas silvestres antes de un cultivo formal; 2) cultivo intencional de las plantas silvestres por parte de los humanos; 3) expansión geográfica de las poblaciones domesticadas, implicando a menudo la adaptación a nuevos entornos, y 4) reproducción deliberada de los individuos ya domesticados (Meyer y Purugganan, 2013; Gaut et al., 2018).

Durante las primeras dos etapas, la relación entre el humano y los organismos en proceso de domesticación es débil, y podemos esperar frecuentes cruces entre los organismos seleccionados y sus formas silvestres. En teoría, lo anterior permite a las poblaciones superar el cuello de botella de domesticación y simultáneamente aumenta la frecuencia de alelos de domesticación. En las últimas

dos etapas, el control del humano sobre de los eventos reproductivos de los organismos seleccionados va aumentando, por lo que aumenta la dependencia de los cultivos (Meyer y Purugganan, 2013; Gaut et al., 2018).

Aunque los detalles de cómo es que se seleccionan las características deseadas durante la domesticación son poco conocidas (Sang y Ge, 2007; Sweeney y McCouch, 2007), se ha observado que en muchos cultivos existe una menor variación genética que en su ancestro silvestre (Gerbault et al., 2014), esto debido a que por lo general la domesticación inicia con unos pocos individuos de la especie (Gepts, 2014; Gerbault et al., 2014). Ejemplo de esto lo podemos ver en cultivos de gran importancia a nivel mundial como la cebada (Morrell et al., 2014), el trigo (Huang et al., 2002; Luo et al., 2007; Würschum et al., 2013; Avni et al., 2017), el arroz (Song et al., 2003; Gao, 2004; Garris et al., 2005; Song et al., 2005), la soya (Lam et al., 2010), las lentejas (Alo et al., 2011), o líneas endógamas Europeas de maíz (Frascaroli et al., 2013) (Tabla 1). Sin embargo, esto no siempre es la regla a seguir.

En cultivos Mesoamericanos se ha observado que la diversidad genética de las especies cultivadas en sistemas de cultivo tradicionales presentan una diversidad genética similar a la del pariente silvestre. Ejemplo de esto lo podemos ver en el chile (Rodríguez et al., 1999; Votava et al., 2006; Aguillar-Meléndez et al., 2009), el jitomate (Labate et al., 2009 ; Bauchet y Causse, 2012), el frijol (Zizumbo-Villarreal et al., 2005; Mamidi et al., 2011; Blair et al., 2013; Müller et al., 2015; M. Rodríguez et al., 2016;), el maíz (Hufford et al., 2013), o el epazote (Blanckaert et al., 2012) (Tabla 1).

A medida que la ciencia ha avanzado, los enfoques genómicos se han vuelto especialmente importantes en el estudio de la domesticación (Verde et al., 2013; Aguirre-Liguori et al., 2017; Guerra-García et al., 2017; Fustier et al., 2017), ya que permiten la secuenciación de múltiples genotipos a un costo relativamente bajo, ampliando así los horizontes de investigación, proporcionando una visión más amplia sobre los efectos de la domesticación en la diversidad genética (Gepts, 2014).

Tabla 1. Comparación de la diversidad genética reportada para diferentes cultivos y sus variedades silvestres.

Nombre común	Nombre científico	Estimador diversidad	Silvestres	Cultivados	Autor
Cebada	<i>Hordeum vulgare</i>	θ_w	0.0069	0.0056	Morrell et al., 2014
Trigo	<i>Triticum turgidum</i>	H_E	0.161	0.131	Luo et al., 2007
		π	0.0011	0.0013	Avni et al., 2017
Arroz	<i>Oryza sativa</i>	H_E	-	0.7	Garris et al., 2005
		H_E	0.48	-	Song et al., 2003
		H_E	0.64	-	Gao 2004
		H_E	0.413	-	Song et al., 2005
Soya	<i>Glycine max</i>	θ_w	0.00296	0.00189	Lam et al., 2010
Lentejas	<i>Lens culinaris</i>	θ_w	0.00951	0.00313	Alo et al., 2011
Chile	<i>Capsicum annuum</i>	θ_w	0.0028	0.00255	Aguilar-Meléndez et al., 2009
Jitomate	<i>Solanum lycopersicum</i>	θ_w	0.0016	0.0017	Labate et al., 2009
		H_E	0.240	0.26	Zizumbo-Villarreal et al., 2005
Frijol	<i>Phaseolus vulgaris</i>	θ_w	0.0086	0.0084	Mamidi et al., 2011
		H_E	0.260	0.157	Rodriguez et al., 2016
		H_E	0.088	0.168	Müller et al., 2015
Maíz	<i>Zea mays</i>	H_E	0.196	0.292	Hufford et al., 2013

1.1 Métodos de secuenciación masiva

Uno de los principales objetivos de la biología es ligar la variación genotípica con la variación fenotípica. Esto ha sido principalmente importante en los estudios con plantas domesticadas, ya que nos permite encontrar cuales son los genes que están relacionados a las características que se están seleccionando.

Desde el descubrimiento del DNA, se ha innovado en tecnologías que permitan obtener conjuntos de datos más grandes a bajos costos para responder a las complejas preguntas biológicas, surgiendo así en la década de los 2000 la secuenciación de nueva generación (NGS), o métodos de secuenciación paralela masiva (Goodwin et al., 2016). Estas técnicas tienen como ventaja que generan datos de secuencias que no requieren datos previos de los marcadores específicos (como puede ser el diseño de primers), lo que reduce el sesgo de verificación inherente en muchos otros enfoques de genotipificación basados en PCR o hibridación que están diseñados para detectar un conjunto predefinido de marcadores (Mardis, 2011; Egan et al., 2012; Ott et al., 2017).

De esta manera, actualmente existen diferentes plataformas de secuenciación las cuales involucran diferentes estrategias para preparar las

bibliotecas para la secuenciación, así como para detectar la señal y finalmente para leer la secuencias de DNA (Buermans y den Dunnen, 2014). De las más usadas son aquellas que requieren una amplificación clonal local de las moléculas de DNA a secuenciar, esto para aumentar la señal de cada base, ya que los sistemas no son lo suficientemente sensibles como para detectar la extensión de una base en el nivel individual (Mitra y Church, 1999; Buermans y den Dunnen, 2014).

Una de estas plataformas de secuenciación es Illumina, anteriormente conocida como Solexa, la cual utiliza nucleótidos marcados con moléculas fluorescentes que se agrupan en esferas de DNA, las cuales se leen a través de una celda de flujo. Durante la secuenciación, en las esferas de DNA en la celda de flujo se lee un nucleótido a la vez en ciclos repetitivos. En circunstancias ideales, todas las bases dentro de un grupo se extenderán en fase. Sin embargo, una pequeña parte de las moléculas no se extienden adecuadamente, y a lo largo de muchos ciclos, estos errores se acumulan, lo que provoca una disminución de la calidad hacia los extremos de las lecturas (Bentley et al., 2008).

Dependiendo de la plataforma Illumina que se utilice, se puede obtener hasta 600 Gb de lecturas de 100 pb en 12 días (HiSeq2000/2500), o 15 Gb de lecturas de 300 pb en 65 horas (MiSeq) (Buermans y den Dunnen, 2014).

Otra plataforma de secuenciación muy usada es Ion Torrent, la cual se basa en la amplificación mediante PCR de emulsión de perlas de DNA con nucleótidos sin ningún marcaje. La detección se realiza a través del ion H⁺ generado durante la incorporación de nucleótidos, utilizando sensores de pH en cada pozo de la placa de secuenciación. En contraste con el método de Illumina, los dNTP utilizados no están bloqueados y cuando en la placa se tiene un tramo de homopolímero, se extenderá todo el tramo de bases idénticas, lo que provocará un cambio de pH más fuerte. Debido a que no están terminados, las corridas de homopolímeros dan como resultado múltiples incorporaciones, que no tienen una señal estrictamente lineal, por lo que esta plataforma de secuenciación presenta un problema sistemático con los homopolímeros. Sin embargo, con esta plataforma se pueden obtener de 1.5 a 5 Gb de lecturas de 200-400 pb en solo 4 horas (Buermans y den Dunnen, 2014).

1.2 Preparación de librerías

La forma más completa de genotipado es la resecuenciación completa de los genomas de individuos de interés a una profundidad suficiente para identificar polimorfismos. Sin embargo, para muchas especies eucariotas, este enfoque todavía tiene un costo muy elevado dado al tamaño de su genoma. Por ello, se han desarrollado varias estrategias de reducción del genoma para dirigir la secuenciación a solo un subconjunto del genoma de los organismos, lo que reduce la cantidad total de datos de secuencia necesarios por individuo (López de Heredia, 2016).

Una de las primeras técnicas de reducción de genoma y de las más populares es la secuenciación de marcadores de DNA asociados a sitios de restricción (RADseq) (López de Heredia, 2016). En esta técnica la reducción del genoma se consigue mediante el uso de endonucleasas de restricción, que son enzimas que reconocen un patrón específico de bases en las secuencias de DNA y fragmentan la doble hélice en sitios específicos de cuatro a ocho nucleótidos, denominados dianas o blancos de restricción. Las lecturas van a cubrir los dos lados de la diana de restricción, y pueden ser simples o pareadas (Davey y Blaxter, 2010). Las dianas de restricción de la secuencia de DNA que se buscan con RADseq son poco frecuentes, de manera que el número de fragmentos que se producen sea tratable para una plataforma de secuenciación y no se genere un número elevado de datos faltantes por locus e individuo (Ott et al., 2017).

Hoy en día existen numerosas modificaciones de la técnica RADseq original, que permiten optimizar la obtención de marcadores genómicos a costos razonables, aumentando el número de individuos que se pueden analizar, mientras se reduce el volumen de datos faltantes o se minimizan las tasas de error (Beissinger et al., 2013; López de Heredia, 2016).

De las más usadas es la técnica de genotipado simple por secuenciación (GBS), la cual se desarrolló para estudios de mapeo genético en maíz (Elshire et al., 2011). Es muy similar a la técnica original, pero modula la composición del

barcode utilizando nucleótidos degenerados para minimizar el error en la identificación de los loci secuenciados. La técnica de GBS consiste en la utilización de endonucleasas de restricción resistentes a metilación, de manera que se evitan regiones repetitivas del genoma, aumentando la profundidad de secuenciación de regiones con un bajo número de copias (Elshire et al., 2011; López de Heredia, 2016).

Sin embargo, la técnica convencional de GBS también presenta altas tasas de datos faltantes y errores de genotipado, particularmente en sitios heterocigotos, por lo que surgió una modificación a esta denominada tGBS (Ott et al., 2017). En este método se emplean dos enzimas de restricción para generar salientes en orientaciones opuestas a las que se ligan los oligos (de una sola hebra) en lugar de los adaptadores (de doble hebra). De tal forma que solo los fragmentos de doble digestión son amplificados y secuenciados. Además, el uso de oligos evita la necesidad de preparar adaptadores y los problemas asociados con el ligamiento entre los mismos adaptadores (Ott et al., 2017). En la preparación de las librerías se emplean cebadores selectivos, donde se pueden usar de 1 a 4 sitios degenerados durante la etapa de PCR. Esto permite una reducción adicional del genoma, por lo que hay más lecturas por sitio disponibles para el llamado de SNPs. Por lo tanto, en comparación con GBS, la técnica de tGBS ofrece una mayor precisión en los llamados de SNP (> 97-99%), incluso en sitios heterocigotos, una menor tasa de datos faltantes por marcador en una población de muestras y una mejor detección de alelos raros (Ott et al., 2017).

1.3 Relevancia del proyecto

Los datos genéticos han demostrado ser un recurso poderoso para extraer inferencias sobre cuestiones relacionadas con el origen geográfico de los cultivos, el número de eventos de domesticación independientes para una determinada especie de cultivo y hace cuánto tiempo ocurrieron, los cambios moleculares específicos que subyacen a los rasgos de domesticación y la naturaleza de la

selección artificial durante la domesticación y el cultivo posterior (Gross y Olsen, 2010).

De los cultivos domesticados con mayor importancia en América se encuentran las calabazas (*Cucurbita* spp.), cuya domesticación inició hace unos 10,000 años, convirtiéndola en una de las especies domesticadas más antiguas de la historia de la humanidad (Smith, 1997). Dentro del género, *C. pepo* es la calabaza económicamente más importante a nivel mundial, ya que es la especie con la mayor cantidad de variedades comerciales y la que puede ser cultivada en una mayor cantidad de ambientes (Nee, 1990; Decker-Walters et al., 1993; Paris, 2000).

Además, *C. pepo* fue una de las primeras especies cultivadas en México, y desde entonces se ha cultivado y manejado en casi todas sus regiones agrícolas bajo el sistema tradicional de cultivo conocido como milpa, lo que trajo como consecuencia el desarrollo de diversidad genética almacenada en las variedades locales (Sánchez de la Vega, 2017).

Recientemente, la milpa se ha ido desplazando por sistemas de cultivo más tecnificados, donde solo se permite crecimiento de las variedades de calabazas comerciales, y se erradican mediante incineración todos aquellos frutos que no presenten el fenotipo comercial deseado por el agricultor. Esto podría traer consigo una disminución en la diversidad genética de las variedades locales, haciéndolas susceptibles a la extinción (Kirkpatrick y Wilson, 1988; Quesada et al., 1996; Decker-Walters et al., 2002; Cohen et al., 2003; Arriaga et al., 2006; Laughlin et al., 2009; Formisano et al., 2010).

A pesar de la gran diversidad que tienen los cultivos tradicionales de calabaza en México, en la mayoría de los estudios previos donde se analiza la diversidad genética de *C. pepo* están pobremente representadas las poblaciones mexicanas, por lo que este estudio es el primero en analizar la diversidad a nivel genómico de las poblaciones criollas de este importante recurso fitogenético. Siendo esto el primer paso para conocer los genes asociados a las adaptaciones locales a las condiciones ecológicas y climáticas locales de donde crecen, los cuales permitirían

mejorar los cultivos de calabazas para adaptarlos a las condiciones ambientales generadas por el cambio climático, mejorando así la productividad futura.

2. Especie de estudio

2.1 Importancia del género

El género *Cucurbita* es nativo de América y es conocido por sus especies cultivadas, las cuales en los países de habla hispana son comúnmente conocidas con nombres como "calabazas" o "zapallos", o mediante numerosos nombres en lenguas indígenas, mientras que en los países anglosajones se les denomina "squashes", "pumpkins" o "gourds" (Lira-Saade, 1995; OCDE, 2012; Paris, 2018). Se caracteriza por tener plantas monoicas, rastreras, trepadoras o subarborescentes en algunas formas cultivadas (Lira-Saade, 1995).

2.2 Generalidades taxonómicas.

Las especies silvestres y cultivadas de *Cucurbita* son plantas monoicas, rastreras, trepadoras o subarborescentes en algunas formas cultivadas. Sus flores son gamopétalas con corolas tubular-campanuladas, muy vistosas, de color amarillo pálido a amarillo-anaranjado brillante (Lira-Saade, 1995).

Las flores de las especies de *Cucurbita* abren muy temprano por la mañana y son polinizadas por especies de abejas solitarias de los géneros *Peponapis* y *Xenoglossa* (Hurd y Linsley, 1964, 1966, 1967; Hurd et al., 1971), aunque también son polinizadas por otras especies, en particular la abeja europea, *Apis mellifera*.

Los frutos son del tipo pepo, nombre aplicado a frutos grandes, redondos y comestible de las cucurbitáceas (Paris, 2018), y en las plantas cultivadas se producen en una gran diversidad de formas, tamaños, colores, tipos de superficies etc., mientras que en las plantas silvestres son relativamente uniformes en cuanto a su forma (globosos, ovoides o raramente piriformes), superficie (generalmente lisa o sin ornamentaciones), coloración (blancos, amarillentos o verdes con o sin manchas y/o franjas) y de tamaño comparativamente pequeño. Actualmente se reconoce que el género *Cucurbita* cuenta con un total de 20 taxa, de los cuales *C. pepo*, *C. ficifolia*, *C. moschata*, *C. maxima* y *C. argyrosperma* cuentan con variedades cultivadas (Lira-Saade, 1995).

2.3 *Cucurbita pepo*

Cucurbita pepo es tal vez la especie más estudiada del género; cuenta con tres subespecies ampliamente reconocidas basadas en datos taxonómicos y genéticos (Wilson et al., 1992; Lira-Saade, 1995; Gong et al., 2012). *Cucurbita pepo* subsp. *fraterna* que no tiene parientes domesticados conocidos – aunque diversos autores han sugerido que podría tratarse del pariente silvestre de *C. pepo* subsp. *pepo* (Andres, 1987; Castellanos-Morales et al., 2018, 2019); *C. pepo* subsp. *ovifera*, también conocida como subespecie *texana*, tiene dos variedades silvestres (*C. pepo* subsp. *ovifera* var. *texana* y *C. pepo* subsp. *ovifera* var. *ozarkana*) y una variedad cultivada (*C. pepo* subsp. *ovifera* var. *ovifera*), la cual se ha derivado de *C. pepo* subsp. *ovifera* var. *ozarkana* (Nesom, 2011); y *C. pepo* subsp. *pepo*, para la cual no se conoce un ancestro silvestre directo, pero abarca la mayor parte del germoplasma cultivado (Lira-Saade, 1995; Paris et al., 2003; Gong et al., 2012; OCDE, 2012; Paris, 2018). Teppner (2000) propuso una cuarta subespecie, *C. pepo* subsp. *gumala*, la cual se considera muy relacionada a *C. pepo* subsp. *pepo*. *C. pepo* subsp. *gumala* son calabazas cultivadas aparentemente autóctonas de Guatemala y el sur de México adyacente. Teppner (2000) observó que los frutos de *C. pepo* subsp. *gumala* son similares a los restos de *C. pepo* de la cueva de Guilá Naquitz en Oaxaca (Smith, 1997), por lo que se considera que esta subespecie puede ser similar a la forma silvestre ancestral *C. pepo* subsp. *pepo*, ya que sus frutos son relativamente pequeños con 13 a 20 cm de diámetro, depresión-globosa, y cáscara extremadamente gruesa, maduración naranja-amarilla y pulpa naranja.

2.3.1 Características morfológicas de las variedades silvestres

Las variedades silvestres de *Cucurbita pepo* son enredaderas herbáceas, de ramas múltiples, procumbentes o trepadoras, que tienen tallos delgados en los cuales las hojas están dispuestas alternativamente (Lira-Saade, 1995; OCDE, 2012; Paris, 1989, 2016). Las láminas palmeadas, usualmente pentalobuladas tienen una longitud de 12 a 15 cm y se extienden sobre pecíolos casi tan largos como ellos

mismos, y en las coyunturas de los pecíolos con tallos (axilas foliares) surgen zarcillos, capullos de flores y primordios de raíz (Lira-Saade, 1995; OCDE, 2012).

Las flores pistiladas crecen rápidamente, llegando a la antesis más rápido que las flores estaminadas. Las corolas son de color naranja-amarillo, de 8 a 12 cm de ancho, y constan de cinco o, a veces, seis pétalos parcialmente fusionados (Lira-Saade, 1995; OCDE, 2012).

