



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Maestría y Doctorado en Ciencias Bioquímicas

Evolución reciente de razas nativas de maíz (*Zea mays* L.) y sus parientes silvestres: Flujo genético con variedades modernas

TESIS

QUE PARA OPTAR POR EL GRADO DE:

Doctora en Ciencias

PRESENTA:

Idalia Claudia Rojas Barrera

TUTOR PRINCIPAL

Daniel Ignacio Piñero Dalmau

Instituto de Ecología, UNAM

MIEMBROS DEL COMITÉ TUTOR

Alejandro Ángel García Rubio Granados  
Instituto de Biotecnología, UNAM

Luis David Alcaraz Peraza  
Facultad de Ciencias, UNAM

Ciudad de México. Febrero, 2020



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

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

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

CGEP/PMDCB/2306/2019

**Rojas Barrera Idalia Claudia**  
**Estudiante de Doctorado en Ciencias Bioquímicas**  
**Presente**

Los miembros del Subcomité Académico en reunión ordinaria del día 07 de octubre del presente año, conocieron su solicitud de asignación de **JURADO DE EXAMEN** para optar por el grado de **DOCTORADO EN CIENCIAS**, con la réplica de la tesis **"Evolución reciente de las razas nativas de maíz (Zea mays L.) y sus parientes silvestres: Flujo genético con variedades mejoradas"**, dirigida por el/la Dr(a). **Piñero Dalmau Daniel**.

De su análisis se acordó nombrar el siguiente jurado integrado por los doctores:

PRESIDENTE	Eguiarte Fruns Luis Enrique
VOCAL	Orozco Ramirez Quetzalcoatl
VOCAL	Jaramillo Correa Juan Pablo
VOCAL	Covarrubias Robles Alejandra Alicia
SECRETARIO	Nieto Sotelo Jorge

Sin otro particular por el momento, aprovecho la ocasión para enviarle un cordial saludo.

**Atentamente**  
**"POR MI RAZA, HABLARÁ EL ESPÍRITU"**  
**Cd. Universitaria, Cd. Mx., a 07 de octubre de 2019.**



**Dra. ANA BRÍGIDA CLORINDA ARIAS ÁLVAREZ**  
**COORDINADORA**

---

Esta tesis se realizó bajo la tutoría del Dr. Daniel Piñero en el Laboratorio de Genética y Ecología del Instituto de Ecología de la Universidad Nacional Autónoma de México. El financiamiento fue otorgado por la Secretaría de Medio Ambiente y Recursos Naturales (SEMARNAT) y la Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO), a través del proyecto “Monitoreo de las razas criollas y de los linajes geográficos del maíz en México usando un enfoque genómico: una estrategia para su uso sustentable y su conservación.”

Idalia Claudia Rojas Barrera fue alumna del Programa de Maestría y Doctorado en Ciencias Bioquímicas de la Universidad Nacional Autónoma de México y fue apoyada por una beca de posgrado CONACYT número 376788, por la beca mixta CONACYT para realizar una estancia de investigación en el Departamento de Botánica de la Universidad de la Columbia Británica, Vancouver, Canadá; y por el Programa de Apoyo a los Estudios del Posgrado (PAEP) para asistir al curso internacional “Temas selectos de domesticación y recursos genéticos” en la Universidad de La Molina, Lima, Perú; y presentar los resultados de la investigación en el Congreso Internacional de Evolución, en Montpellier, Francia.

---

## Agradecimientos académicos

Al Comité Tutorial que asesoró el desarrollo de esta tesis conformado por:

Dr. Alejandro Ángel García Rubio Granados	IBT, UNAM
Dr. Luis David Alcaraz Peraza (2017-2018)	Facultad de Ciencias, UNAM
Dr. León Patricio Martínez-Castilla (2014-2017)	Facultad de Química, UNAM

A la Dra. Ana Laura Wegier Briuolo del Instituto de Biología de la UNAM por participar como tutora invitada.

Al Dr. José de Jesús Sánchez González del Centro Universitario de Ciencias Biológicas y Agropecuarias de la Universidad de Guadalajara, por aportar su valioso conocimiento para el planteamiento y desarrollo de este proyecto.

Al Dr. Juan Manuel Hernández Casillas del Instituto Nacional de Investigaciones Forestales Agrícolas y Pecuarias (INIFAP), por realizar la clasificación racial de las variedades tradicionales de maíz.

Al Dr. Rafael Ortega Paczka de la Universidad Autónoma de Chapingo por recibirme en su oficina para aclarar mis dudas sobre la historia del mejoramiento del maíz en México.

Al Dr. Loren H. Rieseberg del Departamento de Botánica de la *University of British Columbia (UBC)*, por ser mi anfitrión y asesor durante la estancia de investigación realizada en su laboratorio, y por su ayuda en la redacción del manuscrito.

Al Dr. Gregory L. Owens por asesorarme para el llamado de variantes genéticas y los análisis de flujo genético, durante mi estancia en el laboratorio del Dr. Loren H. Rieseberg.

Al Dr. Mauricio Rafael Bellon Corrales por su revisión y críticas a la versión preliminar del manuscrito basado en este trabajo.

A la Dra. Nadia Santini por supervisar la primera versión del manuscrito.

A la Dra. Tania Garrido Garduño por la asesoría técnica prestada en el laboratorio.

A Ernesto Campos Murillo por la asesoría para realizar los análisis bioinformáticos en el clúster de la CONABIO.

A los Doctores Alejandra Covarrubias, Luis Eguiarte, Quetzalcóatl Orozco, Jorge Nieto y Juan Pablo Jaramillo por revisar esta tesis.

---

## Agradecimientos personales

Este proyecto está integrado por un número inmenso de contribuciones a nivel profesional y personal. La persona que soy en este momento es la suma de todas ellas.

Advierto que el texto que sigue es largo, si lo escribo es porque a mi misma me encanta leer los agradecimientos personales, y porque no pienso volver a escribir otra tesis. Creo que estos párrafos nos develan el proceso personal y transformador que tienen los trabajos de investigación. Una persona nunca es la misma después de hacer una investigación, mucho menos una de doctorado.

Empiezo por aclarar que soy biomédica, durante los inicios de mi vida académica imaginaba que haría el doctorado describiendo una vía de señalización en cáncer, no en genética de poblaciones de maíz en un laboratorio que estudia de todo, menos humanos. No me sorprende haber llegado a este punto, lo mío no son las trayectorias lineales, pues al terminar la maestría pensé que ese era el fin de mi vida escolar. El ambiente académico es duro en su esencia misma. Al final de cuentas la ciencia se sostiene sobre un proceso de colonizaciones interminables, en los que una idea o teoría nuevas se construyen sobre los cimientos de otra o en el caso más extremo terminan demoliendola. Y aunque eso sólo debería suceder con las propuestas y teorías, tengo la impresión de que de algunos lo extienden a las personas.

Este proyecto de doctorado nació gracias a la invitación que **Cristina Aguilar** me hizo para colaborar en un proyecto financiado por la iniciativa gubernamental *Masagro* y liderado por el Dr. **Jorge Nieto**. A pesar de lo poco que yo sabía de plantas, el Dr. Nieto me acogió en su grupo y se convirtió en mi mentor durante esa etapa. Fue él quien me introdujo al mundo del maíz y a mis primeros trabajos de campo. Recuerdo sus explicaciones sobre el proceso de apertura y cierre de estomas bajo el sol de mediodía, en los campos experimentales del CIMMYT. Fue así como después de un año y medio en su grupo, me di cuenta de que la investigación es una de mis actividades favoritas, y decidí regresar a la escuela. Aunque me encariñé con el maíces mejorados y los agaves, no quería trabajar con variedades mejoradas, deseaba darle un enfoque evolutivo a mi trabajo y trabajar con maíces "criollos".

Entonces, llegué con el Dr. **Daniel Piñero**, quien ha sido el mejor tutor que hubiera podido imaginar. Me aceptó con toda mi "carga" así me lo dijo él, sobra decir que yo también lo acepté con la suya. Aunque, yo salí ganando, cualquier expresión de agradecimiento y admiración es poco para transmitir mi sentir. Desde que conozco a Daniel me ha dado el apoyo y la libertad para plantear, realizar y concluir este proyecto. Gracias por todo.