Los frutos de la variedades silvestre presentan una variación bastante modesta en forma y color, son pequeños, de 3.5 a 8.0 cm de diámetro, redondos, con exocarpio verde cuando están frescos, y usualmente con 10 franjas oscuras anchas orientadas longitudinalmente que alternan con 10 franjas de luz estrechas que se corresponden con las posiciones de las 10 vías de la vena carpular principal subyacentes. Las franjas anchas y oscuras contribuyen con los fotosintatos al desarrollo de la fruta (Bazzaz, 1979). Las cáscaras están lignificadas y la pulpa de la fruta es invariablemente delgada, fibrosa, de color claro o pálida, generalmente amarga, y se deseca después de la madurez de la semilla, aproximadamente 50 días después de la antesis (Paris y Nerson, 2003).

Dentro de los frutos, están las semillas que son pequeñas y ovaladas, de 8 a 11 mm de largo y 1 a 2 mm de grosor, que a menudo suman 200 a 300. Las frutas silvestres de calabaza pueden permanecer intactas durante bastante tiempo después de que la planta haya muerto, esencialmente convirtiéndose en bolsas de semillas, que consisten simplemente en una cáscara delgada que encierra las semillas (Paris y Nerson, 2003).

2.3.2 Características morfológicas de las variedades cultivadas

Las variedades cultivadas de *C. pepo* crecen en una gran diversidad de ambientes (Kates et al., 2017). Son de las calabazas más tolerantes a bajas temperaturas (Paris, 2018) y así presentan un amplio rango altitudinal que va de 1,000-2,800 msnm, creciendo en ambientes con climas templados a fríos y en zonas cálidas de baja altitud (Montes, 2002; Villanueva, 2007; Velasco y Juárez, 2009).

Son enredaderas herbáceas rastreras, trepadoras o algunas veces de hábito subarborescente. Presenta una raíz principal que alcanza un gran desarrollo en comparación con las raíces secundarias que la acompañan. Los tallos presentan dominancia apical, con forma cilíndrica, ásperos al contacto, angulosos y sulcados. Las hojas grandes y palmeadas de 20 a 30 cm de largo, de color verde que parten directamente del tallo a través del peciolo de manera helicoidal y alterna. El limbo presenta una cara superior suave al tacto y una cara inferior muy áspera, con pelos cortos y fuertes. El borde de la hoja está dentado y con 3 a 5 lóbulos profundos lobuladas o sectadas. El peciolo es largo, hueco y consistente, tiene pelos rígidos en la superficie, por lo que es muy áspero al tacto (Lira-Saade, 1995).

Son monoicas, con flores grandes, de color amarillo intenso y con forma acampanada, que se disponen alrededor del tallo al que se unen por un largo pedúnculo (Lira-Saade, 1995). Las flores estaminadas presentes en pedicelos delgados de 6 a 15 cm de largo; con receptáculo acompañado, ligeramente constreñida por debajo de los sépalos; corola con 5 divisiones hasta casi la mitad de su longitud total. Las flores pistiladas sobre pedicelos robustos, sulcados, 2 a 5 cm de largo; ovario de muy diversas formas (globoso, ovoide, cilíndrico o raramente piriforme o más o menos turbaniforme), con receptáculo muy reducido (Lira-Saade, 1995). Cada flor se abre una vez antes del amanecer y se cierra al medio día, periodo en el cual es visitada por las abejas (Scott, 1934; Whitaker y Davis, 1962).

Los frutos de *Cucurbita pepo* cultivados son muy diversos en tamaño, forma, tamaño y color (Paris, 2000). Las frutas pueden pesar más de 25 kg., van desde redondas hasta extremadamente largas y planas, y tienen costillas longitudinales, surcos, ranuras o crestas. Los colores de las frutas incluyen verde, naranja y amarillo, que pueden variar en intensidad de vívido a pálido y en la oscuridad de casi negro a casi blanco. Los patrones de color incluyen rayas longitudinales en un rango de amplitud y contigüidad, y moteado, así como el patrón bicolor latitudinal, de manera que una sola fruta puede tener cuatro colores en su superficie. El color interior puede variar desde blanco verdoso hasta amarillo intenso o naranja. Las semillas varían de 8 a 25 cm de longitud y de 1.5 a 2.5 veces más largas que anchas,

pero su color es invariablemente beige (Lira-Saade, 1995; Paris, 2001; Paris et al., 2012).

2.4 Historia de domesticación de *Cucurbita*

2.4.1 Descubrimientos arqueológicos

La domesticación de plantas en Mesoamérica se desarrolló en un agroecosistema de cultivo muy característico conocido como *milpa* (Lozada-Aranda et al., 2017). Este sistema se caracteriza por ser un policultivo, donde coexisten diferentes plantas, principalmente maíz (*Zea mays*), frijol (*Phaseolus* spp.) y calabaza (*Cucurbita* spp.) y otras como tomate (*Solanum lycopersicum*), chile (*Capsicum annum*), quelites, etc. Las calabazas, en particular *C. pepo*, fueron unas de las primeras especies que se domesticaron en Mesoamérica, con una historia de domesticación de al menos 10,000 años en el sur de México (Whitaker y Cutler, 1971; Lira-Saade y Montes-Hernandez, 1994; Smith, 1997), 12,000 años para Ecuador (Piperno y Stothert, 2003) y 9,240 a 5,500 años en Perú (Dillehay et al., 2007), y al menos cinco especies diferentes de *Cucurbita* domesticadas, ya que era una fuente importante de alimento y con las que aparentemente existía un gran comercio (Nee, 1990; Sauer, 1950).

Dentro de las especies de calabazas domesticadas destaca *C. pepo*, que presenta la más amplia gama de hábitats de crecimiento, además de ser el taxón más cultivado (Nee, 1990) y la especie de calabaza con mayor importancia económica y más cultivada a nivel mundial (Decker-Walters et al., 1993; Hernández-Bermejo, 1994; OCDE, 2012).

Existen cuatro principales grupos de restos arqueológicos que nos dan pistas sobre la historia de domesticación de *C. pepo*. El primero y con mayor importancia en México para estudios de domesticación es la cueva de Guilá Naquitz en el Valle de Oaxaca, donde se encontraron los restos fósiles más antiguos de calabazas (Flannery, 1986). La mayoría de estos restos datan de 8,400 A.P. (Whitaker y Cutler, 1971), y han sido identificación como pertenecientes a *C. pepo* con base en una serie de características morfológicas distintivas, como el tamaño de la semilla

(Whitaker y Cutler, 1971; 1986; Smith, 1997; Zeder et al., 2006). Estos hallazgos indican que la especie *C. pepo* pudo haber sido domesticada por los primeros habitantes de Guilá Naquitz hace aproximadamente 10,000 años, quienes cosechaban tanto las calabazas silvestres del género *Cucurbita* como la especie domesticada *C. pepo* (Smith, 1997; Zeder et al., 2006).

El segundo grupo de restos arqueológicos son del valle de Tehuacán, Puebla, donde en la cueva de Coxcatlán, zona XIV, se encontró un solo espécimen de semilla de *C. pepo* que data de la fase de El Riego de aproximadamente 5,000 años de antigüedad, probablemente proveniente de una calabaza silvestre cercana al sitio (Byers, 1967). La semilla de la cueva de Coxcatlán es ligeramente más pequeña que cualquiera de los especímenes de Ocampo (Byers, 1967).

El tercer grupo de restos es de hace alrededor de 6,000-5,200 años en el estado de Tamaulipas, al noreste de México en las cuevas de las regiones de Infiernillo y Ocampo, donde se han encontrado fragmentos de semillas, pedúnculos y corteza de *C. pepo* (Hanselka, 2017). Dichos restos se localizaron a menos de 80 km de las poblaciones actuales de la variedad silvestre y posible ancestro *C. pepo* subsp. *fraterna* (Smith, 1997).

El cuarto grupo de hallazgos se encontró en Phillips Spring Site, cerca del río Terre Pomme en el centro oeste de Missouri, EUA (Kay et al., 1980), donde se han descubierto restos de semillas y pedúnculos de *C. pepo* que datan de hace aproximadamente 4,300 años A.P. (Kay et al., 1980; King, 1985; Decker, 1988; Decker-Walters et al., 1990; Wilson et al., 1992; Lira-Saade y Montes-Hernandez, 1994; Paris et al., 2003; Paris et al., 2003; Sanjur et al., 2002; Zeder et al., 2006; Gong et al., 2012)

Dada la distancia geográfica y temporal de estas cuevas y sus restos fósiles, además de la amplia distribución geográfica que tiene *C. pepo*, se ha propuesto que fue domesticada más de una vez por diferentes pueblos indígenas que habitaban en zonas geográficas muy dispersas, y con diferentes propósitos (Whitaker y Carter, 1946).

Las evidencias arqueológicas y biológicas indica dos eventos de domesticación independientes para esta especie (Decker, 1988; Wilson et al., 1992; Decker et al., 1993; Ferriol et al., 2003; Paris et al., 2003; Sanjur et al., 2002; Gong et al., 2012; Castellanos-Morales et al., 2018). Una domesticación en el sureste de Estados Unidos corresponde a *C. pepo* subsp. *ovifera* var. *ovifera*, cuyo ancestro silvestre fue *C. pepo* subsp. *ovifera* var. *texana* (Smith, 2006). La otra domesticación de *C. pepo* subsp. *pepo* sucedió en México, tal vez en el sureste, cuyo ancestro silvestre se desconoce (Smith, 1997; Sanjur et al., 2002), aunque algunos autores han sugerido que el pariente silvestre de *C. pepo* subsp. *pepo* es *C. pepo* subsp. *fraterna* (Andres, 1987; Castellanos-Morales et al., 2018, 2019). Análisis genéticos basados en secuencias de cloroplasto y en loci de microsatélites nucleares recientes sugieren que *C. pepo* subsp. *fraterna* es el probable ancestro silvestre de *C. pepo* subsp. *pepo*, con una subsiguiente y constante hibridación con *C. pepo* subsp. *ovifera*, lo que da una señal confusa para la identificación del ancestro silvestre de *C. pepo* subsp. *pepo* (Castellanos-Morales et al., 2019).

2.4.2 Síndrome de domesticación

Los restos de calabazas hallados en la cueva de Guilá Naquitz, del más antiguo al más joven, muestran cambios graduales en las características morfológicas de éstas, como aumento constante en el grosor de la corteza, en el tamaño de la semilla, y en el diámetro del pedúnculo (Smith, 1997; Zeder et al., 2006). Estos cambios parecen indicar que los seres humanos estaban seleccionando deliberadamente ciertas características de la fruta en *C. pepo* hace de 7,000 a 10,000 años (Whitaker y Cutler, 1971; Lira-Saade y Montes-Hernandez, 1994; Smith, 1997).

Estos hallazgos han sugerido que la domesticación de las calabazas probablemente se inició porque la forma esférica y rayada de los frutos atraieron a los primeros cazadores-recolectores (Nee, 1990), los cuales en un principio consumían solo las semillas, que se limpian fácilmente del tejido placentario, donde se encuentra la cucurbitacina que da el sabor amargo que caracteriza a las

calabazas silvestres (Whitaker y Cutler, 1965; Lira-Saade y Montes-Hernandez, 1994).

Posteriormente, por encuentros ocasionales con frutos no-amargos que facilitaron su consumo, se dio la crianza de plantas con frutos no-amargos, siendo un paso importante en la domesticación de las calabazas, ya que la pulpa de los frutos maduros de las plantas silvestres, además de ser amarga, es delgada, fibrosa y se deseca rápidamente. Por lo tanto, la carne de la fruta madura probablemente se convirtió en una fuente de alimento sólo después de pasos adicionales en el proceso de domesticación (Paris, 2018).

2.5 Trabajos previos de diversidad y estructura genética en *C. pepo*

El primer estudio genético de *Cucurbita pepo* se publicó en 1984, cuando Ignart y Weeden (1984) realizaron un análisis de isoenzimas para evaluar la diversidad existente entre cinco variedades comerciales de *C. pepo* (*Zuccchini*, *Pumpkin*, *Spaghetti*, *Acorn*, *Scallop* y *Straightneck*), procedentes de 21 cultivares diferentes en Ginebra, Nueva York, EUA. Estos autores encontraron que no existía mucha variación genética dentro de estas variedades comerciales, ni mucha diferenciación entre las mismas. De ahí en adelante, los análisis con isoenzimas con *C. pepo* fueron extremadamente útiles para comprender sus relaciones filogenéticas y taxonomía, gracias a estos estudios se reveló la existencia de al menos dos eventos de domesticación independiente y la diferenciación genética entre las subespecies cultivadas (Decker y Wilson, 1987; Decker, 1988; Decker-Walter et al., 1993).

Numerosos estudios posteriores emplearon diferentes marcadores genéticos (como RFLPs, AFLPs, ISSRs, RAPDs, ESTs) para caracterizar la diversidad y estructura genética de *C. pepo* en diferentes cultivares (Ruiz y Hemleben, 1991; Decker-Walters et al., 2002; Katzir et al., 2002; González et al., 2010; Formisano et al., 2012; Gong et al., 2012; Barzegar et al., 2013; Ntuli et al., 2015). Así, se analizó la estructura genética de los diferentes morfotipos del fruto descritos para las variedades comerciales (Ferriol et al., 2003; Paris et al., 2004), se evaluó la introgresión de genes de resistencia a patógenos (Formisano et al., 2010), y se llegó

a la construcción y comparación de los primeros mapas genéticos para la especie (Zraidi et al., 2007).

Gong y colaboradores (2008a) desarrollaron una serie de marcadores tipo microsatélites (o SSRs) para *C. pepo* y *C. moschata*, con los cuales Gong et al. (2008b) crearon un mapa genético para las dos especies de *Cucurbita*, el cual al compararlo solo encontraron cuatro de 76 marcadores SSR en grupos cromosómicos no homólogos, indicando que la mayoría de los marcadores han conservado órdenes y distancias genéticas similares, representando loci ortólogos. Más recientemente, los microsatélites se han empleado para recabar evidencias genéticas que permitan la detección de variedades comerciales y locales de calabazas en el mercado agroalimentario (Verdone et al., 2018).

Asimismo, se han desarrollado trabajos de evaluación de la diversidad y la estructura genética en variedades de cultivo más tradicional de Mesoamérica. Enríquez et al. (2018) evaluaron el efecto de la fragmentación de las áreas conservadas y el manejo agrícola sobre la diversidad y el flujo de genes de *C. pepo* en áreas del centro de Guatemala, encontrando que las prácticas de manejo tienen un efecto muy fuerte sobre la diversidad y estructura genética de las poblaciones cultivadas, además, encontraron evidencia de cruza de variedades tradicionales locales con variedades comerciales, lo que podría poner en riesgo la diversidad e identidad genética local de *C. pepo*.

Ruiz Mondragón (2017) realizó un estudio para conocer la historia evolutiva de *C. pepo* por medio de un enfoque filogeográfico, con base en dos regiones de cloroplasto y una región de la mitocondria para las tres subespecies. Para *C. pepo* subsp. *pepo* se analizaron un total de 25 poblaciones provenientes de cultivos tradicionales en 13 diferentes estados de México, así como de cinco variedades comerciales y 11 secuencias arqueológicas de las tres subespecies. Los resultados de este análisis sugieren que el linaje de *C. pepo* mexicano (subsp. *pepo* y subsp. *fraterna*) divergió del linaje de los EUA (subsp. *texana*) en el Pleistoceno, hace 1.54 MA. También reportó que *C. pepo* subsp. *fraterna*, es el pariente silvestre más cercano a *C. pepo* subsp. *pepo* de México, señalándolo como su posible ancestro,

y que la mayor diversidad y diferenciación genética está en el centro-occidente de México, por lo que este podría ser el centro de diversidad actual de la especie cultivada *C. pepo* subsp. *pepo*. También se han realizado estudios de relaciones filogenéticas y tiempos de divergencia para todo el género *Cucurbita*. Paredes-Torres (2016) evaluó estas relaciones mediante seis fragmentos de cloroplasto de 20 de los 22 taxa existentes para el género, encontrando que dentro del grupo *Pepo* se diferenciaron las subespecies mexicanas (*C. pepo* subsp. *pepo* y *C. pepo* subsp. *fraterna*) de los taxa de Estados Unidos (*C. pepo* subsp. *ovifera* var. *texana* y *C. pepo* subsp. *ovifera* var. *ozarkana*) hace 100 000 años.

Kates y colaboradores (2017) obtuvieron una filogenia molecular basada en 44 loci derivados de intrones de genes nucleares de una sola copia, encontrando que las tres subespecies de *C. pepo* forman un clado hermano de *C. okeechobeensis* y *C. lundelliana*, pero no lograron resolver las relaciones filogenéticas entre las tres subespecies de *C. pepo*. Castellanos-Morales y colaboradores (2018), retomando y ampliando los datos de Paredes-Torres (2016), analizaron los patrones temporales, espaciales y ecológicos de la diversificación del género *Cucurbita*, siendo los primeros en incluir a la especie silvestre *C. radicans* en un análisis filogenético. Ellos también encontraron que *C. pepo* y sus subespecies forman un claro grupo dentro del género, que se separó del resto de las calabazas hace 1.33 MA, formando dos clados, uno para las subespecies que se distribuyen en México (*C. pepo* subsp. *fraterna* y *C. pepo* subsp. *pepo*), y otro para las subespecies y variedades que se distribuyen en los Estados Unidos (*C. pepo* subsp. *ovifera*, *C. pepo* subsp. *ovifera* var. *ozarkana* y *C. pepo* subsp. *ovifera* var. *texana*).

El trabajo más reciente con *C. pepo* es el de Castellanos-Morales y colaboradores, (2019), quienes mediante la implementación de dos regiones de cloroplasto y nueve loci microsatélites nucleares, dilucidan la historia de domesticación de *C. pepo* en México. Ellos encuentran que *C. pepo* subsp. *pepo* presenta alta variación y baja diferenciación genética, y que las tres subespecies están bien diferenciadas para los loci de microsatélites. También encuentran que *C. pepo* subsp. *fraterna* es probablemente el ancestro silvestre de *C. pepo* subsp.

pepo, pero debido a una posterior hibridación entre los diferentes taxones de *C. pepo*, se complica su definición como el ancestro silvestre de *C. pepo* subsp. *pepo*.

Con el avance en las técnicas de secuenciación, se ha logrado conocer diferentes características genómicas de *C. pepo*, las cuales nos sirven como herramientas básicas para lograr un mejor entendimiento de su historia evolutiva y de domesticación. Alverson et al. (2010) secuenciaron el genoma mitocondrial de *C. pepo* (Zucchini: 982,833 nt), donde el gran tamaño del genoma mitocondrial reflejó una considerable acumulación de regiones del genoma de cloroplasto y de secuencias cortas repetidas.

Esteras et al. (2013) realizaron la primera construcción de un mapa genético para *C. pepo* basado en SNPs mapeando loci de rasgos cuantitativos (QTL). Para ello emplearon una población derivada del cruce de dos variedades con fenotipos contrastantes, que representan los principales grupos de cultivares de las dos subespecies: Zucchini (*C. pepo* subsp. *pepo*) × Scallop (*C. pepo* subsp. *ovifera*). Más recientemente, Montero-Pau et al. (2018) generaron el ensamble del genoma de *C. pepo* subsp. *pepo* var. *zucchini*, con un tamaño de 263 Mb, estimando una cobertura del 93% del genoma, y organizándolo en 20 pseudocromosomas que representan el 81.4% del ensamble.