Como mencioné antes, este trabajo se compone de innumerables contribuciones. La Dra. **Ana Wegier**, con el arrojo que la caracteriza, me ayudó a plantear el diseño experimental para evaluar el flujo genético de las variedades mejoradas hacia los maíces criollos y sus parientes silvestres. Y aunque el escrutinio para ingresar al doctorado no estuvo libre de la pregunta incómoda ¿Qué hace una biomédica estudiando maíces? Mi comité me dio el voto de confianza.

Este trabajo no hubiera arrancado de no ser por la participación del Dr. **José de Jesús Sánchez**, quién ha dedicado su carrera al estudio del género *Zea*; su calidad científica se refleja en el rigor de sus registros de colecta, en la colección de germoplasma con la que ha contribuido a este país y en la generosidad

---

enorme para compartir su conocimiento.

Mi llegada al laboratorio de Genética y Ecología resultó intimidante por estar en un ambiente en donde todo mundo parecía saber lo que hacía, menos yo que soy biomédica. Sin embargo, mis colegas hicieron del laboratorio un lugar increíble para la vida académica y en donde cultivé nuevas amistades. Mis compañeras de generación, después de un tiempo se volvieron mis hermanas y me guiaron durante todo el proceso. Gracias a **Laura** quién me enseñó a hacer mi primer mapa, yo me perdía hasta con google maps, y hasta la fecha necesito ver el punto azul que se desplaza conmigo. Gracias a **Myriam** por enseñarme a planear mis salidas a campo, pues no tenía ni la menor idea de como hacerlo. Y gracias a **Azalea**, a quien le debo tanto que no acabaría, pero empiezo por agradecerle por ser mi maestra de campo, por discutir conmigo mis ideas y proyectos, y hasta por adoptarme en su casa.

Poco a poco las filas del grupo de Daniel se renovaron, así llegaron **Nancy, Raquel, Verito, Christian y Ruth**, no necesariamente en ese orden. Gracias **Ruth** por nuestras esporádicas carreras vespertinas y por nuestras largas charlas nocturnas. Gracias **Verónica** por ser un ejemplo de tenacidad, por tu honestidad y amistad. Gracias a **Chris y Oso** por ser los compañeros de cubículo más *cool ever*.

En ese ambiente también estaban los fantásticos y queridos hermanos Vázquez. Gracias a **Tania, Alex, Marco, Carlos, Madisson, Susette, Oscar y Gaby**. Y por supuesto, a la líder este talentoso clan la **Dra. Ella Vázquez**. Gracias a **JC, Erick, Mariana Rojas, Mariana Hernández, Jorge, Miguel** y a todos los miembros del laboratorio con los que compartí durante mi estancia.

El trabajo de campo, recolectando maíces fue la etapa más divertida y quizá también la más dura. Por una parte, está el salir a campo en un país donde no hay garantías de que podamos volver sanos y salvos; y por otra el estrellarse con la dura realidad del campo mexicano, que parecía sacada de un cuento de Juan Rulfo. Donde el desánimo y a veces el resentimiento, por el abandono del campo mexicano, se cuelan en las frases cortas que expresan los y las campesinas. A pesar de esto, ellos y ellas continúan con este proceso que mantiene la herencia cultural y biológica que son nuestras especies y variedades domesticadas. Conservo la esperanza de que con la suma de que nuestras acciones como academic@s y sociedad civil, podamos contribuir a un futuro más alentador. Esta etapa requirió de muchas manos y largas horas en carretera, gracias a **Myriam, Alicia, Azalea, Daniel, Diana, Verónica, Alfredo, Leopoldo y Alejandro**. Pero, sobre todo gracias a **los agricultores** que me hicieron creer en este proyecto.

Después del campo vino el laboratorio y el análisis de datos. ¡Oh bioinformática! te amo tanto como te odio. Gracias a **Alicia Mastretta** por introducirme al mundo del análisis de datos. Y gracias a **Loren H Rieseberg y Greg Owens**, por ser unos geniecillos generosos, que me permitieron enriquecer mi vida académica en mi paso por la UBC. Gracias a **Mariana Pascual** por ser mi anfitriona, mi revisora de estilo, mi Pepe Grillo de la ciencia.