3. Objetivos

3.1 Objetivo general

- Caracterizar los niveles de diversidad, estructura genética y flujo génico en loci neutrales de las diferentes variedades de *Cucurbita pepo* y cultivos de *C. pepo* subsp. *pepo* en México utilizando una aproximación genómica.

3.2 Objetivos particulares

- Identificar loci neutrales en cultivos mexicanos de *C. pepo* subsp. *pepo*.
- Caracterizar la diversidad genética de los cultivos mexicanos de *C. pepo* subsp. *pepo* para loci neutrales.
- Caracterizar la estructura genética de los cultivos mexicanos de *C. pepo* subsp. *pepo* para loci neutrales.
- Estimar el flujo génico entre cultivos mexicanos de *C. pepo*.

4. Hipótesis

- La diversidad genómica (a nivel SNPs) en *C. pepo* será mayor que la de otros cultivos de calabazas, debido a que ha mostrado mayor diversidad con marcadores de cloroplasto y microsatélites nucleares.
- La diversidad genética evaluada con SNPs de los cultivos en México será similar a la de la *C. pepo*. subsp. *fraterna* debido a que éste es el pariente silvestre más cercano de *C. pepo*. subsp. *pepo*.
- La estructura genética estimada con SNPs no reflejará la división taxonómica tradicional entre las diferentes subespecies de *C. pepo*, ya que ha existido un constante flujo genético entre las poblaciones cultivadas y las poblaciones silvestres.
- Se encontrarán altos niveles de polimorfismos compartidos evaluados con SNPs entre las diferentes subespecies de *C. pepo*, ya que aún mantienen polimorfismos por ancestría y por contacto reproductivo reciente.

5. Método

5.1 Obtención de muestras

Para *C. pepo* subsp. *pepo* se determinaron 14 puntos de muestreo a partir de la distribución documentada por la CONABIO (2016), de tal manera que se abarcara la mayoría de su distribución en México (Figura 1). Algunas muestras provienen de donaciones del Banco de Germoplasma del Campo Experimental Bajío del INIFAP localizado en Celaya, Guanajuato (datos en la Tabla 2). Otras muestras fueron colectadas durante los años 2014 - 2015 directamente de mercados y milpas de distintas zonas del país (Tabla 2) por los miembros del proyecto *Cucurbita* de la UNAM, apoyado por la CONABIO (KE004 y PE001), del cual esta tesis forma parte. Las localidades y muestras empleadas en esta tesis son las mismas trabajadas por Ruiz Mondragón (2017) y Castellanos-Morales et al. (2019). De estas 14 localidades muestreadas se analizaron un total de 69 individuos, 3 a 8 por localidad provenientes de distintos frutos.

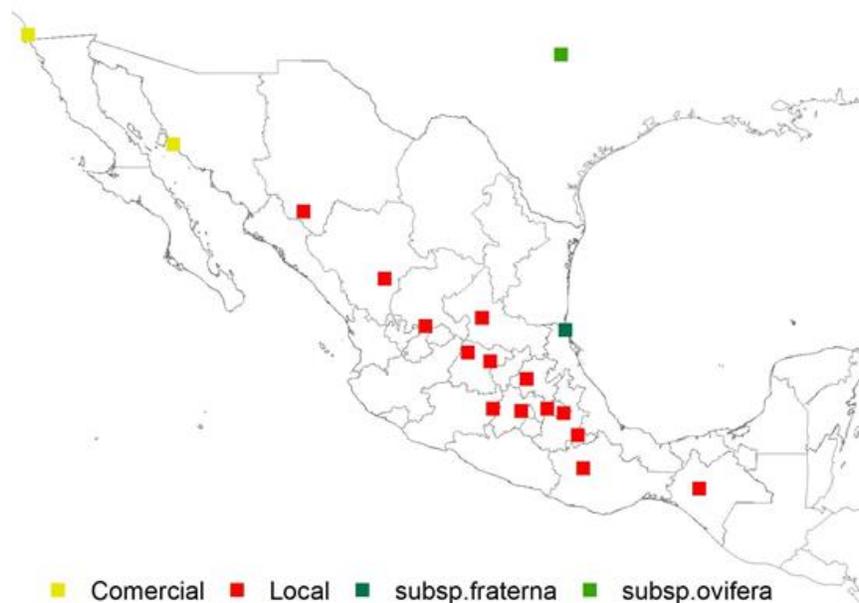


Figura 1. Mapa de distribución de las localidades donde se colectaron las diferentes muestras de *Cucurbita pepo* subsp. *pepo*, *C. pepo* subsp. *fraterna* y *C. pepo* subsp. *ovifera* var. *ovifera*.

Adicionalmente, se utilizaron ocho individuos de *C. pepo* subsp. *fraterna*, y seis individuos de *C. pepo* subsp. *ovifera* var. *ovifera*. Además, se incluyeron en los

análisis ocho individuos de cinco variedades comerciales pertenecientes a EUA de *C. pepo* subsp. *pepo*: *Black Beauty Zuchinni*, *Cocozele*, *Delicata Honey Boat*, *Straightneck* y *Spaghetti*. En total analicé 95 individuos de las distintas subespecies y variedades de *Cucurbita pepo* (Tabla 2).

Tabla 2. Localidades muestreadas con número de individuos para cada localidad y su procedencia.

Subespecie	Estado	Procedencia	No. individuos
subsp. <i>pepo</i>	Chihuahua	BG CU0879, 0880,0898	5
subsp. <i>pepo</i>	Durango	BG CU0654	4
subsp. <i>pepo</i>	Zacatecas	Colecta Nov. 2015	3
subsp. <i>pepo</i>	San Luis Potosí	Colecta Feb. 2016	2
subsp. <i>pepo</i>	Jalisco	BG CU0759	3
subsp. <i>pepo</i>	Guanajuato	BG CU0757,0568, 0766	4
subsp. <i>pepo</i>	Hidalgo	Colecta Nov. 2014	6
subsp. <i>pepo</i>	Michoacán	BG CU0972	6
subsp. <i>pepo</i>	Estado de México	BG CU1114	6
subsp. <i>pepo</i>	Tlaxcala (Ixtenco)	Colecta Nov. 2014	3
subsp. <i>pepo</i>	Tlaxcala (Vicente Guerrero)	Colecta Nov. 2014	5
subsp. <i>pepo</i>	Puebla	Colecta	8
subsp. <i>pepo</i>	Oaxaca	BG CU1052	6
subsp. <i>pepo</i>	Chiapas	Colecta Nov. 2014	8
Variedades mejoradas de <i>C. pepo</i> subsp. <i>pepo</i>			
Var. <i>Spaghetti</i>	Sonora	Colecta Nov. 2015	4
Var. <i>Cocozele</i>	California	Comerciales	3
Var. <i>Straightneck</i>	California	Comerciales	1
Var. <i>Delicata</i>	California	Comerciales	2
Var. <i>Zucchini</i>	California	Comerciales	2
Grupo externo (<i>C. pepo</i> subsp. diferentes)			
subsp. <i>fraterna</i>	Tamaulipas	BG CU0629, 0633	8
subsp. <i>ovifera</i> var. <i>ovifera</i>	México	Ornamentales	6

5.2 Extracción del DNA

Para lograr una extracción de DNA de alta calidad, se germinaron en el invernadero del Instituto de Ecología de la UNAM de 15 a 20 semillas por localidad. En las localidades donde se contaba con más de un fruto se tomaban semillas de los diferentes frutos. En el caso de las muestras donadas por el Banco de Germoplasma, cuando no se contaba con la información del número de frutos del que procedían las semillas, se sembraban de 15 a 20 semillas por sobre. A las tres semanas de germinación se colectó una muestra de tejido foliar de aproximadamente 2 cm², el cual fue utilizado para realizar la extracción de DNA, el resto de la hoja fue almacenada a -80°C como respaldo.

La extracción de DNA se realizó con el kit de extracción de QIAGEN “Plant Mini Kit” siguiendo las instrucciones del proveedor. Para evaluar la integridad del DNA se corrieron geles de agarosa al 1%, para conocer la pureza de las extracciones se utilizó nanodrop (Thermo Fisher Scientific), y para estimar la concentración de DNA por muestra se empleó Qubit (Thermo Fisher Scientific). Las extracciones que tuvieron valores de pureza de 1.8 – 2.0 para A260/A280 y concentraciones de DNA > 20 ng/μl se usaron en la preparación de librerías, las cuales se prepararon mediante el método de tGBS empleando enzimas de restricción con 2 sitios degenerados, para aumentar las lecturas por sitio disponibles en el llamado de SNPs (Ott et al., 2017). Las bibliotecas se obtuvieron utilizando la tecnología de Data2Bio con secuenciación por Ion Torrent (www.data2bio.com), dando un total de 124,783,053 reads.

5.3 Procesamiento de datos

Para revisar la calidad de las librerías se utilizó el programa FastQC (Andrews, 2010), que nos proporciona un primer acercamiento del conjunto de datos por individuo, de tal forma que podamos detectar si nuestras secuencias tienen algún problema antes de empezar a hacer cualquier análisis.

La limpieza de los datos se hizo con el programa *fastxtools* (Gordon y Hannon, 2010), reteniendo las secuencias con valores PHRED > 15 en el 80% de sus bases. Se mapearon las secuencias restantes utilizando *segemehl* 0.3.4, que es el programa recomendado para datos secuenciados con Ion Torrent (Caboche et al., 2014), utilizando a *C. pepo* subsp. *pepo* var. *zucchini* (Montero-Pau et al., 2018) como genoma de referencia, y se conservaron las secuencias con mapeo único.

El llamado de SNPs se realizó con *bcftools*, donde la calidad de mapeo mínima y la calidad PHRED mínima por sitio fuera de 20, para asegurarnos de que el SNP encontrado no sea un error de secuenciación, reteniendo únicamente los sitios polimórficos. Posteriormente, con *vcftools* se retuvieron solo los sitios con profundidad mínima de secuenciación de 5X. Para el filtrado de SNPs se empleó el

programa PLINK 1.9 (Purcell et al., 2007), eliminando los SNPs y los individuos con más del 50% de datos faltantes, y los SNPs con frecuencia mínima de alelos (MAF) < 0.02 , esto para que cada alelo estuviera representado en al menos dos individuos. Estos filtros se realizaron para asegurarnos que los SNPs encontrados sean verdaderos, y no errores de secuenciación. En total se obtuvieron 12,932 SNPs, que fueron empleados en análisis de flujo génico.

Para separar los loci neutrales de los loci con señales de selección, se empleó el programa BayeScan 2.1 (Fischer et al., 2011), ya que a diferencia de otros programas que detectan loci con señales de selección mediante una correlación con características ambientales, este nos permite detectar loci candidatos bajo selección a partir de las diferencias en las frecuencias de alelos entre poblaciones, sin la necesidad de incluir datos ambientales. Las diferencias en las frecuencias de alelos entre un conjunto común de genes y cada subpoblación se miden mediante un coeficiente F_{ST} específico de la subpoblación. Por lo tanto, este programa puede considerar escenarios ecológicos realistas en los que el tamaño efectivo y la tasa de migración pueden diferir entre las subpoblaciones (Foll, 2012).

Se usó un valor de corte de 0.05 para el q -value, resultando en 12,881 loci neutrales y 51 loci con señales de selección, de los cuales 19 parecen estar sujetos a selección purificadora y 32 a selección direccional. Con base en estos resultados, y mediante el programa vcftools, se separaron los loci con señales de selección de los loci neutrales. El juego de datos con loci neutrales se empleó para los análisis de diversidad y estructura genética.

5.4 Análisis de datos

5.4.1 Diversidad genética

Con el módulo *populations* de STACKS (Catchen et al., 2013), y usando el juego de datos con 12,881 SNPs se calcularon las medidas básicas de diversidad genética (alelos privados (A_p), heterocigosis observada (H_o), heterocigosis esperada (H_E), y

coeficiente de endogamia (F_{IS}), para las diferentes accesiones de *C. pepo* tanto con el juego de datos de loci neutrales, como los loci con señales de selección.

5.4.2 Estructura genética

Para conocer la estructura genética de *C. pepo*, y usando el juego de datos con 12,881 SNPs con la librería *hierfstat* de R (Goudet, 2005), se realizó una prueba de F_{ST} entre las diferentes subespecies. Además, para saber cómo se agrupaban genéticamente las diferentes accesiones se realizaron tres pruebas diferentes:

- 1) Análisis de componentes principales (PCA) que nos permite conocer cómo se agrupan los individuos de *C. pepo* de acuerdo con su similitud genética. Para este análisis se implementó la librería *adegenet* (Jombart et al., 2018) de R, reteniendo solo los primeros dos componentes. Para determinar el número de componentes a usar se realizó una prueba de Cross-validation usando la paquetería de R *mdatools* (Kucheryavskiy, 2017).
- 2) Red de identidad por estado (IBS), que al igual que el PCA, nos agrupa a los individuos de *C. pepo* de acuerdo con su similitud genética, pero además nos muestra los probables eventos de flujo genético entre los individuos. Para esta prueba primero calculamos las distancias genéticas pareadas para cada individuo con el programa TASSEL 5 (Bradbury et al., 2007), y la matriz obtenida se graficó en el programa SplitsTree 4 (Huson y Bryant, 2006), empleando el método de NeighborNet.
- 3) Prueba de ancestría individual por máxima verosimilitud, la cual nos muestra cuantos grupos genéticos existen en nuestro universo muestral y en qué proporción se encuentran en cada individuo. Para ello empleamos el programa ADMIXTURE 1.3 (Alexander et al., 2015), donde se probaron valores K de 2 a 12 y se realizó la prueba de Cross-validation para cada una. Para el análisis se realizó una sola corrida con los valores predeterminados por el programa, es decir un bootstrap de 200, con selección de número semilla al azar, con un criterio de finalización por corrida de detener el análisis

cuando la probabilidad de registro aumenta en menos de $\epsilon = 10^{-4}$ entre iteraciones.

5.4.3 Flujo génico

Para evaluar el flujo génico entre las diferentes accesiones locales y de las variedades *C. pepo* subsp. *fraterna* y *C. pepo* subsp. *ovifera* var. *ovifera*, se empleó el programa Treemix (Pickrell y Pritchard, 2012), el cual infiere patrones de división poblacional y mezcla a partir de datos de frecuencias de alelos genómicos, con los cuales realiza un árbol de máxima verosimilitud para el conjunto de poblaciones y prueba el número de eventos migratorios que se le indique. Se graficaron con R las probabilidades obtenidas del análisis, probando de 0 a 19 eventos migratorios, y se seleccionó el primer evento donde el logaritmo natural de la probabilidad (LN(P)) se aproxima a cero.

Debido a que sabemos que las variedades comerciales se producen mediante la hibridación de las diferentes subespecies de *C. pepo*, o incluso con otras especies de calabazas (Paris, 2000, 2016), se empleo el total de los 12,932 SNPs. Para este análisis, se excluyeron las variedades comerciales y la accesión de San Luis Potosí, la cual en el análisis de ancestría individual por máxima verosimilitud y la red IBS presentó evidencias de ser una variedad comercial.

6. Resultados

6.1 Calidad de datos

Los datos crudos enviados por el laboratorio para los 95 individuos de *C. pepo* presentaron un total de 163,842,323 reads por muestra, con un promedio de reads por muestra de 1,706,690. En cuanto a la calidad de secuenciación PHRED, el promedio por muestra fue de 22, con un mínimo de 4 y un máximo de 35.

Después del primer filtro de calidad (retención de secuencias con valores PHRED > 15 en al menos el 80% de sus pares de bases) se obtuvo un total de 149,283,580 reads, con un promedio de reads por muestra de 1,555,037.

Estos datos fueron los que se utilizaron para el llamado inicial de SNPs, en el que se encontró un total de 513,237 SNPs para *C. pepo*. De esos SNPs, el 10.17% (52,226 SNPs) presentaron una profundidad mínima de 5X. Cuando se eliminaron los sitios y los individuos con más del 50% de datos faltantes, se retuvo el 48.37% de los SNPs (25,263 SNPs) y se eliminaron 6 de los 95 individuos muestreado (Dos del Estado de México, dos de Zacatecas, uno de Guanajuato y uno de la variedad comercial *Cocozele*). Para el último filtro de calidad (MAF < 0.02) se retuvo el 51.18% de SNPs, dando un total de 12,932 SNPs en 89 individuos de *C. pepo*.

6.2 Diversidad genética

Para calcular las diferentes medidas de diversidad genética se empleó el juego de 12,881 SNPs, excluyendo los 51 SNP's que presentaron señales de selección en la prueba de BayeScan 2.1 (Fischer et al., 2011). En promedio, la heterocigosis esperada (H_E) encontrada para las 14 localidades de cultivos de *C. pepo* subsp. *pepo* en México fue de $H_E = 0.1627$, donde la heterocigosis más baja la presenta San Luis Potosí ($H_E = 0.0508$), seguida de Zacatecas ($H_E = 0.0940$), mientras que la heterocigosis más alta la presenta Michoacán ($H_E = 0.2575$), siendo también la localidad con mayor número de alelos privados (AP = 37) (Tabla 3).

El valor de H_E en *C. pepo* subsp. *pepo* fue ligeramente menor que la encontrada para la subespecie silvestre *C. pepo* subsp. *fraterna* ($H_E = 0.1735$), y mucho más pequeña que la encontrada para la variedad ornamental *C. pepo* subsp. *ovifera* var. *ovifera* ($H_E = 0.2344$) (Tabla 3). Además, si los comparamos con los valores H_E encontrados para las variedades comerciales de *C. pepo* subsp. *pepo*, observamos que todas presentan una diversidad menor que los cultivos mexicanos ($H_E = 0.0732 - 0.1226$), a excepción de la variedad *Spaghetti*, que presenta una $H_E = 0.1877$.

Tabla 3. Valores de diversidad genética para las diferentes accesiones de *C. pepo* subsp. *pepo* provenientes de 14 localidades de México, además de las accesiones para los grupos externos *C. pepo* subsp. *fraterna* y *C. pepo* subsp. *ovifera* var. *ovifera*, y las diferentes accesiones de las variedades comerciales. N = número de individuos, AP = alelos privados, H_o = heterocigosis observada, H_E = heterocigosis esperada y F_{IS} = coeficiente de endogamia.