Aquí hago una pausa para agradecer a los que participaron en la obra de teatro **“Sueño de una noche en la milpa”**, financiada por la Sociedad para el Estudio de la Evolución (SSE), fuimos un rotundo éxito, y sino me divertí mucho. Es la clase de proyectos en los que me embarqué por el amor a la ciencia, y porque tenía miedo de someter el trabajo de investigación doctoral al escrutinio de la comunidad internacional. Gracias a **Raquel y Aza** por ser mis hermanas



---

frijol y maíz, respectivamente, a **Christian** por ser el Sr. Escarabajo, a **Oscar** por ser el dulce Huitzilin, a **Leopoldo** por ser Chava, a **Susana** por ser nuestra mamá, a **Ray** por ser el hombre dos, al querido **Carlos** por diseñar el póster, a **Aketzalli González** por escribir y dirigir el guión de la obra y a **Rodolfo** por ser el *floor manager* de la obra y de mi vida.

En este intermedio llegó **Heriberto Vázquez**, un talentoso y entusiasta agrónomo, quien me dio el honor de ser su mentora y me llevó a explorar el entorno de las variedades de maíces acriolladas. Los campesinos mexicanos son experimentadores incansables, siempre en busca de mejorar sus semillas al margen de un sistema que pocas veces los considera.

Sin embargo, me pasó como a muchos estudiantes de doctorado, se terminó la beca pero no el doctorado. En ese momento me ofrecieron un trabajo-pseudo-postdoc en Alemania. Yo decidí tomarlo, porque ya había mandado el artículo y pensé ¿Cuánto puede demorar el proceso de publicación?, ¡nos tardamos casi un año! Pero también lo hice porque creo que al paso del tiempo, uno se arrepiente más de lo que no hizo. Y aunque nada sucedió como lo esperaba, pues la historia termina conmigo haciendo sola El Camino de Santiago en el Norte de España. Ese proceso que me sirvió para exorcizar ciertos demonios y juntar fondos para terminar el doctorado. Le agradezco a **Andreas** por ser el alemán más amable del mundo y por hacerme un espacio en su casa desde el primer día que lo conocí, te recuerdo con mucho cariño. Gracias a la pandilla latino-alemana **Yaxal, Alexandra, Gaby, Claudio y Jan**.

Voy llegando al final de este proceso, es probable que muchos se queden fuera de estos extensos agradecimientos, me disculpo de antemano. Hacer esta tesis significó mucho en mi vida, durante el doctorado la vida cambia mucho, uno sale con más responsabilidades y nuevas expectativas, siendo un adulto en toda la extensión de la palabra. No me queda más que enlistar a todos los que a pesar de no estar involucrados directamente en este trabajo, forman parte de las historias alternas.

A mis padres **Catalina y Bernardino**, y a mi tía **Adela**, por el apoyo y amor incondicionales.

A mi hermana **Nadia** y su esposo **Rubén**, por su cariño.

A **Inés y Bernardo** cuya sonrisa siempre es suficiente para recargar baterías.

A los miembros del Laboratorio de Fisiología Molecular del Jardín Botánico de la UNAM, que nunca dejó de sentirse como mi casa: **Jorge Nieto, Cris, Sergio, Lucy, Leopoldo, Alejandro, Alberto, Alejandra, Sol y Ray**.

Al grupo SC **Sandy, Maruca, Andy y Micke**; por esta década que llevamos juntos, los adoro amigos.

A **Alfredo** por tu amistad, a **Marco** por tu hermandad, mi cariño por ti es inmenso.

Al grupo de amigos que he ido recolectando a lo largo de la vida, bajo las circunstancias más diversas, y con los que siempre puedo compartir una conversación amena sin importar el tiempo que demoremos en vernos **Favi, Priscilla, Dafné, Lilian, Sandra, Guillermo Arzate, Guillermo Noriega, Eva, Nelly, Helena y Cecilia**.

A **Rodolfo**, los dedos de las manos ya no me alcanzan para contar el tiempo que llevamos juntos, gracias por compartir tu vida conmigo, te amo.

A **todos mis profesores y mentores**.

A la **UNAM** que, a pesar de todos sus problemas los cuales no son menores, es un oasis en este país y que constituye uno de los pilares en la movilidad social.





---

## ÍNDICE

Índice de figuras y tablas.....	1
Lista de abreviaturas.....	2
Resumen.....	3
Abstract.....	5
1. Introducción.....	6
1.1 Centros de origen.....	6
1.2 Mesoamérica: centro de origen del maíz y otras plantas cultivadas.....	7
1.3 Flujo genético en plantas.....	8
1.4 Flujo de transgenes hacia variedades nativas y parientes silvestres.....	9
1.5 Flujo entre domesticados no transgénicos y parientes silvestres.....	11
1.6 Barreras al flujo genético.....	13
1.7 Sistemas de incompatibilidad.....	14
1.8 Métodos para la detección de flujo genético.....	15
1.9 Planteamiento del problema: panorama socioeconómico, político y tecnológico.....	18
1.10 Sustitución de variedades tradicionales y/o flujo genético con variedades modernas.....	21
1.11 Flujo genético de variedades modernas hacia razas de maíz.....	23
1.12 Acervo de germoplasma de maíz en México.....	25
Hipótesis.....	27
Objetivo general.....	27
Objetivos particulares.....	27
2. Materiales y Métodos.....	28
2.1 Muestras y accesiones.....	28
2.2 Extracción de DNA y genotipificación.....	31
2.4 Análisis de ancestría.....	32
2.5 Evaluación del flujo genético (ABBA-BABA).....	33
2.6 Análisis de agrupamiento.....	34
2.7 Divergencia ( $F_{ST}$ ) a lo largo del genoma de RNs, VMs y PSs.....	35
2.8 Diversidad genética.....	35
3 Resultados.....	36
3.1 El conjunto de polimorfismos se distribuye a lo largo de los diez cromosomas de las RNs, VMs y PSs.....	36
3.2 La ancestría de los individuos cambia de acuerdo con la altitud y el periodo de colecta.....	39
3.2 Flujo genético de variedades mejoradas hacia razas nativas y parientes silvestres.....	43
3.3 La divergencia ( $F_{ST}$ ) entre VMs con RNs y PSs decae en el tiempo.....	46
3.4 MV1 y MV2 se agrupan con los acervos genéticos tropicales desarrollados en México y Camerún.....	47
3.5 La diversidad genética de las RNs y PSs es mayor para los genotipos colectados después del 2000.....	50
4. Discusión.....	55
4.1 La introducción de variedades mejoradas ha modificado la estructura poblacional de las razas de maíz.....	55
4.2 Flujo genético e introgresión de variedades modernas hacia razas	

---

nativas y parientes silvestres de maíz.....	57
4.3 La divergencia entre razas y parientes silvestres con VMs disminuye en el tiempo.....	61
4.4 Las CMLs son basales a los grupos VM1 y VM2.....	61
4.5 La diversidad nucleotídica incrementa en las razas y parientes silvestres muestreados después del 2000.....	62
5. Conclusiones.....	65
6. Perspectivas.....	66
7. Referencias.....	68
Anexo 1.....	79
Anexo 2.....	84
Anexo 3.....	87
Anexo 4.....	91

## Índice de figuras y tablas

Figura 1. Distribución de las poblaciones y accesiones de razas (RNs), variedades modernas (VMs) y parientes silvestres (PSs) incluidas en este trabajo.

Figura 2. Densidad de los SNPs en para los 10 cromosomas de maíz.

Figura 3. Desequilibrio de ligamiento ( $r^2$ ) calculado en ventanas de 100 Kb para razas, variedades modernas, *parviglumis*, *mexicana* y *Z. diploperennis*.

Figura 4. Estructura poblacional.

Figura 5. Rearreglo de los genotipos incluidos en el análisis de estructura poblacional.

Figura 6. Correlación de los 7 grupos genéticos ( $K=7$ ) identificados para razas nativas (RNs), variedades modernas (VMs) y parientes silvestres.

Figura 7. Modelos de flujo genético.

Figura 8. Flujo genético para las siete unidades experimentales distribuidas en un gradiente altitudinal.

Figura 9. Divergencia poblacional ( $F_{ST}$ ) a lo largo del genoma.