Localidad	N	AP	H_o	H_E	F_{IS}
Chihuahua	5	4	0.1145	0.1920	0.2547*
Durango	4	6	0.1145	0.2096	0.2794
Zacatecas	2	2	0.1178	0.094	0.0219
San Luis Potosí	2	2	0.0749	0.0508	-0.0056
Jalisco	3	9	0.1301	0.2164	0.2508
Guanajuato	3	2	0.1306	0.1148	0.0333
Michoacán	6	37	0.1080	0.2575	0.4025*
Estado de México	6	22	0.1096	0.2170	0.2762
Hidalgo	4	11	0.1156	0.1092	0.0187
Tlaxcala (Ixtenco)	3	1	0.1360	0.1039	0.0021
Tlaxcala (Vicente Guerrero)	5	16	0.1240	0.1133	0.0152
Puebla	8	13	0.1205	0.2236	0.2889*
Oaxaca	6	12	0.1292	0.2009	0.2413*
Chiapas	8	6	0.1163	0.1753	0.2283
Promedio razas locales	4.6	10.21	0.1173	0.1627	0.1648
Grupo externo					
<i>C. pepo</i> subsp. <i>fraterna</i>	8	14	0.1275	0.1735	0.1812
<i>C. pepo</i> subsp. <i>ovifera</i> var. <i>ovifera</i>	5	9	0.1611	0.2344	0.2146
Variedades comerciales					
<i>C. pepo</i> subsp. <i>pepo</i> var. <i>Spaghetti</i>	4	5	0.1431	0.1877	0.1605*
<i>C. pepo</i> subsp. <i>pepo</i> var. <i>Cocozelle</i>	2	33	0.1308	0.0973	0.008
<i>C. pepo</i> subsp. <i>pepo</i> var. <i>Straightneck</i>	1	2	0.1465	0.0732	0
<i>C. pepo</i> subsp. <i>pepo</i> var. <i>Delicata</i>	2	4	0.1307	0.1089	0.0391
<i>C. pepo</i> subsp. <i>pepo</i> var. <i>Zucchini</i>	2	0	0.1463	0.1226	0.0402
Promedio accesiones comerciales	2.2	8.8	0.13948	0.11794	0.04956

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

En cuanto a los valores de coeficiente de endogamia, observamos que los cultivos de México de *C. pepo* subsp. *pepo* presentan en promedio una $F_{IS} = 0.1648$, siendo San Luis Potosí la localidad con el valor de endogamia más bajo ($F_{IS} = -0.0056$), y Michoacán la localidad con mayor endogamia ($F_{IS} = 0.4025$). El F_{IS}

promedio para *C. pepo* subsp. *pepo* fue menor a los estimados para *C. pepo* subsp. *fraterna* ($F_{IS} = 0.1812$) y *C. pepo* subsp. *ovifera* var. *ovifera* ($F_{IS} = 0.2146$).

Si comparamos los niveles de coeficiente de endogamia de *C. pepo* subsp. *pepo* de las accesiones locales contra las variedades comerciales, observamos que la variedad comercial *Spaghetti* ($F_{IS} = 0.1605$) presenta un valor de F_{IS} ligeramente más bajo que el de las accesiones locales, mientras que el resto de las variedades comerciales presentan valores de F_{IS} bajos ($F_{IS} = 0 - 0.0402$).

6.3 Estructura genética

Para los valores de F_{ST} pareadas entre las tres subespecies de *C. pepo* y sus variedades comerciales, se encontró que el valor más grande es entre las subespecies *C. pepo* subsp. *fraterna* y *C. pepo* subsp. *ovifera* ($F_{ST} = 0.0953$). El valor más pequeño fue entre las subespecies *C. pepo* subsp. *pepo* colectada en México y *C. pepo* subsp. *fraterna* ($F_{ST} = 0.0127$) (Tabla 4).

Tabla 4. Valores de F_{ST} pareadas para las tres subespecies de *C. pepo* y sus variedades comerciales. Comerciales = Variedades comerciales de *C. pepo* (*C. pepo* subsp. *pepo* var. *Spaguetti*; var. *Cocozelle*; var. *Straightneck*; var. *Delicata*; var. *Zucchini*), Locales = *C. pepo* subsp. *pepo* (colectada en México), Fraterna = *C. pepo* subsp. *fraterna*, Ovifera = *C. pepo* subsp. *ovifera*.

	Comerciales	Locales	Fraterna	Ovifera
Comerciales	0			
Locales	0.0144	0		
Fraterna	0.0650	0.0127	0	
Ovifera	0.0610	0.0157	0.0953	0

p < 0.05; *p < 0.01; ****p < 0.001

En el análisis de componentes principales (PCA), observamos que en general se forman tres grupos genéticos (Figura 2). El primer grupo se encuentra al centro de los cuadrantes, y está formado por 73 de los 89 individuos, todos de diferentes subespecies. El segundo grupo se ubica en el cuadrante superior derecho, y está formado por ocho individuos pertenecientes a *C. pepo* subsp. *pepo* de varias localidades de México (Estado de México, Puebla, Chiapas y Michoacán). El último grupo se ubica en el cuadrante inferior derecho, y está formado por los ocho individuos restantes, de los cuales uno pertenece a *C. pepo* subsp. *fraterna*, dos a *C. pepo* subsp. *ovifera* var. *ovifera*, dos a *C. pepo* subsp. *pepo* variedades

comerciales y tres a *C. pepo* subsp. *pepo*, uno de Durango, uno de Michoacán y otro de Oaxaca.

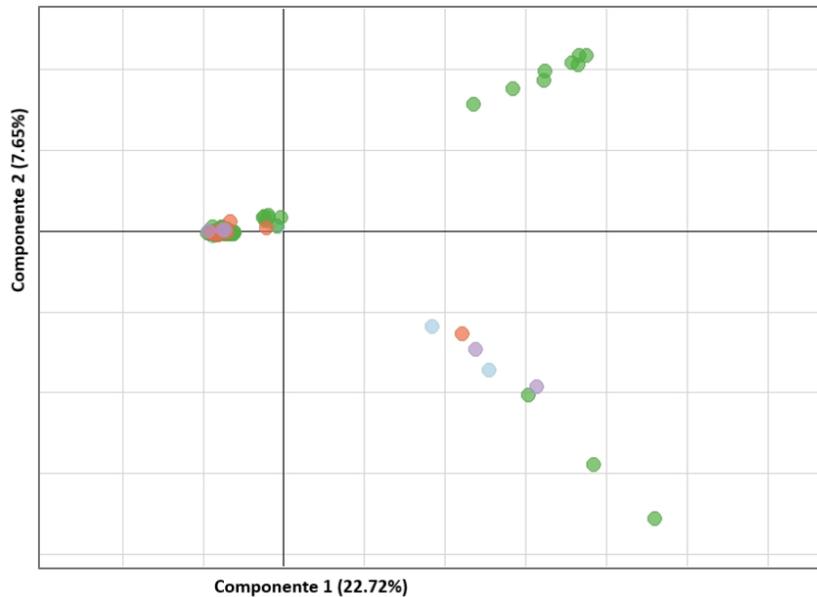


Figura 2. Gráfico del PCA. De color verde se encuentran las accesiones locales de México, de color anaranjado las accesiones de *C. pepo* subsp. *fraterna*, de color lila las accesiones de *C. pepo* subsp. *ovifera* var. *ovifera*, de color azul cielo las accesiones de las variedades comerciales. El componente 1 explican el 22.72% de la variación, y el componente 2 explica el 7.65% de la variación.

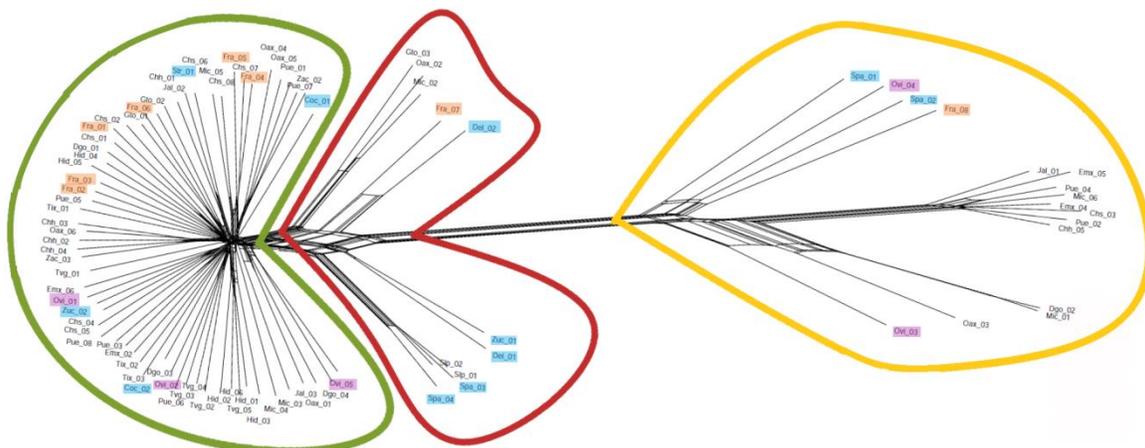


Figura 3. Análisis IBS, donde cada línea representa un individuo y las líneas que los conectan son los diferentes eventos de hibridación estimados entre los genotipos de esos individuos. El primer grupo genético está señalado de color verde, el segundo de color rojo y el tercero de color amarillo. Las accesiones de la subespecie *fraterna* se señalan con color anaranjado, las de la subespecie *ovifera* de color lila, y las accesiones comerciales de color azul.

En la red de identidad por estado (IBS), encontramos que también se forman tres grupos genéticos (Figura 3). El primero y más grande de los grupos, ubicado a la izquierda de la red señalado por el color verde, donde están 62 de los 89 individuos. El segundo y más chico de los grupos, ubicado al centro de la red señalado por el color rojo, muy cercano al primer grupo, donde se encuentran 11 de los 89 individuos. Y el tercer grupo, ubicado en el extremo derecho de la red señalado por el color amarillo, formado por 16 de los 89 individuos.

Al observar detalladamente la Figura 3, podemos notar que el primer grupo marcado en verde está formado principalmente por accesiones locales mexicanas, seis de los ocho individuos muestreados de la subespecie *fraterna*, tres de los cinco individuos muestreados de la subespecie *ovifera*, y cuatro individuos de variedades comerciales (*Zucchini*, *Cocozelle* y *Straightneck*).

El segundo grupo, marcado en rojo, está conformado por cinco individuos de accesiones locales mexicanas, un individuo de la subespecie *fraterna*, y cinco individuos de variedades comerciales (*Zucchini*, *Delicata* y *Spaghetti*). Dentro de ese mismo grupo observamos que las accesiones pertenecientes a la localidad de San Luís Potosí, se agrupan junto con la variedad comercial *Spaghetti*.

En el tercer grupo, marcado en amarillo, encontramos 11 individuos de accesiones locales mexicanas, dos individuos de la variedad comercial *Spaghetti*, dos individuos de la subespecie *ovifera* y un solo individuo de la subespecie *fraterna*.

En la prueba de ancestría individual por máxima verosimilitud, encontramos las asignaciones de los individuos grupos genéticos más probable fueron $K = 2$ ($CV = 0.42704$) y $K = 3$ ($CV = 0.4285$) (Figura 4). Debido a que en el análisis de PCA y de IBS se encontraron tres grupos genéticos, se decidió trabajar con $K = 3$.

El primer grupo genético, representado por el color verde, es el grupo más común, al estar presente en la mayoría de los individuos en al menos una quinta parte de su genotipo. El segundo grupo genético representado por el color amarillo, presente en mucha menor proporción de los individuos. El tercer grupo genético, representado por el color rojo, es el genotipo de la variedad comercial *Spaghetti* y la accesión local de San Luis Potosí, pero también lo encontramos en gran parte del

genotipo de la variedad comercial *Delicata* y en un solo individuo de la variedad comercial *Zucchini*, y en bajas proporciones del genotipo de algunos individuos de accesiones locales.

También podemos notar que existe una correspondencia entre los grupos genéticos formados en la prueba Admixture, con los formados en la prueba IBS. Los individuos que forman el primer y más grande grupo de la red IBS, son los individuos que su genotipo es mayoritariamente verde. El segundo y más pequeño grupo de la red IBS, son los individuos cuyo genotipo tiene evidencia de correspondencia con el grupo genético rojo. Dentro de este segundo grupo se encuentran las accesiones locales de Oax_02, Gto_03 y Mic_02 que, aunque su genotipo es mayoritariamente verde; también presentan cierto porcentaje de genotipo amarillo y rojo, lo que las hace más cercanas a la variedad comercial *Delicata* y a un individuo de la subespecie *fraterna*. El tercer grupo de la red IBS, está formado por individuos cuyo genotipo está más representado por el grupo genético amarillo.

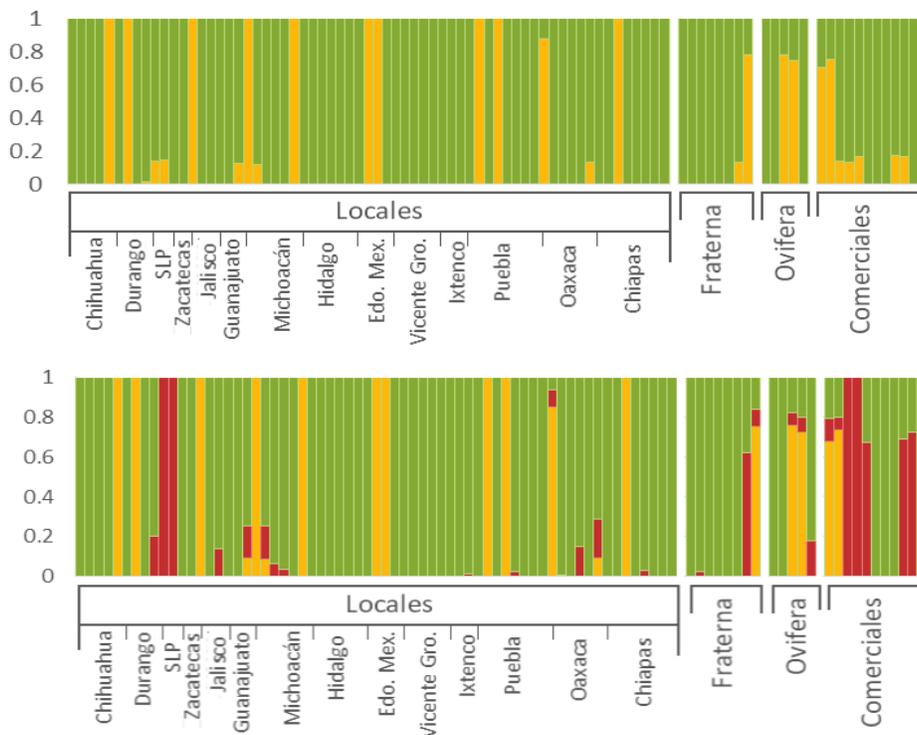


Figura 4. Histograma de $K = 2$ y $K = 3$ de análisis Admixture, donde en ninguno de los dos existe una correlación entre el origen de las accesiones y la asignación a algún grupo genético.

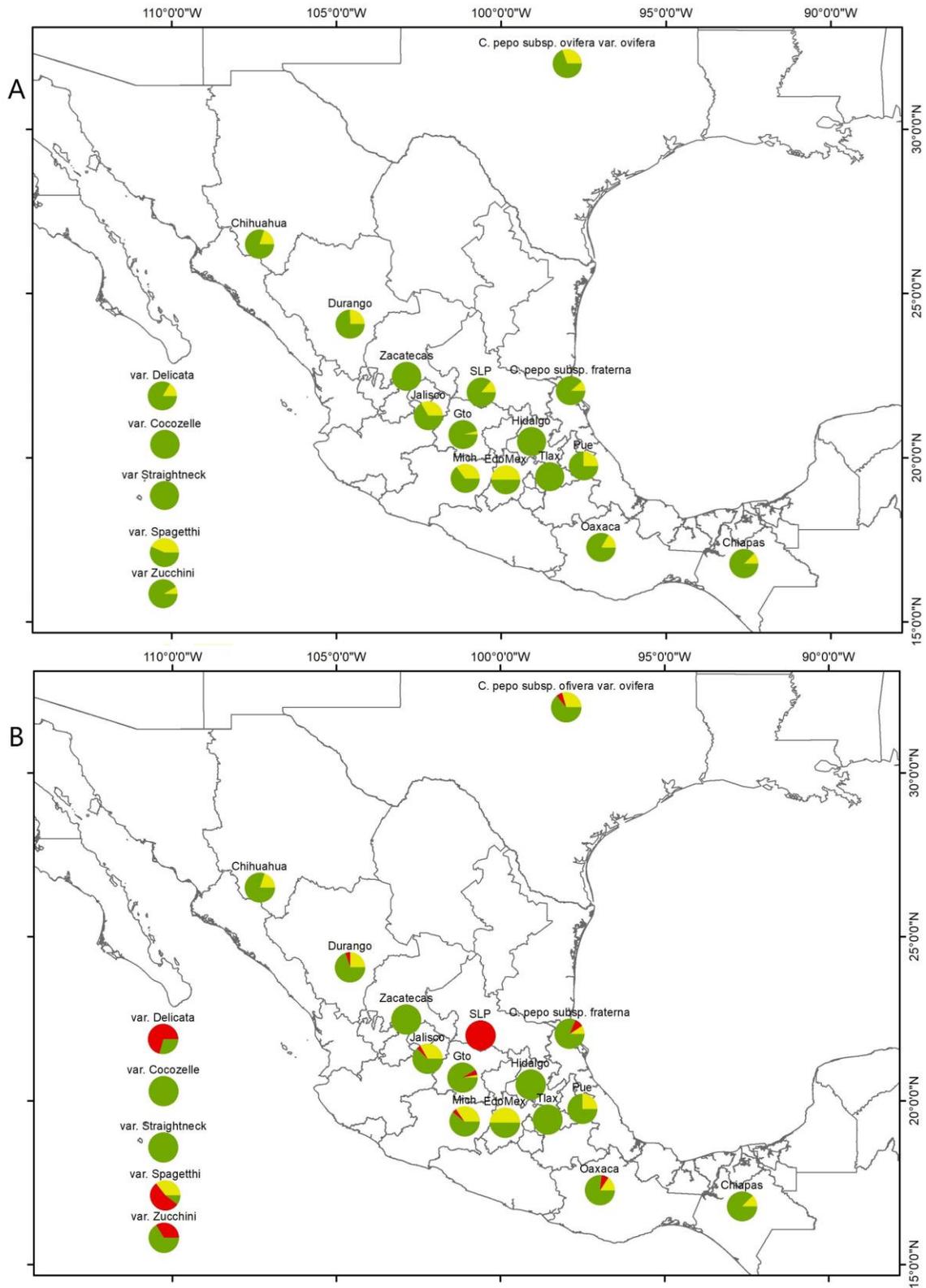


Figura 5. A) Proporciones de K = 2 sugeridos por el análisis Admixture, por localidad muestreada, y para las accesiones comerciales. B) Proporciones de K = 3 sugeridos por el análisis Admixture, por localidad muestreada, y para las accesiones comerciales.

Al graficar las proporciones de los genotipos de las diferentes accesiones en la zona geográfica donde fueron colectados para $K = 2$ y $K = 3$ (Figura 5), se puede notar que no hay un patrón claro de distribución geográfica de los genotipos.

En $K = 2$ (Figura 5 A) se puede observar que predomina el genotipo verde en todas las accesiones de *C. pepo* (82.67 %), y parece estar más relacionado a *C. pepo* subsp. *fraterna*. Mientras que el genotipo amarillo se encuentra en menor proporción (17.32 %), pero parece estar más relacionado a *C. pepo* subsp. *ovifera* var. *ovifera*.

Por otro lado, para $K = 3$ (Figura 5 B), podemos observar que *C. pepo* subsp. *fraterna* y *C. pepo* subsp. *ovifera* var. *ovifera* comparten los 3 genotipos, pero en diferentes proporciones. La subespecie *fraterna* tiene principalmente genotipo verde (81.5 %), y un poco de rojo (9.1 %) y amarillo (9.4 %). Por otro lado, la subespecie *ovifera* presenta mayoritariamente genotipo verde (55.0 %), seguido de amarillo (37.1 %) y en baja proporción rojo (7.9 %). Mientras que la accesión de San Luis Potosí resalta del resto de las accesiones locales, al ser la única que presenta un 100% de genotipo rojo.