Figura 10. Análisis de componentes principales.

Figura 11. Análisis de ancestría de los grupos genéticos VM1 y VM2.

Figura 12. Estimados de diversidad genética  $\pi$  y  $\theta_w$ .

Tabla 1. Asignación temporal de las colectas y accesiones.

Tabla 2. Parámetros promedio de profundidad por sitio e individuo

Tabla 3. Diversidad nucleotídica ( $\pi$ ).

Tabla 4. Estimados de theta ( $\theta_w$ ).

Tabla 5. Valor promedio de la D de Tajima

## Lista de abreviaturas

EVC	Error de validación cruzada
CIMMYT	Centro Internacional de Mejoramiento de Maíz y Trigo
CML	Líneas mejoradas del CIMMYT
GTR	Modelo general tiempo reversible (Generalised Time-Reversible)
GBS	Genotipificación por secuenciación (Genotyping By Sequencing)
GM	Genéticamente modificados
INIA	Instituto Nacional de Investigaciones Agrícolas
INIFAP	Instituto Nacional de Investigaciones Forestales y Agrícolas
IITA	Instituto Internacional de Agricultura Tropical
IRA	Instituto de Investigaciones Agronómicas De Camerún
RNs	Razas Nativas
MAF	Frecuencia alélica mínima (Minimum Allele Frequency)
VMs	Variedades modernas
VM1	Variedades modernas grupo 1
VM2	Variedades modernas grupo 2
PCA	Análisis de componentes principales (Principal Component Analysis)
PSs	Parientes Silvestres
PRONASE	Productora Nacional de Semillas
RNs	Razas nativas
SNP	Polimorfismos de un sólo nucleótido (Single Nucleotide Polymorphism)
VTs	Variedades Tradicionales
Zmx	<i>Zea mays ssp. mexicana</i>
Zpr	<i>Zea mays ssp. parviglumis</i>

## Resumen

México es el centro de origen y domesticación del maíz. El uso y la introducción de variedades modernas (VMs) de maíz durante la década de 1960 dio pie a varios cuestionamientos acerca de los posibles efectos adversos del flujo genético de VMs hacia las razas nativas (RNs) y sus parientes silvestres (PSs), conocidos genéricamente como teocintles. Sin embargo, a más de 60 años de la liberación de las primeras VMs de maíz, no hay evidencia genética de las consecuencias de la simpatria entre VMs, RNs y sus PSs. En este trabajo evaluamos los cambios en el genoma de siete RNs, y dos PSs empleando colectas que abarcan un período mayor a 70 años. Los individuos se agruparon en tres períodos de acuerdo con el año en que fueron colectados: 1) El primero incluye aquellas variedades tradicionales colectadas antes de la introducción de VMs (< 1960); 2) el segundo abarca las décadas de 1960-1980 (1960-1980); y 3) el último son las colectas realizadas después del año 2000 (>2000). Para cumplir nuestro objetivo, los genotipos de las accesiones y colectas fueron obtenidos por el método de Genotipificación por Secuenciación (GBS). En total analizamos 316,294 SNPs para 385 muestras: 190 variedades tradicionales de maíz, 47 VMs, 71 teocintles subespecie *mexicana*, 65 teocintles subespecie *parviglumis*, y 12 teocintles perennes, estos últimos fueron empleados como grupo externo. Nuestros resultados sugieren introgresión de las VMs hacia las RNs y las poblaciones de *Zea mays* ssp. *mexicana* muestreadas después del 2000. También observamos una disminución en el índice de divergencia pareada ( $F_{ST}$ ) a través del periodo evaluado entre VMs-RNs (VM-RN1960:  $0.051 \pm 0.074$ ; VM-RN1960-1980:  $0.049 \pm 0.077$ ; VM-RN2000:  $0.036 \pm 0.054$ ), VMs-*Z. m. ssp. mexicana* (VM-*Z. m. ssp. mexicana* 1980:  $0.161 \pm 0.211$ ; VM-*Z. m. ssp. mexicana* 2000:  $0.141 \pm 0.145$ ) y VM-*Zea mays* ssp. *parviglumis* (VM-*Z. m. ssp. parviglumis* 1980:  $0.125 \pm 0.185$  y VM-*Z. m. ssp. parviglumis* 2000:  $0.118 \pm 0.130$ ). Además, la diversidad nucleotídica resultó mayor para las RNs colectadas después del 2000 ( $2.03E-04 \pm 4.3E-04$ ) en comparación con las RN colectadas antes de 1960 ( $1.87E-04 \pm 4.01E-04$ ) y en la década de 1980 ( $1.84E-04 \pm 3.92E-04$ ), como consecuencia probable del flujo de las VMs introducidas en los 90s hacia las RNs. Nuestros hallazgos brindan evidencia de la evolución reciente de los acervos genéticos domesticado y silvestre de *Zea*

*mays*, y concuerdan con trabajos previos que han evaluado el flujo genético de las VMs hacia las RNs de maíz y otros cultivos. Consideramos que este tipo de acercamientos pueden ser empleados como herramienta para proponer y monitorear las estrategias de conservación de los recursos genéticos silvestres y domesticados en regiones con condiciones agrícolas heterogéneas.



## Abstract

Mexico is recognized as the center of origin and domestication of maize. Introduction of modern maize varieties (VMs) into Mexico raised concerns regarding the possible effects of gene flow from VMs into native maize landraces (RNs) and wild relatives (PSs), teosintes. However, after more than 60 years from the introduction of the first VMs, the impact of the sympatry with RNs and PSs has not been explored with genomic data. In this work, we assessed changes in the genomes of seven maize RNs and two PSs subspecies from collections spanning over 70 years. We compared the genotypes obtained by Genotyping-by-Sequencing (GBS) for RNs and PSs, before and after the adoption of VMs, and observed significant introgression from sympatric VMs into both RNs and the PS, *Zea mays* ssp. *mexicana* sampled after the year 2000. We also found a decreased paired divergence index ( $F_{ST}$ ) between VMs-RNs (VM-RN1960:  $0.051 \pm 0.074$ ; VM-RN1960-1980:  $0.049 \pm 0.077$ ; VM-RN2000:  $0.036 \pm 0.054$ ), VMs-*Zea mays* ssp. *mexicana* (VM-Z. m. ssp. *mexicana* 1980:  $0.161 \pm 0.211$ ; VM-Z. m. ssp. *mexicana* 2000:  $0.141 \pm 0.145$ ) and VMs-*Zea mays* ssp. *parviglumis* (VM-Z. m. ssp. *parviglumis* 1980:  $0.125 \pm 0.185$  y VM-Z. m. ssp. *parviglumis* 2000:  $0.118 \pm 0.130$ ), over the same time frame. Moreover, we determined that RN's genetic diversity increased after 2000 ( $2.03E-04 \pm 4.3E-04$ ) in comparison with RNs sampled before 1960 ( $1.87E-04 \pm 4.01E-04$ ) and during 1980 decade ( $1.84E-04 \pm 3.92E-04$ ), probably as a result of gene flow from VMs introduced in the 90s. Such findings allowed identifying ongoing changes in the domesticated and wild maize genetic pools, which concur with previous works that have evaluated short-term gene flow from VMs into RNs in other crops. Our approach represents a useful tool for tracking evolutionary change in wild and domesticated genetic resources, and for developing strategies for their conservation in heterogeneous conditions.

# 1. Introducción

## 1.1 Centros de origen

En 1926 el agrónomo y genetista Nikolái Vavílov propuso en su ensayo “Sobre el origen de plantas cultivadas” que era posible determinar su centro de origen analizando los patrones de distribución de la variación genética. Aunque, a la fecha es aceptado que un centro de origen puede diferir del centro de diversificación, los ocho centros de origen propuestos por Vavilov sentaron las bases para su identificación; estas regiones incluyen: 1. China, 2. India, 2a. Región Indo-Malaya, 3. Asia Central (Pakistán, Punjab, Kashmir, Afganistán y Turquestán) 4. Cercano Oriente, 5. Mediterráneo, 6. Etiopía, 7. Sur de México y Centroamérica, y 8. Sudamérica (8. Ecuador, Perú, Bolivia), 8a. Chile; 8b. Brasil-Paraguay (CONABIO; Harlan 1971).