6.4 Flujo génico

El análisis de diferenciación y mezcla poblacional resultó en que la mayor probabilidad de eventos de migración entre las diferentes accesiones locales mexicanas y las subespecies *fraterna* y *ovifera* es elevado (10 eventos migratorios, $\ln = -69.8068$) (Figura 6). Los eventos migratorios más probables se sugiere fueron de la localidad Estado de México a Chihuahua y Michoacán. Mientras que los menos probables se fueron de la subespecie *ovifera* a la localidad de Guanajuato y de la localidad Jalisco a la subespecie *ovifera*.

La mayoría de los eventos migratorios inferidos supuestamente fueron dentro de las mismas localidades mexicanas, siendo la localidad de Estado de México la que más eventos migratorios presenta. Pero en general, no se observa un patrón claro de intercambio de individuos entre las localidades.

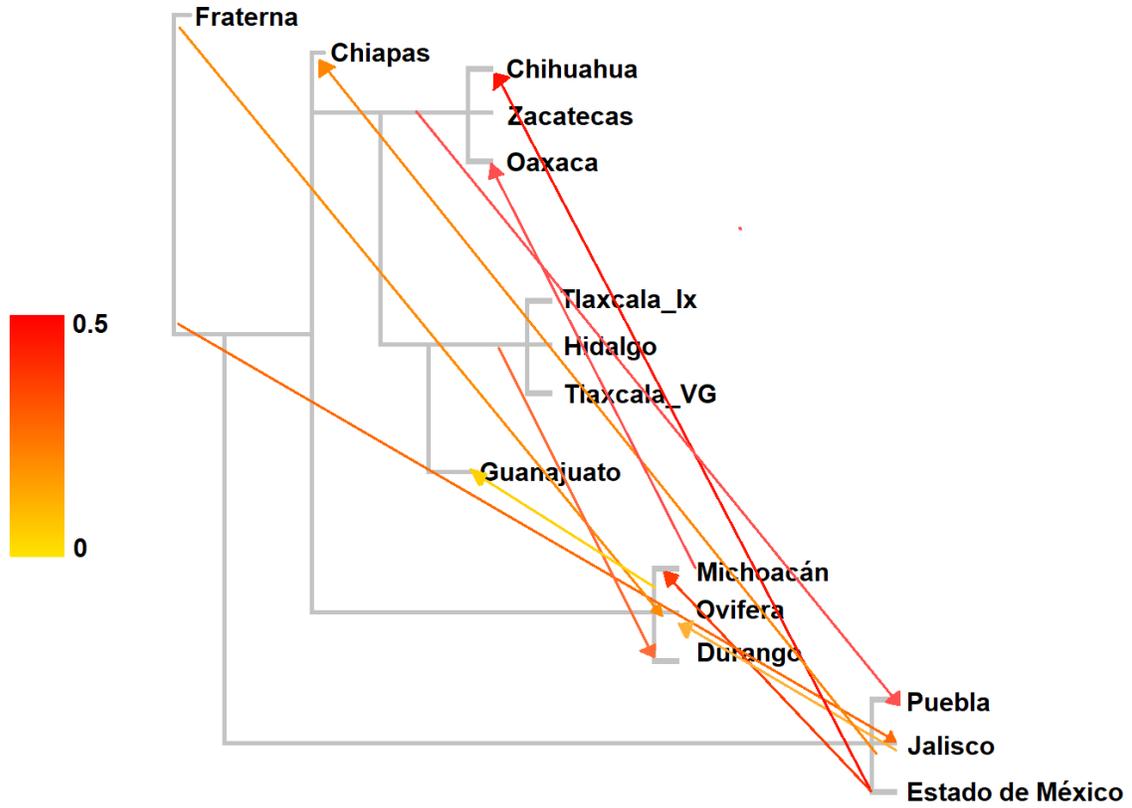


Figura 6. Análisis de los 10 eventos migratorios históricos más probables entre los diferentes grupos genéticos sugeridos por el programa TreeMix. Las flechas señalan en qué dirección se dio el evento migratorio, y el color la probabilidad de éste.

7. Discusión

7.1 Calidad de datos

En este trabajo se analizó solo el 2.5% del total de SNPs encontrados en el llamado inicial. En comparación con otros estudios donde trabajan con SNPs, este porcentaje es relativamente bajo (Frascaroli et al., 2013 (80%); Hufford et al., 2013 (69%); Würschum et al., 2013 (16%); Müller et al., 2015 (11 %)). En los trabajos mencionados, para la obtención de SNPs emplean diferentes estrategias, como el uso de chips de secuenciación Illumina o preparación de librerías por GBS, lo que les permite una mayor obtención de reads, además las plataformas de secuenciación empleadas son de Illumina. La forma de preparación de las librerías y la plataforma de secuenciación tienen un alto impacto en la obtención final de los SNPs (Jones et al., 2015). En este trabajo las librerías se prepararon por la técnica de tGBS, la cual al emplear cebadores selectivos, donde se pueden usar de 1 a 4 sitios degenerados durante la etapa de PCR, permite una reducción adicional del genoma, resultando en más lecturas por sitio disponibles para el llamado de SNPs (Ott et al., 2017). Nuestros datos fueron generados plataforma de secuenciación, Ion Torrent que es más rápida que Illumina, y puede secuenciar fragmentos más grandes, lo que permite que a la hora de mapear los fragmentos en un genoma de referencia obtengamos resultados de mejor calidad. Presenta la desventaja de que la calidad PHRED de secuenciación es menor que la de Illumina (Quail et al., 2012). Como consecuencia de esto, no se podía ser tan estricto con los primeros filtros de calidad, ya que se corría el riesgo de perder demasiados datos. Sin embargo, como la preparación de librerías con tGBS nos da una mayor profundidad en las lecturas, esto permite que los siguientes filtros de calidad aplicados durante el mapeo y el llamado de SNPs sean más robustos, de tal forma que aunque se pierdan muchos de los SNPs del llamado inicial, podemos aseverar que los SNPs empleados en los análisis son verdaderos y no errores de secuenciación.

7.2 Diversidad genética

En este trabajo se encontró que el promedio de la diversidad genética de *C. pepo* subsp. *ovifera* var. *ovifera* ($H_E = 0.2344$) es mayor que el promedio encontrado por nosotros para la subespecie silvestre *C. pepo* subsp. *fraterna* ($H_E = 0.1735$) y el promedio de las poblaciones locales de México de *C. pepo* subsp. *pepo* ($H_E = 0.1627$) Esto puede deberse al hecho de que las accesiones empleadas en este estudio son una variedad comercial usada con fines ornamentales, y no es directamente la forma silvestre de Estados Unidos (*C. pepo* subsp. *ovifera* var. *texana*), por lo que estas han pasado por un manejo agrícola, que tal vez incluya hibridación con otras variedades y selección artificial (Andres, 1987; Robinson y Shail, 1987; Cohen et al., 2003; Paris et al., 2003; Formisano et al., 2010). Por otro lado, *C. pepo* subsp. *ovifera* var. *ovifera* es ampliamente cultivada en Estados Unidos (Nee, 1990; Decker-Walters, et al., 2002; Quemada et al., 2008). Aunado a esto, existen estudios previos que sugieren que las poblaciones de *C. pepo* subsp. *ovifera* no sufrieron grandes cambios demográficos históricos durante el Pleistoceno, a diferencia de *C. pepo* subsp. *fraterna* (Castellanos-Morales et al., 2018).

En cuanto a que la diversidad genética de las variedades cultivadas en México de *C. pepo* subsp. *pepo* y la subespecie silvestre *C. pepo* subsp. *fraterna* sean similares puede deberse a que *C. pepo* subsp. *fraterna* es el ancestro silvestre más probable de *C. pepo* subsp. *pepo*, (Andres, 1987; Castellanos-Morales et al., 2019), por lo que es esperable que sus niveles de diversidad genética sean similares.

Dentro de las accesiones mexicanas de *C. pepo* subsp. *pepo*, encontramos que si bien el promedio de $H_E = 0.1627$, este valor presenta un amplio rango, que va de $H_E = 0.0508$ en San Luis Potosí a $H_E = 0.2575$ en Michoacán. Las diferencias de heterocigosis esperada entre las diferentes localidades de México puede deberse al tipo de manejo que se realiza en cada región. Mientras que en ciertos lugares de México aún se practica el cultivo tradicional de milpa, que ayuda a mantener la diversidad genética de las poblaciones, en otros lugares se ha

sustituido por cultivos más tecnificados, donde se emplean semillas de variedades mejoradas (Sánchez de la Vega, 2017). Sin embargo, debido a que la mayoría de nuestras muestras provienen del banco de germoplasma del Campo Experimental Bajío del INIFAP, no es posible aseverar esto.

Si comparamos la diversidad genética encontrada en *C. pepo* con la encontrada en otras especies de calabazas cultivadas en México empleando SNPs (Tabla 5), podemos ver que *C. pepo* subsp. *ovifera* var. *ovifera*, sigue siendo la más diversa de todas.

En cuanto a *C. pepo* subsp. *pepo*, la diversidad genética de la variedad cultivada en México parece ser alta si la comparamos con *C. moschata*, la cual es una especie de calabaza de la que se desconoce su ancestro silvestre. Pero si la comparamos con *C. argyrosperma* subsp. *argyrosperma*, su heterocigosis esperada es muy parecida a la de las variedades de *C. pepo* subsp. *pepo* cultivadas en México, pero mayor a las variedades comerciales de Estados Unidos.

Por otro lado, al comparar a la subespecie silvestre *C. argyrosperma* subsp. *sororia*, podemos notar que esta presenta una mayor heterocigosis que la subespecie silvestre *C. pepo* subsp. *fraterna* y que la subespecie cultivada *C. pepo* subsp. *pepo*, tanto las variedades de México como las comerciales de Estados Unidos. Por lo que podemos decir que en general la especie de calabaza *C. pepo* presenta una diversidad genética alta, aunque ésta no es homogénea entre sus subespecies.

Tabla 5. Valores de diversidad genética reportados para otras especies de calabazas con SNPs. Heterocigosis esperada (H_E); Coeficientes de endogamia (F_{IS}).

Especie	Subespecie	H_E	F_{IS}	No. SNPs	Autor
<i>C. pepo</i>	subsp. <i>pepo</i> de México	0.1627	0.1648	12,881	En este trabajo
	comerciales de EUA	0.11794	0.04956		
	subsp. <i>fraterna</i>	0.1735	0.1812		
	subsp. <i>ovifera</i> var. <i>ovifera</i>	0.2344	0.2146		
<i>C. argyrosperma</i>	subsp. <i>argyrosperma</i>	0.156		16,191	Sánchez-de la Vega, en Proceso;
	subsp. <i>sororia</i>	0.1914			Barrera, en proceso
<i>C. moschata</i>		0.132	0.06		Hernández-Rosales, en proceso

La diversidad genética en *C. pepo* ya había sido previamente reportada para cloroplasto y microsatélites (Castellanos-Morales et al., 2019; Ruiz Mondragón, 2017), siendo más alta con microsatélites, debido a la alta tasa de mutación que estos presentan (Schlötterer, 2004). El patrón de diversidad genética presente en *C. pepo* puede estar relacionado a múltiples factores, como que *C. pepo* fue una de las primeras especies cultivadas en Mesoamérica, con registros arqueológicos en Oaxaca y Tamaulipas que datan desde hace 10,000 y 4,500 años respectivamente (Decker, et al., 1993; Flannery, 1986; Nee, 1990; Smith, 1997), pasando por un proceso rápido de expansión y diversificación después de la domesticación (Castellanos-Morales et al., 2019). Durante este tiempo pudo haberse acumulado variación genética, sin resultar muy afectada por el cuello de botella que implica la domesticación, por ser cultivadas bajo el sistema de milpa.

Cuando analizamos la diversidad genética reportada para otras especies de la milpa, como el maíz (*Zea mays*) y el frijol común (*Phaseolus vulgaris*), encontramos que en general parece que las variedades que se cultivan tradicionalmente en Mesoamérica presentan una mayor diversidad genética en comparación con cultivos de las mismas especies en otros países o de líneas mejoradas (Blair et al., 2013; Frascaroli et al., 2013; Hufford et al., 2013; Müller et al., 2015).

El hecho de que tanto las variedades comerciales como las variedades cultivadas en México de *C. pepo* subsp. *pepo* presenten una diversidad genética menor que *C. pepo* subsp. *ovifera* var. *ovifera*, puede deberse al hecho de que históricamente, las poblaciones de las diferentes subespecies de *C. pepo* han sufrido eventos de contracción y expansión poblacional, siendo que la subespecie *C. pepo* subsp. *ovifera* la que ha presentado una menor reducción de sus poblaciones en comparación a la subespecie *C. pepo* subsp. *fraterna* (Castellanos-Morales et al., 2018), y al ser *C. pepo* subsp. *fraterna* el pariente silvestre más probable de *C. pepo* subsp. *pepo*, es lógico que ésta también se viera afectada por dichos eventos de contracción y expansión poblacional.

Sin embargo, a pesar de esta alta diversidad genética, encontramos altos valores de endogamia para las tres subespecies, especialmente si los comparamos con los valores obtenidos con SNPs reportados para otras calabazas cultivadas de manera tradicional en México (Tabla 5).

Por otro lado, al comparar los valores obtenidos para calabazas cultivadas en México con microsatélites (Tabla 6), observamos que los coeficientes de endogamia de *C. pepo* son mayores que los reportados para otras especies de calabazas, excepto por la subespecie *C. pepo* subsp. *ovifera* var. *ovifera*, lo que nos indica que en realidad la endogamia de esta calabaza es alta.

Tabla 6. Valores de diversidad genética reportados para otras especies de calabazas con microsatélites nucleares. Heterocigosis esperada (H_E); Coeficientes de endogamia (F_{IS}).

Especie	Subespecie	H_E	F_{IS}	No. loci	Autor
<i>C. pepo</i>	subsp. <i>pepo</i> México	0.573	0.123	10	Castellanos-Morales et al., 2019
	comerciales de EUA	0.638	0.101		
	subsp. <i>fraterna</i>	0.292	0.559		
	subsp. <i>ovifera</i> var. <i>ovifera</i>	0.564	-0.089		
<i>C. argyrosperma</i>	subsp. <i>argyrosperma</i>	0.41	0.033	9	Sánchez-de la Vega et al., 2018
	subsp. <i>sororia</i>	0.428	0.077		
<i>C. moschata</i>		0.395	0.18	11	Hernández-Rosales, en proceso

Las estimaciones de coeficiente de endogamia entre microsatélites y SNPs no son comparables de manera directa, debido al tipo de muestreo del genoma que realizan cada uno de estos marcadores (Waples y Gaggiotti, 2006; Allendorf y Luikart, 2007; Butler, 2012). Con los SNPs tenemos un muestreo más amplio de todo el genoma, que nos da una mejor idea de la variación genética global y local a lo largo del genoma, permitiéndonos en principio distinguir entre selección y eventos demográficos como la endogamia (Chakraborty et al., 1999; Gill, 2001; Helyar et al., 2011).

7.3 Estructura, diferenciación y flujo genético

En cuanto a la estructura genética, para la prueba de F_{ST} pareada para las diferentes subespecies y variedades de *C. pepo* encontramos que no existe estructura

genética entre estas. Los valores de F_{ST} más altos, pero sin tener diferencias estadísticamente significativas, fueron entre *C. pepo* subsp. *fraterna* y *C. pepo* subsp. *ovifera*, y entre *C. pepo* subsp. *fraterna* y las variedades comerciales de Estados Unidos. La falta de estructura genética resultante del análisis de F_{ST} , puede deberse a que dentro de cada grupo asignado para la prueba se encuentran individuos que presentan evidencias de flujo genético con otras subespecies y/o variedades, o que incluso pertenecen a otros grupos genéticos, como lo reflejan las pruebas de IBS (Figura 3) y Admixture (Figura 4).

En los análisis de asignación de los individuos de *C. pepo* a grupos genéticos, encontramos que existen tres claros grupos genéticos, como se había reportado previamente, pero los grupos son diferentes (Paris et al., 2003; Castellanos-Morales et al., 2019). Encontrados que con SNPs los grupos que se forman no presentan una asignación clara ni por subespecie ni por morfología. Por ejemplo, los tres individuos de *C. pepo* subsp. *ovifera* var. *ovifera* que en el IBS se agrupan en el grupo verde (Figura 3), junto con *C. pepo* subsp. *fraterna* y *C. pepo* subsp. *pepo*. La asignación de estos tres individuos al grupo genético verde podría deberse cuando menos en parte al origen polifilético de *C. pepo* subsp. *ovifera*, sugerido en la filogenia de 44 loci nucleares de Kates y colaboradores (2017).

En general, podemos notar que *C. pepo* subsp. *fraterna* presenta una mayor proporción de genotipo verde, aunque dos individuos de *C. pepo* subsp. *ovifera* var. *ovifera* son mayoritariamente genotipo amarillo, mientras que el genotipo rojo se encontró en mayor frecuencia en las variedades comerciales. Esto sugiere que la estructura genética de *C. pepo* encontrada con SNPs está relacionada a la proporción del genotipo que compone a los individuos, que pueden ser más cercanos a la subespecie *C. pepo* subsp. *fraterna* (grupo verde), más cercanos a la subespecie *C. pepo* subsp. *ovifera* var. *ovifera* (grupo amarillo) y a las variedades comerciales mejoradas (grupo rojo) (Figuras 3 y Figura 4).

Esta asignación de los individuos puede deberse a diferencias en la resolución que dan los marcadores empleados. Por un lado, Paris et al. (2003) emplearon marcadores AFLPs, los cuales son especialmente útiles para analizar la

diversidad genética de especies no modelo — ya que no requieren un conocimiento previo sobre las secuencias de cebadores para la especie objetivo — pueden tener limitaciones, por ejemplo debido a que en algunos casos son difíciles de reproducir (Zabeau y Vos, 1993; Vos et al., 1995; Schlötterer, 2004). Además, una desventaja que muestran es que al ser un marcador dominante, reduce su poder en los análisis de diversidad y la consanguinidad dentro de especies cultivadas (Marsjan y Oldenbroek, 2006).

Por otro lado, Castellanos-Morales et al., (2019) emplearon microsatélites nucleares, los cuales se caracterizan por ser altamente polimórficos, abundantes y estar distribuidos de manera uniforme en todos los genomas. Además, presentan una tasa de mutación alta, lo que permite tener una mejor aproximación de la genética poblacional de la especie o taxa analizado, especialmente si es de origen reciente (a veces hasta cientos de años o aún menos), como pueden ser las especies domesticadas (Schlötterer, 2004; Helyar et al., 2011). Sin embargo, la incertidumbre acerca de las tasas de mutación y el patrón existente en los microsatélites complica enormemente el análisis y la robustez de estos (Ellegren, 2000; Schlötterer, 2000).

Por último, los SNPs tienen como ventaja principal su alto potencial de obtener muchos marcadores a un costo moderado (Schlötterer, 2004), pero son menos variables en comparación con los microsatélites, ya que tienen una tasa de mutación menor (Waples y Gaggiotti, 2006; Allendorf y Luikart, 2007; Foll y Gaggiotti, 2008). Además, presentan la desventaja de que actualmente el poder computacional y los programas disponibles sólo permiten analizar un locus dos alelos, lo que tiene por consecuencia que la probabilidad de que se identifique un SNPs en el panel de verificación está en función de la frecuencia de alelos mínimos (MAF). Es decir, aquellos SNPs con un MAF muy bajo — que suelen ser los más recientes — tienen menor probabilidad de ser detectados, en comparación con aquellos con un MAF más alto — que suelen ser los más antiguos — (Helyar et al., 2011). Por lo que todo análisis dependiente de las frecuencias alélicas — como lo es la estructura genética — puede verse afectado (Würschum et al., 2013).

Sin embargo, al tener un muestreo mucho más amplio de todo el genoma, en lugar de tener menos de 50 marcadores, los SNPs nos permiten tener cientos o miles de marcadores genéticos. Análisis computacionales han demostrado que, en promedio, sólo se necesitan de 25 a 50 loci SNPs para obtener probabilidades equivalentes de coincidencia aleatoria como las obtenidas con microsatélites (Chakraborty et al., 1999; Gill, 2001). Por lo que los SNPs no reflejarán una estructura genética muy reciente, sino una más relacionada al flujo genético histórico y la retención de caracteres ancestrales de *C. pepo*. Como lo podemos observar en las pruebas de IBS (Figura 3) y de Treemix (Figura 6), donde encontramos que en *C. pepo* ha existido mucho flujo genético histórico, relacionado a su domesticación (Castellanos-Morales et al., 2019). El cual en su mayoría se lleva a cabo para realizar mejoramiento de las variedades (Cohen et al., 2003; Paris et al., 2003; Formisano et al., 2010).

Ejemplo de esto son las accesiones de *C. pepo* subsp. *fraterna* y las muestras de *C. pepo* subsp. *pepo* para las localidades de México, que no se encuentran dentro del grupo genético verde del IBS (Figura 3), y que los análisis sugiere flujo genético con variedades comerciales. Éstas aunque en el análisis de ancestría presentan genotipo mayoritariamente verde, también están compuestas por genotipo rojo y/o amarillo en distintas proporciones (Figura 4).

El mismo patrón encontramos para las variedades comerciales de *C. pepo*, las cuales se distribuyen entre los tres grupos genéticos encontrados en el IBS (Figura 3), pero al observar su composición genotípica existe una correspondencia de la proporción del genotipo con asignación al grupo genético del IBS. Esta composición genética tan variada que presentan las accesiones comerciales puede ser el resultado de un proceso de hibridación entre las diferentes subespecies y variedades de *C. pepo*. Aunque también puede deberse al hecho de que a veces se toma una de las diferentes subespecies y variedades de *C. pepo* como uno de los parentales y se cruza con alguna otra especie del género *Cucurbita*, especialmente para producir formas resistentes a enfermedades (Andres, 1987; Robinson y Shail, 1987; Cohen et al., 2003; Paris et al., 2003; Paris, 2018; Castellanos-Morales et al., 2019).

La localidad San Luis Potosí presenta evidencias claras de que representa un cultivo de la variedad comercial *Spaghetti* (Figura 3 y 4). *C. pepo* es la calabaza más comercializada a nivel mundial, por lo que el acceso a las semillas de variedades comerciales para cultivo es relativamente sencillo. Por ejemplo, personas que migran de las grandes ciudades a zonas más rurales, pueden acceder a terrenos y sembrar sus propias "milpas" sin que necesariamente las semillas provengan de un cultivo tradicional, o incluso pueden conseguir semillas en supermercados y emplearlas para cultivo de consumo personal o comercio muy local (Sr. José Luis Aboites, San Miguel de Allende, com. pers.). Por lo tanto, es muy probable que las muestras de esta localidad provengan o de un monocultivo tecnificado, los cuales son cada vez más comunes en la región del Bajío Mexicano, o de algún terreno de personas que implementaron sus propios cultivos con semillas de esta variedad comercial.

En cuanto a la mayoría de las variedades de *C. pepo* subsp. *pepo* cultivadas en México, podemos notar que en la red IBS (Figura 3) éstas se agrupan con las muestras asignadas al grupo genético verde en el análisis de ancestría (Figura 4). Además, en el análisis de Treemix (Figura 6) podemos notar que hay dos líneas cruzadas, que representan los eventos de flujo genético con *C. pepo* subsp. *fraterna*. Siendo esto concordante con la hipótesis de que *C. pepo* subsp. *fraterna* es el ancestro silvestre más probable para las variedades cultivadas de manera tradicional en México de *C. pepo* subsp. *pepo*, con eventos de flujo genético históricos entre estas dos subespecies (Andres, 1987; Castellanos-Morales et al., 2019).

7.4 Relevancia de los recursos genéticos y su conservación

La diversidad fitogenética se centra principalmente en regiones conocidas como centros de diversidad – por ejemplo, Mesoamérica – (Harlan, 1971; Piperno y Pearsall, 1998; Purugganan y Fuller, 2009; Meyer y Purugganan, 2013). Los agricultores de centros de diversidad, suelen practicar la agricultura tradicional,

cultivando variedades locales que han sido seleccionadas a lo largo de muchas generaciones (Bastías Marín, 2008).

En México, este sistema de cultivo tradicional es conocido como milpa, el cual conserva una importante cantidad de diversidad genética, ayudando así a la conservación de los recursos fitogenéticos relevantes para el desarrollo de las sociedades humanas (Bastías Marín, 2008; Lobo, 2008). Desafortunadamente, el cambio de la sociedad ha puesto en peligro la diversidad producida durante el proceso de domesticación y el continuo proceso evolutivo desde las especies silvestres ancestrales hasta las variedades modernas (Bastías Marín, 2008; Lobo, 2008; Zizumbo et al., 2010). Con el paso de los años, la milpa ha sido gradualmente desplazada por sistemas de cultivo más tecnificados, en los cuales se ve disminuida la diversidad genética de los cultivos, ya que se maneja una sola variedad que suele ser una comercial mejorada, con menor diversidad genética, la cual puede substituir a las variedades nativas con mayor diversidad (Sánchez de la Vega, 2017).

En los últimos años se ha modificado radicalmente la distribución y estructura de la diversidad agrícola para la alimentación, perdiéndose más del 80%, debido a los procesos de desplazamiento de variedades tradicionales por mejoradas, y la eliminación de las poblaciones de las especies silvestres que dieron origen a las variedades domesticadas (Lobo, 2008; Zizumbo et al., 2010).

El cultivo de variedades comerciales de calabaza en México comenzó hace aproximadamente 25 años, principalmente en las regiones agrícolas de Sonora y Sinaloa (Sánchez de la Vega, 2017). En este trabajo encontramos evidencia de esta substitución, ya que las variedades de *C. pepo* subsp. *pepo* cultivadas en México de forma tradicional aparentemente tienen señales de flujo genético con variedades comerciales de Estados Unidos como es el caso de Zacatecas, Durango y Michoacán. Inclusive, en el caso de nuestra muestra de San Luis Potosí, encontramos un desplazamiento total de las variedades tradicionales mexicanas por variedades comerciales (*C. pepo* subsp. *pepo* var. Spaguetti).

Es posible que en el futuro la diversidad genética de las variedades locales de México se vea seriamente afectada por el aumento de sistemas de cultivo

intensivos, que tienen una mayor productividad al corto plazo, pero pone en riesgo la sustentabilidad del sistema agrícola (Zizumbo et al., 2010), debido a que se ha sugerido que la sustentabilidad de un sistema agrícola debería depender de que tan diverso en número de especies sea éste, ya que de esta diversidad dependen los servicios reguladores esenciales, y por tanto el soporte de la producción (Jarvis et al., 2007).

La diversidad genética de los recursos agrícolas permite a los cultivos adaptarse a diferentes ambientes y condiciones de crecimiento. Si consideramos la velocidad a la que se están perdiendo estos recursos genéticos (Bastías Marín, 2008; Zizumbo et al., 2010), además de la alta dependencia de las poblaciones rurales a la agricultura, –se estima que las poblaciones rurales dependen en 90% de los recursos biológicos para satisfacer sus necesidades de supervivencia (Shand, 1997) –es indispensable plantear un plan de conservación de los recursos genéticos almacenados en los cultivos tradicionales de nuestro país. Por ejemplo, se podría renovar y aumentar la colección de semillas de *C. pepo* subsp. *pepo* del banco de germoplasma con semillas que sean realmente de cultivos tradicionales, además de aumentar el número de muestras de las especies y subespecies silvestres de calabazas como *C. pepo* subsp. *fraterna*.

8. Conclusiones

En este trabajo se analizó con datos genómicos la estructura genética de las diferentes subespecies y variedades de *C. pepo*. Los resultados difieren de los previamente reportado (Paris et al., 2003; Castellanos Morales et al., 2019). Si bien se mantienen tres grupos genéticos, estos no se separan según la clasificación taxonómica, sino que aparentemente correlacionan con su cercanía evolutiva o geográfica, no queda claro! a *C. pepo* subsp. *fraterna*, *C. pepo* subsp. *ovifera*, o a las variedades comerciales.

Es muy probable que la estructura genética que encontramos con SNPs este determinada, cuando menos en parte, por retención de caracteres ancestrales o por incomplete lineage sorting (separación incompleta del linajes = polimorfismo ancestrales), ya que la divergencia entre estas subespecies es muy reciente, y la resolución a la que trabajan los SNPs no nos permite ver con tanta facilidad estructuras genéticas muy recientes como lo hacen los microsatélites.

También encontramos que evidencia que se puede interpretar como que las variedades tradicionales cultivadas en México de *C. pepo* subsp. *pepo* están siendo desplazadas por variedades comerciales mejoradas, especialmente en las regiones del bajío y norte de México. Sabemos que las variedades comerciales suelen tener una mayor resistencia a ciertas plagas (Paris, 2008), además pueden tener mayor valor económico en el mercado, pero ponen en riesgo la diversidad genética de las variedades tradicionales, afectando también el futuro de este valioso recurso fitogenético.

9. Perspectivas

Entender cómo funcionan los procesos de selección artificial de caracteres a nivel genómico en las especies domesticadas permitirá realizar mejoras en los cultivos para hacerlos más eficientes en su producción.

Este trabajo es un primer acercamiento para entender estos procesos en la calabaza *C. pepo*. Pero aún hace falta conocer exactamente cuáles son los genes que están siendo seleccionados por la domesticación. Para trabajar solo con los loci neutrales y poder observar la diversidad y estructura genética propia de la especie *C. pepo* sin ver los efectos de los loci con selección, se realizó la prueba de Bayescan, encontrando 51 loci bajo selección, de los cuales 19 presentaron señales de selección direccional, y 32 presentaron evidencia de selección balanceadora (Figura 7).

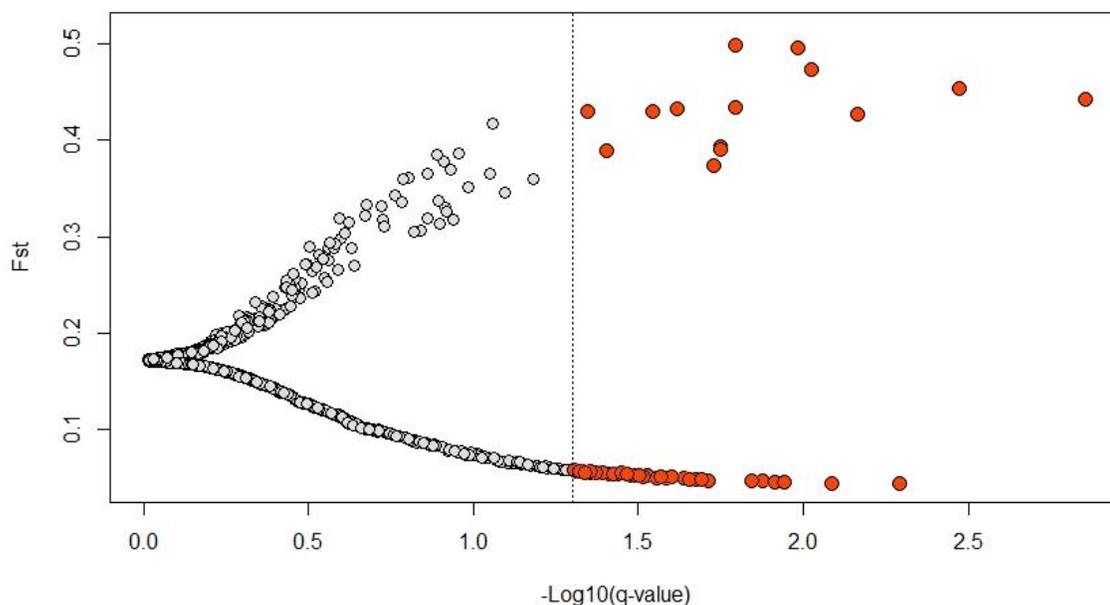


Figura 7. Gráfico de la probabilidad de los SNPs a presentar señales de selección VS sus valores de F_{ST} . De color naranja se encuentran todos los SNPs con $p < 0.05$ de presentar señales de selección.

Aun haría falta mapear y anotar estos loci en el genoma de referencia *C. pepo* subsp. *pepo* var. *Zucchini*, para entender las regiones y genes involucrados y de esta forma comenzar a entender realmente que es lo que se está seleccionando durante el proceso de domesticación y mejoramiento de esta calabaza.

Otro paso a dar será entender que genes comparten entre las diferentes subespecies y variedades de *C. pepo*. Para ello se podría hacer un análisis genómico de cada una de las subespecies reconocidas (*C. pepo* subsp. *pepo*, *C. pepo* subsp. *fraterna*, *C. pepo* subsp. *ovifera* var. *texana*, *C. pepo* subsp. *ovifera* var. *ozarkana* y *C. pepo* subsp. *ovifera* var. *ovifera*), de tal forma que se pueda ver que regiones del genoma se comparten entre ellas y cuales difieren, cuáles son las regiones genómicas que se comparten de forma ancestral entre las tres subespecies, y cuales se han ido modificando con el tiempo e incluso muestran señales de flujo genético antiguo o reciente entre ellas. Esto permitirá entender mejor que es lo que da la estructura genética de esta especie, y como ha sido su historia de domesticación.

Bibliografía

- Aguilar-Meléndez, A., Morrell, P. L., Roose, M. L., y Kim, S. C. (2009). Genetic diversity and structure in semiwild and domesticated chiles (*Capsicum Annuum*; Solanaceae) from Mexico. *American Journal of Botany*, 96(6), 1190–1202. <https://doi.org/10.3732/ajb.080015>
- Aguirre-Liguori, J. A., Tenaillon, M. I., Vázquez-Lobo, A., Gaut, B. S., Jaramillo-Correa, J. P., Montes-Hernandez, S., ... Eguiarte, L. E. (2017). Connecting genomic patterns of local adaptation and niche suitability in teosintes. *Molecular Ecology*, 26(16), 4226–4240. <https://doi.org/10.1111/mec.14203>
- Alexander, D. H., Shringarpure, S. S., Novembre, J., y Lange, K. (2015). *Admixture 1.3 Software Manual*. 3–4.
- Allendorf, F. W., y Luikart, G. (2007). *Conservation and the Genetics of Populations*. Oxford, UK.: Blackwell Publishing.
- Alo, F., Furman, B. J., Akhunov, E., Dvorak, J., y Gepts, P. (2011). Leveraging genomic resources of model species for the assessment of diversity and phylogeny in wild and domesticated lentil. *Journal of Heredity*, 102(3), 315–329. <https://doi.org/10.1093/jhered/esr015>
- Alverson, A. J., Wei, X., Rice, D. W., Stern, D. B., Barry, K., y Palmer, J. D. (2010). Insights into the evolution of mitochondrial genome size from complete sequences of *Citrullus lanatus* and *Cucurbita pepo* (Cucurbitaceae). *Molecular Biology and Evolution*, 27(6), 1436–1448. <https://doi.org/10.1093/molbev/msq029>
- Andres, T. C. (1987). *Cucurbita fraterna*, the closest wild relative and progenitor of *C. pepo*. *Cucurbit Genet Coop Rep*, 10, 69–71.
- Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data.
- Arriaga, L., Huerta, E., Lira-Saade, R., Moreno, E., y Alarcón, J. (2006). Assessing the risk of releasing transgenic *Cucurbita* spp. in Mexico. *Agriculture, Ecosystems and Environment*, 112(4), 291–299. <https://doi.org/10.1016/j.agee.2005.07.007>
- Avni, R., Nave, M., Barad, O., Baruch, K., Twardziok, S. O., Gundlach, H., ... Distelfeld, A. (2017). Wild emmer genome architecture and diversity elucidate wheat evolution and domestication. *Science*, 97, 93–97. <https://doi.org/10.1126/science.aan0032>
- Barzegar, R., Peyvast, G., Mohamad, A., Rabiei, B., Akbar, A., y Babagolzadeh, A. (2013). Biochemical systematic, population structure and genetic variability studies among Iranian *Cucurbita* (*Cucurbita pepo* L.) accessions, using genomic SSRs and implications for their breeding potential. *Biochemical Systematics and Ecology*, 50, 187–198. <https://doi.org/10.1016/j.bse.2013.03.048>

- Bastías Marín, E. (2008). Biodiversidad Y Recursos Fitogenéticos En La Agricultura. *Idesia (Arica)*, 26(1), 5–7. <https://doi.org/10.4067/S0718-34292008000100001>
- Bauchet, G., y Causse, M. (2012). Genetic Diversity in Tomato (*Solanum lycopersicum*) and Its Wild Relatives. In *Genetic Diversity in Plants*. <https://doi.org/10.5772/33073>
- Bazzaz, F. (1979). The physiological ecology of plant succession. *Annual Review of Ecology and Systematics*, 10(1), 351–371.
- Beissinger, T. M., Hirsch, C. N., Sekhon, R. S., Foerster, J. M., Johnson, J. M., Muttoni, G., ... de Leon, N. (2013). Marker density and read depth for genotyping populations using genotyping-by-sequencing. *Genetics*, 193(4), 1073–1081. <https://doi.org/10.1534/genetics.112.147710>
- Bentley, D. R., Balasubramanian, S., Swerdlow, H. P., Smith, G. P., Milton, J., Brown, C. G., ... Smith, A. J. (2008). Accurate whole human genome sequencing using reversible terminator chemistry. *Nature*, 456(7218), 53–59. <https://doi.org/10.1038/nature07517>
- Blair, M. W., Cortés, A. J., Penmetsa, R. V., Farmer, A., Carrasquilla-Garcia, N., y Cook, D. R. (2013). A high-throughput SNP marker system for parental polymorphism screening, and diversity analysis in common bean (*Phaseolus vulgaris* L.). *Theoretical and Applied Genetics*, 126(2), 535–548. <https://doi.org/10.1007/s00122-012-1999-z>
- Blancas, J., Casas, A., Rangel-Landa, S., Moreno-Calles, A., Torres, I., Pérez-Negrón, E., ... Dávila, P. (2010). Plant Management in the Tehuacán-Cuicatlán Valley, Mexico. *Economic Botany*, 64(4), 287–302. <https://doi.org/10.1007/s12231-010-9133-0>
- Blanckaert, I., Paredes-Flores, M., Espinosa-García, F. J., Piñero, D., y Lira, R. (2012). Ethnobotanical, morphological, phytochemical and molecular evidence for the incipient domestication of Epazote (*Chenopodium ambrosioides* L.: Chenopodiaceae) in a semi-arid region of Mexico. *Genetic Resources and Crop Evolution*, 59(4), 557–573. <https://doi.org/10.1007/s10722-011-9704-7>
- Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., y Buckler, E. S. (2007). TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics*, 23(19), 2633–2635. <https://doi.org/10.1093/bioinformatics/btm308>
- Buermans, H. P. J., y den Dunnen, J. T. (2014). Next generation sequencing technology: Advances and applications. *Biochimica et Biophysica Acta - Molecular Basis of Disease*, 1842(10), 1932–1941. <https://doi.org/10.1016/j.bbadis.2014.06.015>
- Butler, J. M. (2012). Single Nucleotide Polymorphisms and Applications. In *Advanced Topics in Forensic DNA Typing* (pp. 347–369). <https://doi.org/10.1016/b978-0-12-374513-2.00012-9>
- Byers, D. S. (1967). *The prehistory of the Tehuacan Valley, Vol 1 Environment and Subsistence* (D. S. Byers, Ed.). Londres: University of Texas Press.

- Caboche, S., Audebert, C., Lemoine, Y., y Hot, D. (2014). Comparison of mapping algorithms used in high-throughput sequencing: Application to Ion Torrent data. *BMC Genomics*, 15(1), 1–16. <https://doi.org/10.1186/1471-2164-15-264>
- Casas, A., Vázquez, C., J. L. Viveros y Caballero, J. (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(4), 455–478.
- Casas, A., Torres-Guevara, J., y Parra, F. (Eds.). (2016). *Domesticación en el continente americano* (Vol. 1). Universidad Nacional Autónoma de México.
- Castellanos-Morales, G., Paredes-Torres, L. M., Gámez, N., Hernández-Rosales, H. S., Sánchez-de la Vega, G., Barrera-Redondo, J., ... Eguiarte, L. E. (2018). Historical biogeography and phylogeny of *Cucurbita*: Insights from ancestral area reconstruction and niche evolution. *Molecular Phylogenetics and Evolution*, 128(April), 38–54. <https://doi.org/10.1016/j.ympev.2018.07.016>
- Castellanos-Morales, G., Ruiz-Mondragón, K. Y., Hernández-Rosales, H. S., Sánchez-de la Vega, G., Gámez, N., Aguirre-Planter, E., ... Eguiarte, L. E. (2019). Tracing back the origin of pumpkins (*Cucurbita pepo* ssp. *pepo* L.) in Mexico. *Proceedings of the Royal Society B: Biological Sciences*, 286(1908), 20191440. <https://doi.org/10.1098/rspb.2019.1440>
- Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., y Cresko, W. A. (2013). Stacks: An analysis tool set for population genomics. *Molecular Ecology*, 22(11), 3124–3140. <https://doi.org/10.1111/mec.12354>
- Chakraborty, R., N., S. D., Birg, S., Yixi, Z., y Bruce, B. (1999). The utility of short tandem repeat loci beyond human identification: Implications for development of new DNA typing systems. *Electrophoresis*, 20(8), 1682–1696. [https://doi.org/doi:10.1002/\(SICI\)1522-2683\(19990101\)20:8<1682::AID-ELPS1682>3.0.CO;2-Z](https://doi.org/doi:10.1002/(SICI)1522-2683(19990101)20:8<1682::AID-ELPS1682>3.0.CO;2-Z)
- Chan, K. F., y Sun, M. (1997). Genetic diversity and relationships detected by isozyme and RAPD analysis of crop and wild species of *Amaranthus*. *Theoretical and Applied Genetics*, 95, 865–873. <https://doi.org/10.1007/s001220050637>
- Cohen, R., Hanan, A., y Paris, H. (2003). Single-gene resistance to powdery mildew in zucchini squash (*Cucurbita pepo*). *Euphytica*, 130(3), 433–441. <https://doi.org/10.1023/A:1023082612420>
- Davey, J. L., y Blaxter, M. W. (2010). RADseq: Next-generation population genetics. *Briefings in Functional Genomics*, 9(5–6), 416–423. <https://doi.org/10.1093/bfpg/elq031>
- Decker-Walters D S, T W Walters, U Poluszny, P. G. K. (1990). Genealogy and gene flow among annual domesticated species of *Cucurbita*. *Can. J. Bot.*, 68, 782–789.

- Decker-Walters S, D., Staub, J. E., Chung, S.-M., Naata, E., y Quemada, H. D. (2002). Diversity in free-living populations of *Cucurbita pepo* (Cucurbitaceae) as assessed by Random Amplified Polymorphic DNA. *Systematic Botany*, 27(1), 19–28.
- Decker-Walters S, D., y Wilson, H. D. (1987). Allozyme Variation in the *Cucurbita pepo* Complex : *C. pepo* var. *ovifera* vs. *C. texana*. Published by : American Society of Plant Taxonomists
Stable URL : <http://www.jstor.org/stable/2419320> JSTOR is a n. *Systematic Botany*, 12(2), 263–273.
- Decker, D. S. (1988). Origin(s), Evolution , and Systematics of *Cucurbita pepo* (Cucurbitaceae) 1. 42(August 1987), 4–15.
- Decker, D. S., Walters, T. W., Wesley Cowan, C., y Smith, B. D. (1993). Isozymic characterization of wild populations of *Cucurbita pepo*. *J. Ethnobiol*, 13(131), 55–72. Retrieved from <https://ethnobiology.org/sites/default/files/pdfs/JoE/13-1/Decker-Walters.et.al.pdf>
- Dillehay, T. D., Rossen, J., Andres, T. C., y Williams, D. E. (2007). Pre-ceramic adoption of peanut, squash, and cotton in northern Peru. *Science*, 316(5833), 1890–1893.
- Egan, A.N., Schlueter, J. and Spooner, D. M. (2012). Applications of next-generation sequencing in plant biology. *American Journal of Botany*, 99, 175–185.
- Ellegren, H. (2000). Microsatellite mutations in the germline: implications for evolutionary inference. *Trends in Genetics*, 16, 551–558.
- Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., y Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One*, 6(5), e19379. <https://doi.org/10.1371/journal.pone.0019379>
- Enríquez, E., Landaverde-González, P., Lima-Cordón, R., Solórzano-Ortíz, E., Tapia-López, R., y Nuñez-Farfán, J. (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(3), 979–991. <https://doi.org/10.1007/s10722-017-0589-y>
- Esteras, C., Formisano, G., Roig, C., Díaz, A., Blanca, J., Garcia-Mas, J., ... Picó, B. (2013). SNP genotyping in melons: Genetic variation, population structure, and linkage disequilibrium. *Theoretical and Applied Genetics*, 126(5), 1285–1303. <https://doi.org/10.1007/s00122-013-2053-5>
- Ferriol, M., Picó, B., y Nuez, F. (2003). Genetic diversity of a germplasm collection of *Cucurbita pepo* using SRAP and AFLP markers. *Theoretical and Applied Genetics*, 107(2), 271–282. <https://doi.org/10.1007/s00122-003-1242-z>
- Fischer, M. C., Foll, M., Excoffier, L., y Heckel, G. (2011). Enhanced AFLP genome scans detect local adaptation in high-altitude populations of a small rodent (*Microtus arvalis*). *Molecular*

- Ecology, 20(7), 1450–1462. <https://doi.org/10.1111/j.1365-294X.2011.05015.x>
- Flannery, K. V. (1986). *Guilá Naquitz: Archaic foraging and early agriculture in Oaxaca, Mexico*. Emerald Group Publishing.
- Flannery, Kent V. (1986). *Guilá Naquitz: Archaic foraging and early agriculture in Oaxaca, Mexico*. Emerald Group Publishing.
- Foll, M. (2012). BayeScan v2.1 user manual. *Ecology*, 20, 1450–1462.
- Foll, M., y Gaggiotti, O. (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics*, 180(2), 977–993. <https://doi.org/10.1534/genetics.108.092221>
- Formisano, G., Paris, H., Frusciante, L., y Ercolano, M. R. (2010). Commercial *Cucurbita pepo* squash hybrids carrying disease resistance introgressed from *Cucurbita moschata* have high genetic similarity. *Plant Genetic Resources: Characterisation and Utilisation*, 8(3), 198–203. <https://doi.org/10.1017/S1479262110000183>
- Frascaroli, E., Schrag, T. A., y Melchinger, A. E. (2013). Genetic diversity analysis of elite European maize (*Zea mays* L.) inbred lines using AFLP, SSR, and SNP markers reveals ascertainment bias for a subset of SNPs. *Theoretical and Applied Genetics*, 126(1), 133–141. <https://doi.org/10.1007/s00122-012-1968-6>
- Fustier, M. A., Brandenburg, J. T., Boitard, S., Lapeyronnie, J., Eguiarte, L. E., Vigouroux, Y., ... Tenailon, M. I. (2017). Signatures of local adaptation in lowland and highland teosintes from whole-genome sequencing of pooled samples. *Molecular Ecology*, 26(10), 2738–2756. <https://doi.org/10.1111/mec.14082>
- Gao, L. (2004). Population structure and conservation genetics of wild rice *Oriza rufipogon* (Poaceae): a region-wide perspective from microsatellite variation. *Molecular Ecology*, 13, 1009–1024.
- Garris, A. J., Tai, T. H., Coburn, J., Kresovich, S., y McCouch, S. (2005). Genetic structure and diversity in *Oryza sativa* L. *Genetics*, 169(3), 1631–1638. <https://doi.org/10.1534/genetics.104.035642>
- Gaut, B. S., Seymour, D. K., Liu, Q., Zhou, Y. (2018). Demography and its effects on genomic variation in crop domestication. *Nature Plants*, 4(August). <https://doi.org/10.1038/s41477-018-0210-1>
- Gepts, P. (2014a). The contribution of genetic and genomic approaches to plant domestication studies. *Current Opinion in Plant Biology*, 18(Figure 2), 51–59. <https://doi.org/10.1016/j.pbi.2014.02.001>

- Gepts, P. (2014b). The contribution of genetic and genomic approaches to plant domestication studies. *Current Opinion in Plant Biology*, Vol. 18, pp. 51–59.
<https://doi.org/10.1016/j.pbi.2014.02.001>
- Gepts, P., y Papa, R. (2003). Evolution during Domestication. *Encyclopedia of Life Sciences*, 1–7.
<https://doi.org/10.1038/npg.els.0003071>
- Gerbault, P., Allaby, R. G., Boivin, N., Ruzdinski, A., Grimaldi, I. M., Pires, J. C., ... Thomas, M. G. (2014). Storytelling and story testing in domestication. *Proceedings of the National Academy of Sciences*, 111(17), 6159–6164. <https://doi.org/10.1073/pnas.1400425111>
- Gill, P. (2001). An assessment of the utility of single nucleotide polymorphisms (SNPs) for forensic purposes. *International Journal of Legal Medicine*, 114(4–5), 204–210.
<https://doi.org/10.1007/s004149900117>
- Gong, L., Stift, G., Kofler, R., Pachner, M., y Lelley, T. (2008). Microsatellites for the genus *Cucurbita* and an SSR-based genetic linkage map of *Cucurbita pepo* L. *Theoretical and Applied Genetics*, 117(1), 37–48. <https://doi.org/10.1007/s00122-008-0750-2>
- Gong, L., Pachner, M., Kalai, K., y Lelley, T. (2008). SSR-based genetic linkage map of *Cucurbita moschata* and its synteny with *Cucurbita pepo*. *Genome*, 51(11), 878–887.
<https://doi.org/10.1139/g08-072>
- Gong, Li, Paris, H., Nee, M. H., Stift, G., Pachner, M., Vollmann, J., y Lelley, T. (2012). Genetic relationships and evolution in *Cucurbita pepo* (pumpkin, squash, gourd) as revealed by simple sequence repeat polymorphisms. *Theoretical and Applied Genetics*, 124(5), 875–891.
<https://doi.org/10.1007/s00122-011-1752-z>
- González-Insuasti, M. S., y Caballero, J. (2007). Managing plant resources: How intensive can it be? *Human Ecology*, 35(3), 303–314. <https://doi.org/10.1007/s10745-006-9063-8>
- Goodwin, S., McPherson, J. D., y McCombie, W. R. (2016). Coming of age: Ten years of next-generation sequencing technologies. *Nature Reviews Genetics*, 17(6), 333–351.
<https://doi.org/10.1038/nrg.2016.49>
- Gordon, A., y Hannon, G. (2010). Fastx-toolkit. FASTQ/A short-reads pre-processing tools. Unpublished http://hannonlab.cshl.edu/fastx_tool.
- Goudet, J. (2005). HIERFSTAT, a package for R to compute and test hierarchical F-statistics. *Molecular Ecology Notes*, 5, 184–186. <https://doi.org/10.1111/j.1471-8278>
- Gross, B. L., y Olsen, K. M. (2010). Genetic perspectives on crop domestication. *Trends in Plant Science*, 15(9), 529–537. <https://doi.org/10.1016/j.tplants.2010.05.008>
- Guerra-García, A., Suárez-Atilano, M., Mastretta-Yanes, A., Delgado-Salinas, A., y Piñero, D.

- (2017). Domestication Genomics of the Open-Pollinated Scarlet Runner Bean (*Phaseolus coccineus* L.). *Frontiers in Plant Science*, 8(November), 1–15.
<https://doi.org/10.3389/fpls.2017.01891>
- Hanselka, J. K. (2017). Revisiting the Archaeobotanical Record of Romero's Cave in the Ocampo Region of Tamaulipas, Mexico. *Journal of Ethnobiology*, 37(1), 37–59.
<https://doi.org/10.2993/0278-0771-37.1.37>
- Harlan, J. R. (1971). Origins: agricultural centers and noncenters. *Science*, 174(4008), 468–474.
- Helyar, S. J., Hemmer-Hansen, J., Bekkevold, D., Taylor, M. I., Ogden, R., Limborg, M. T., ... Nielsen, E. E. (2011). Application of SNPs for population genetics of nonmodel organisms: New opportunities and challenges. *Molecular Ecology Resources*, 11(SUPPL. 1), 123–136.
<https://doi.org/10.1111/j.1755-0998.2010.02943.x>
- Hernández-Bermejo, J. G. (1994). Neglected crops: 1492 from a different perspective. In *Neglected Crops 1942 from a different perspective*. Retrieved from
<http://www.fao.org/docrep/018/t0646e/t0646e.pdf>
- Huang, X. Q., Börner, A., Röder, M. S., y Ganal, M. W. (2002). Assessing genetic diversity of wheat (*Triticum aestivum* L.) germplasm using microsatellite markers. *Theoretical and Applied Genetics*, 105(5), 699–707. <https://doi.org/10.1007/s00122-002-0959-4>
- Hufford, M. B., Lubinsky, P., Pyhäjärvi, T., Devengenzo, M. T., Ellstrand, N. C., y Ross-Ibarra, J. (2013). The Genomic Signature of Crop-Wild Introgression in Maize. *PLoS Genetics*, 9(5).
<https://doi.org/10.1371/journal.pgen.1003477>
- Hurd, P., y Linsley, E. (1964). The squash and gourd bees—genera *Peponapis* Robertson and *Xenoglossa* Smith—inhabiting America north of Mexico (Hymenoptera: Apoidea). *Hilgardia*, 35, 177–375.
- Hurd, P., y Linsley, E. (1966). The Mexican squash and gourd bees of the genus *Peponapis* (Hymenoptera: Apoidea). *Annals of the Entomological Society of America*, 59(6), 835–851.
- Hurd, P., y Linsley, E. (1967). South American squash and gourd bees of the genus *Peponapis* (Hymenoptera: Apoidea). *Annals of the Entomological Society of America*, 60(3), 647–661.
- Hurd, P., Linsley, E., y Whitaker, T. W. (1971). Squash and gourd bees (*Peponapis*, *Xenoglossa*) and the origin of the cultivated *Cucurbita*. *Evolution*, 25(1), 218–234.
- Huson, D. H., y Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, 23(2), 254–267. <https://doi.org/10.1093/molbev/msj030>
- Ignart, F., y Weeden, N. (1984). Allozyme variation in cultivars of *Cucurbita pepo* L. *Euphytica*, 33, 779–785.

- Jarvis, D. I., Padoch, C., y H.D., C. (2007). Biodiversity, Agriculture and Ecosystem Services. In C. H. Jarvis DI, Padoch C (Ed.), *Managing Biodiversity in Agricultural Ecosystems* (pp. 1–12). New York: Bioersivity International. Columbia University Press.
- Jombart, A. T., Kamvar, Z. N., Collins, C., Lustrik, R., Beugin, P., Knaus, B. J., ... Jombart, M. T. (2018). Package ‘adegenet’ R topics documented :
- Jones, M. B., Highlander, S. K., Anderson, E. L., Li, W., Dayrit, M., Klitgord, N., ... Venter, J. C. (2015). Library preparation methodology can influence genomic and functional predictions in human microbiome research. *Proceedings of the National Academy of Sciences*, 112(45), 14024–14029. <https://doi.org/10.1073/pnas.1519288112>
- Kareiva, P., Watts, S., McDonald, R., Boucher, T. (2007). Domesticated nature: Shaping landscapes and ecosystems for human welfare. *Science*, 316(5833), 1866–1869.
- Kates, H. R., Soltis, P. S., y Soltis, D. E. (2017). Evolutionary and domestication history of *Cucurbita* (pumpkin and squash) species inferred from 44 nuclear loci. *Molecular Phylogenetics and Evolution*, 111, 98–109. <https://doi.org/10.1016/j.ympev.2017.03.002>
- Katzir, N., Tadmor, Y., Tzuri, G., Leshzeshen, E., Mozes-Daube, N., Danin-Poleg, Y., y Paris, H. (2002). Further ISSR and preliminary SSR analysis of relationships among accessions of *Cucurbita pepo*. In Nurit Katzir y H. S. Paris (Eds.), *Proc Cucurbitaceae 2000*, Acta Hort 510, Israel.
- Kay, M., King, F. B., y Robinson, C. K. (1980). Cucurbits from Phillips Spring: new evidence and interpretations. *American Archaeology*, 45(4), 806–822.
- King, F. B. (1985). *Prehistoric Food Production in Eastern North America* (R. I. Ford, Ed.). useum of Anthropology, University of Michigan, Ann Arbor, MI.
- Kirkpatrick, K. J., y Wilson, H. D. (1988). Interspecific Gene Flow in *Cucurbita*: *C. texana* vs. *C. pepo*. *American Journal of Botany*, 75(4), 519–527.
- Kucheryavskiy, S. (2017). mdatools: Multivariate Data Analysis for Chemometrics,. R Package Version 0.8. 2.
- Labate, J. A., Robertson, L. D., y Baldo, A. M. (2009). Multilocus sequence data reveal extensive departures from equilibrium in domesticated tomato (*Solanum lycopersicum* L.). *Heredity*, 103(3), 257–267. <https://doi.org/10.1038/hdy.2009.58>
- Lam, H. M., Xu, X., Liu, X., Chen, W., Yang, G., Wong, F. L., ... Zhang, G. (2010). Resequencing of 31 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection. *Nature Genetics*, 42(12), 1053–1059. <https://doi.org/10.1038/ng.715>
- Laughlin, K. D., Power, A. G., Snow, A. A., y Spencer, L. J. (2009). Risk assessment of genetically

- engineered crops: Fitness effects of virus-resistance transgenes in wild *Cucurbita pepo*. *Ecological Applications*, 19(5), 1091–1101. <https://doi.org/10.1890/08-0105.1>
- Lira-Saade, R. (1995). Estudios taxonómicos y ecogeográficos de las cucurbitaceae latinoamericanas de importancia económica. Roma, Italia.: International Plant Genetic Resources Institute.
- Lira-Saade, R., y Montes-Hernandez, S. (1994). Cucurbits (*Cucurbita* spp.). *Neglected Crops*, 1492, 63–77.
- Lobo, A. M. (2008). Importancia de los recursos genéticos de la agrobiodiversidad en el desarrollo de producciones sostenibles. *Revista Corpoica*, 26(1), 5–7. <https://doi.org/10.4067/S0718-34292008000100001>
- López de Heredia, U. (2016). Las técnicas de secuenciación masiva en el estudio de la diversidad biológica. *Munibe Ciencias Naturales*, 64. <https://doi.org/10.21630/mcn.2016.64.07>
- Lozada-Aranda, M., Rojas-Barrer, I., Mastretta-Yanes, A., Ponce-Mendoza, A., Burgeff, C., Orjuela-R, M. A., y Oliveros, O. (2017). Las milpas de México. *Oikos*, 17, 10–12.
- Luo, M. C., Yang, Z. L., You, F. M., Kawahara, T., Waines, J. G., y Dvorak, J. (2007). The structure of wild and domesticated emmer wheat populations, gene flow between them, and the site of emmer domestication. *Theoretical and Applied Genetics*, 114(6), 947–959. <https://doi.org/10.1007/s00122-006-0474-0>
- Mamidi, S., Rossi, M., Annam, D., Moghaddam, S., Lee, R., Papa, R., y McClean, P. (2011). Investigation of the domestication of common bean (*Phaseolus vulgaris*) using multilocus sequence data. *Functional Plant Biology*, 38(12), 953. <https://doi.org/10.1071/fp11124>
- Mardis, E. R. (2011). A decade's perspective on DNA sequencing technology. *Nature*, 470, 198–203.
- Marsjan, P. A., y Oldenbroek, J. K. (2006). Molecular markers, a tool for exploring genetic diversity (Section C in part 4). *State of the Art in the Management of Animal Genetic Resources*, 359–379.
- Meyer, R. S., y Purugganan, M. D. (2013). Evolution of crop species: Genetics of domestication and diversification. *Nature Reviews Genetics*, 14(12), 840–852. <https://doi.org/10.1038/nrg3605>
- Mitra, R. D., y G.M. Church. (1999). In situ localized amplification and contact replication of many individual DNA molecules. *Nucleic Acids Research*, 27, e34–e39.
- Montero-Pau, J., Blanca, J., Bombarely, A., Ziarolo, P., Esteras, C., Martí-Gómez, C., ... Cañizares, J. (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(6), 1161–1171. <https://doi.org/10.1111/pbi.12860>

- Montes, H. S. (2002). Flujo génico en calabaza (*Cucurbita* spp.) dentro del sistema milpa en el occidente de México. Universidad Nacional Autónoma de México.
- Morrell, P. L., Gonzales, A. M., Meyer, K. K. T., y Clegg, M. T. (2014). Resequencing data indicate a modest effect of domestication on diversity in barley: A cultigen with multiple origins. *Journal of Heredity*, 105(2), 253–264. <https://doi.org/10.1093/jhered/est083>
- Müller, B. S. F., Pappas, G. J., Valdisser, P. A. M. R., Coelho, G. R. C., de Menezes, I. P. P., Abreu, A. G., ... Vianello, R. P. (2015). An operational SNP panel integrated to SSR marker for the assessment of genetic diversity and population structure of the common bean. *Plant Molecular Biology Reporter*, 33(6), 1697–1711. <https://doi.org/10.1007/s11105-015-0866-x>
- Nee, M. (1990). The domestication of *cucurbita* (Cucurbitaceae). *Economic Botany*, 44(3 Supplement), 56–68. <https://doi.org/10.1007/BF02860475>
- Nesom, G. L. (2011). Toward Consistency of Taxonomic Rank in Wild/Domesticated Cucurbitaceae. *Phytoneuron*, 13(March), 1–33. Retrieved from <http://www.phytoneuron.net/PhytoN-CucurbConsistency.pdf>
- Ntuli, N. R., Tongoona, P. B., y Zobolo, A. M. (2015). Genetic diversity in *Cucurbita pepo* landraces revealed by RAPD and SSR markers. *Scientia Horticulturae*, 189, 192–200. <https://doi.org/10.1016/j.scienta.2015.03.020>
- OCDE, M. (2012). La Educación Superior en Colombia.
- Organization for Economic Cooperation and Development. (2012). Consensus Document on the Biology of *Cucurbita* L. (Squashes, Pumpkins, Zucchini and Gourds). Series on Harmonisation of Regulatory Oversight in Biotechnology, 53.
- Ott, A., Liu, S., Schnable, J. C., Yeh, C. T., Wang, K. S., y Schnable, P. S. (2017). tGBS® genotyping-by-sequencing enables reliable genotyping of heterozygous loci. *Nucleic Acids Research*, 45(21). <https://doi.org/10.1093/nar/gkx853>
- Paredes-Torres, L. M. (2016). Filogenia molecular del género *Cucurbita* L. (Cucurbitaceae) usando secuencias de cloroplasto. Tesis de Licenciatura. Facultad de Estudios Superiores Iztacala, Universidad Nacional Autónoma de México.
- Paris, H. (2001). History of the cultivar-groups of *Cucurbita pepo*. *Horticultural Reviews-Westport Then New York*, 25, 71–170. <https://doi.org/10.1002/9780470650783.ch2>
- Paris, H. (2018). Consumer-oriented exploitation and conservation of genetic resources of pumpkins and squash, *Cucurbita*. *Israel Journal of Plant Sciences*, 65(3–4), 202–221.
- Paris, H., Yonash, N., Portnoy, V., Mozes-Daube, N., Tzuri, G., y Katzir, N. (2003). Assessment of

- genetic relationships in *Cucurbita pepo* (Cucurbitaceae) using DNA markers. *Theoretical and Applied Genetics*, 106(6), 971–978. <https://doi.org/10.1007/s00122-002-1157-0>
- Paris, H., Portnoy, V., Mozes-Daube, N., Tzuri, G., Katzir, N., y Yonash, N. (2004). AFLP, ISSR, and SSR polymorphisms are in accordance with botanical and cultivated plant taxonomies of the highly polymorphic *Cucurbita pepo*. *Acta Horticulturae*, 634, 167–173. <https://doi.org/10.17660/ActaHortic.2004.634.20>
- Paris, H. (2000). Paintings (1769-1774) by A. N. Duchesne and the history of *Cucurbita pepo*. *Annals of Botany*, 85(6), 815–830. <https://doi.org/10.1006/anbo.2000.1147>
- Paris, H. (2016a). Germplasm enhancement of *Cucurbita pepo* (pumpkin, squash, gourd: Cucurbitaceae): progress and challenges. *Euphytica*, 208(3), 415–438. <https://doi.org/10.1007/s10681-015-1605-y>
- Paris, H. (2016b). Germplasm enhancement of *Cucurbita pepo* (pumpkin, squash, gourd: Cucurbitaceae): progress and challenges. *Euphytica*, Vol. 208, pp. 415–438. <https://doi.org/10.1007/s10681-015-1605-y>
- Paris, H., Lebeda, A., Křistkova, E., Andres, T. C., y Nee, M. H. (2012). Parallel Evolution Under Domestication and Phenotypic Differentiation of the Cultivated Subspecies of *Cucurbita pepo* (Cucurbitaceae). *Economic Botany*, Vol. 66, pp. 71–90. <https://doi.org/10.1007/s12231-012-9186-3>
- Paris, H. (1989). Historical Records, Origins, and Development of the Edible Cultivar Groups of *Cucurbita pepo* (Cucurbitaceae). *Economic Botany*, 43(4), 423–443.
- Paris, H., y Nerson, H. (2003). Seed dimensions in the subspecies and cultivar-groups of *Cucurbita pepo*. *Genetic Resources and Crop Evolution*, 50(6), 615–625. <https://doi.org/10.1023/A:1024464831595>
- Pickrell, J. K., y Pritchard, J. K. (2012). User Manual for TreeMix v1 . 0. 0–7.
- Piperno, D. R., y Pearsall, D. M. (1998). The origins of agriculture in the lowland Neotropics.
- Piperno, D. R., y Stothert, K. E. (2003). Phytolith evidence for early Holocene *Cucurbita* domestication in southwest Ecuador. *Science*, 299(5609), 1054–1057.
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., ... Sham, P. C. (2007). PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *The American Journal of Human Genetics*, 81(3), 559–575. <https://doi.org/10.1086/519795>
- Purugganan, M. D., y Fuller, D. Q. (2009). The nature of selection during plant domestication. *Nature*, 457(7231), 843–848. <https://doi.org/10.1038/nature07895>

- Quail, M. A., Smith, M., Coupland, P., Otto, T. D., Harris, S. R., Connor, T., y Al, E. (2012). A tale of three next generation sequencing platforms: comparison of Ion torrent, pacific biosciences and illumina MiSeq sequencers. *BMC Genomics*, 13(1), 341. <https://doi.org/10.1186/1471-2164-13-341>
- Quesada, M., Winsor, J. A., y Stephenson, A. G. (1996). Effects of pollen selection on progeny vigor in a *Cucurbita pepo* x *C. texana* hybrid. *Theoretical and Applied Genetics*, 92(7), 885–890. <https://doi.org/10.1007/BF00221902>
- Robinson, R. W., y Shail, J. W. (1987). Genetic variability for compatibility of an interspecific cross. *Cucurbit Genetics Cooperative Report*, 10, 88–89.
- Rodriguez, J. M., Berke, T., Engle, L., y Nienhuis, J. (1999). Variation among and within *Capsicum* species revealed by RAPD markers. *Theoretical and Applied Genetics*, 99(1–2), 147–156. <https://doi.org/10.1007/s001220051219>
- Rodriguez, M., Rau, D., Bitocchi, E., Bellucci, E., Biagetti, E., Carboni, A., ... Attene, G. (2016). Landscape genetics , adaptive diversity and population structure in *Phaseolus vulgaris*. 1781–1794.
- Ruiz Mondragón, K. Y. (2017). Filogeografía y diversidad genética en *Cucurbita pepo* L. 1753 en México. Tesis de Licenciatura. Facultad de Ciencias, Universidad Nacional Autónoma de México.
- Sánchez de la Vega, G. (2017). De la cueva a la mesa, y ahora al laboratorio genómico: la diversidad de calabazas de México. *Oikos*, 17, 11–17.
- Sang, T., y Ge, S. (2007). The puzzle of rice domestication. *Journal of Integrative Plant Biology*, 49, 760–768.
- Sanjur, O. I., Piperno, D. R., Andres, T. C., y 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*, 99(1), 535–540. <https://doi.org/10.1073/pnas.012577299>
- Sauer, J. D. (1950). The grain amaranths: a survey of their history and classification. *Annals of the Missouri Botanical Garden*, 37(4), 561–632.
- Schlötterer, C. (2000). Evolutionary dynamics of microsatellite DNA. *Chromosoma*, 109, 365–371.
- Schlötterer, Christian. (2004). The evolution of molecular markers. *Nature Reviews Genetics*, 5(January), 63–69.
- Scott, G. (1934). Observations on some inbred lines of bush types of *Cucurbita pepo*. In *Proc. Amer. Soc. Hort. Sci*, 32, 408.

- Shand, H. (1997). Human Nature: Agricultural Biodiversity and Farm-Based Food Security. In RAFI.
- Smith, B. D. (2006). Eastern North America as an independent center of plant domestication. *Proceedings of the National Academy of Sciences*, 103(33), 12223–12228. <https://doi.org/10.1073/pnas.0604335103>
- Smith, Bruce D. (1997). The initial domestication of *Cucurbita pepo* in the Americas 10,000 years ago. *Science*, 276(5314), 932–934.
- Smith, Bruce D., y Zeder, M. A. (2013). The onset of the Anthropocene. *Anthropocene*, 4, 8–13. <https://doi.org/10.1016/j.ancene.2013.05.001>
- Song, Z. P., Xu, X., Wang, B., Chen, J. K., y Lu, B. R. (2003). Genetic diversity in the northernmost *Oryza rufipogon* populations estimated by SSR markers. *Theoretical and Applied Genetics*, 107, 1492–1499. <https://doi.org/10.1007/s00122-003-1380-3>
- Song, Z., Li, B., Chen, J., y Lu, B. (2005). Genetic diversity and conservation of common wild rice (*Oryza rufipogon*) in China. *Plant Species Biology*, 83–92.
- Sweeney, M. McCouch, S. (2007). The complex history of the domestication of rice. *Annals of Botany*, 100, 951–957.
- Taitano, N., Bernau, V., Jardón-Barbolla, L., Leckie, B., Mazourek, M., Mercer, K., ... van der Knaap, E. (2018). Genome-wide genotyping of a novel Mexican Chile Pepper collection illuminates the history of landrace differentiation after *Capsicum annuum* L. domestication. *Evolutionary Applications*, 12(1), 78–92. <https://doi.org/10.1111/eva.12651>
- Teppner, H. (2000). *Cucurbita pepo* (Cucurbitaceae)-history, seed coat types, thin coated seeds and their genetics. *Phyton (Horn)*, 40(1), 1–42.
- Torres Ruiz, R. A., y Hemleben, V. (1991). Use of ribosomal DNA spacer probes to distinguish cultivars of *Cucurbita pepo* L. and other Cucurbitaceae. *Euphytica*, 53(1), 11–17. <https://doi.org/10.1007/BF00032026>
- Velasco Gutiérrez, K., y Juárez Sierra, J. (2009). Etnobotánica del género *Cucurbita* en dos localidades mixtecas de Oaxaca, México. *Etnobiología*, 7(1), 63–85. Retrieved from <https://dialnet.unirioja.es/servlet/articulo?codigo=5294428>
- Verde, I., Abbott, A. G., Scalabrin, S., Jung, S., Shu, S., Marroni, F., ... Rokhsar, D. S. (2013). The high-quality draft genome of peach (*Prunus persica*) identifies unique patterns of genetic diversity, domestication and genome evolution. *Nature Genetics*, 45(5), 487–494. <https://doi.org/10.1038/ng.2586>
- Verdone, M., Rao, R., Coppola, M., y Corrado, G. (2018). Identification of zucchini varieties in commercial food products by DNA typing. *Food Control*, 84, 197–204.

<https://doi.org/10.1016/j.foodcont.2017.07.039>

- Villanueva Verduzco, C. (2007). Calabazas cultivadas. Identificación de especies, caracterización y descripción varietal. Universidad Autónoma Chapingo, 123.
- Vos, P., Hogers, R., Bleeker, M., Reijans, M., Lee, T. V. D., Hornes, M., et al. (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research*, 23, 4407–4414.
- Votava, E. J., Baral, J. B., y Bosland, P. W. (2006). Genetic Diversity of Chile (*Capsicum Annuum* var. *annuum* L.) Landraces from Northern New Mexico, Colorado, and Mexico. *Economic Botany*, 59(1), 8–17. [https://doi.org/10.1663/0013-0001\(2005\)059\[0008:gdocca\]2.0.co;2](https://doi.org/10.1663/0013-0001(2005)059[0008:gdocca]2.0.co;2)
- Waples, R. S., y Gaggiotti, O. (2006). What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology*, 15(6), 1419–1439. <https://doi.org/10.1111/j.1365-294X.2006.02890.x>
- Wells, J. C. K., y Wells, J. C. K. (2016). The Emergence of Agriculture. In *The Metabolic Ghetto* (Vol. 44, pp. 255–274). <https://doi.org/10.1017/cbo9780511972959.014>
- Whitaker, T. W., y Carter, F. G. (1946). Critical notes on the origin and domestication of the cultivated species of *Cucurbita*. *American Journal of Botany*, 33(1), 10–15.
- Whitaker, Thomas W., y Cutler, H. C. (1965). Cucurbits and cultures in the Americas. *Economic Botany*, 19(4), 344–349. <https://doi.org/10.1007/BF02904804>
- Whitaker, Thomas W., y Cutler, H. C. (1971). Pre-historic cucurbits from the Valley of Oaxaca. *Economic Botany*, 25(2), 123–127. <https://doi.org/10.1007/BF02860073>
- Whitaker, Thomas W., y Cutler, H. C. (1986). Cucurbits from preceramic levels at Guila Naquitz. Academic Press, Incorporated.
- Whitaker, Thomas W., y Davis, G. (1962). Cucurbits. Cucurbits.
- Wilson, H. D., Doebley, J., y Duvall, M. (1992). Chloroplast DNA diversity among wild and cultivated members of *Cucurbita* (Cucurbitaceae). *Theoretical and Applied Genetics*, 84(7–8), 859–865. <https://doi.org/10.1007/BF00227397>
- Würschum, T., Langer, S. M., Longin, C. F. H., Korzun, V., Akhunov, E., Ebmeyer, E., ... Reif, J. C. (2013). Population structure, genetic diversity and linkage disequilibrium in elite winter wheat assessed with SNP and SSR markers. *Theoretical and Applied Genetics*, 126(6), 1477–1486. <https://doi.org/10.1007/s00122-013-2065-1>
- Zabeau, M., y Vos, P. (1993). Selective restriction fragment amplification: a general method for DNA fingerprinting. European Patent, 0 534858 A1.
- Zeder, M. A. (2006). Central questions in the domestication of plants and animals. *Evolutionary Anthropology*, 15(3), 105–117. <https://doi.org/10.1002/evan.20101>

- Zeder, M. A. (2015). Core questions in domestication research. *Proceedings of the National Academy of Sciences*, 112(11), 3191–3198. <https://doi.org/10.1073/pnas.1501711112>
- Zeder, M. A., Emshwiller, E., Smith, B. D., y Bradley, D. G. (2006). Documenting domestication: The intersection of genetics and archaeology. *Trends in Genetics*, 22(3), 139–155. <https://doi.org/10.1016/j.tig.2006.01.007>
- Zizumbo-Villarreal, D., Colunga-GarcíaMarín, P., Payró De La Cruz, E., Delgado-Valerio, P., y Gepts, P. (2005). Population structure and evolutionary dynamics of wild-weedy-domesticated complexes of common bean in a Mesoamerican region. *Crop Science*, 45(3), 1073–1083. <https://doi.org/10.2135/cropsci2004.0304>
- Zizumbo-Villarreal, D., y Colunga, P. (1993). Tecnología agrícola tradicional, conservación de recursos naturales y desarrollo sustentable. In E. Leff y J. Carabias (coords.). *Cultura y manejo sustentable de los recursos naturales*. Vol I. (pp. 165–202). <https://doi.org/10.13140/RG.2.1.3090.8962>
- Zizumbo Villarreal, D., Colunga García-Marín, P., Pat May, F., Martíne Castillo, J., y Mijangos Cortés, J. O. (2010). Recursos fito-genéticos para la alimentación y la agricultura. *Biodiversidad y Desarrollo Humano En Yucatán*, 334–339. Retrieved from <http://www.cicy.mx/Documentos/CICY/Sitios/Biodiversidad/pdfs/Cap7/01 Recursos fitogeneticos.pdf>
- Zraidi, A., Stift, G., Pachner, M., Shojaeiyan, A., Gong, L., y Lelley, T. (2007). A consensus map for *Cucurbita pepo*. *Molecular Breeding*, 20(4), 375–388. <https://doi.org/10.1007/s11032-007-9098-6>