A pesar del reconocimiento de estas regiones, la definición de centros de origen se ha ampliado y contraído, como fue el caso de los “microcentros primarios de genes” y los “megacentros secundarios de genes” propuestos por P. M. Zhukovsky (1970) ( en Harris 1990). En otros casos el concepto de centro se ha redefinido; el ejemplo más memorable es el artículo publicado por Harlan en 1971, que introdujo los conceptos “centros” y “no-centros”; donde propuso que la agricultura se originó de manera independiente en tres regiones, cada una dividida en un centro y un no-centro. En estas, las actividades asociadas a la domesticación alcanzan un rango de 5-10 mil km. De acuerdo con la propuesta de Harlan (1971), un centro es aquella región con límites definibles que cuenta con suficiente evidencia arqueológica, botánica y lingüística para permitir establecer el origen de uno o más cultivos; mientras que un no-centro es una región extensa, con límites difusos, en donde no es posible ubicar en el tiempo ni el espacio el origen de las especies domesticadas. Las tres regiones propuestas por Harlan son: 1) centro: El cercano Oriente, no-centro: África, 2) centro: El centro de China, no-centro: El sureste asiático, y 3) centro: Mesoamérica y no-centro: Sudamérica (Harlan 1971).

La siguiente modificación al concepto de centro de origen fue realizada por Hawkes (1983), quien dio fin a la confusión entre los centros de origen de la agricultura y las áreas de gran diversidad; definiendo a los

primeros como centros nucleares de la agricultura y a los segundos como regiones de diversificación, las cuales surgieron después del evento de domesticación, cuando las poblaciones cultivadas se dispersaron más allá de los centros de origen (Harris 1990).

De acuerdo con la visión de Harlan 1998 *“No es posible encontrar un tiempo y lugar para el origen de la agricultura [...] La agricultura ha emergido como el resultado de largos periodos de íntima coevolución entre las plantas y el hombre [...]; esta coevolución ha tenido lugar a lo largo de milenios y en vastas regiones [...]; por tanto, los orígenes son difusos en el tiempo y el espacio”*. Es probable que este contexto haya permitido que el concepto de centro de diversidad haya prevalecido sobre el del centro de origen, pues posibilita comprender y evaluar las condiciones que dieron origen a una alta concentración de especies cultivadas en ciertas regiones (Khoury et al. 2016).

La importancia de la definición de un centro de origen y/o de diversidad recae en las implicaciones científicas, sociales y políticas para la conservación y estudio de estas regiones. La globalización de los sistemas para la producción y distribución de los alimentos, requieren la aplicación de los conocimientos adquiridos sobre la evolución de los organismos domesticados en aras de la conservación y el aprovechamiento de la diversidad genética distribuida a en los diferentes centros de origen y diversidad (Khoury et al. 2016; Gepts 2008).

## 1.2 Mesoamérica: centro de origen del maíz y otras plantas cultivadas

Mesoamérica es una región cultural que se extiende desde las cuencas de los ríos Santiago y Pánuco, hasta la región del Valle Central de Costa Rica. Posee una vasta diversidad cultural, geográfica y biológica, además de concentrar un gran número de especies domesticadas (Farraia et al. 2019; Zizumbo-Villarreal & Colunga-García Marín 2010). El manejo de recursos vegetales por los pobladores locales, que perdura hasta hoy en día, ha influenciado el curso evolutivo de las especies domesticadas. Actualmente, los grupos indígenas de Mesoamérica utilizan más de 6000 especies vegetales con fines alimenticios, medicinales y ornamentales, entre otros (Delgado-Salinas, Caballero & Casas 2004).

En Mesoamérica, y específicamente en territorio mexicano, se domesticaron el maíz (*Zea mays* ssp. *mays* L.) (Piperno et al. 2009; Matsuoka et al. 2002), el frijol común (*Phaseolus vulgaris* L.) (Kwak, Kami, & Gepts 2009) y varias calabazas (*Cucurbita* spp.) (Smith 1997); la distribución simpátrica de las poblaciones ancestrales de estas especies eventualmente permitió que fueran cultivadas en conjunto dando origen al sistema de policultivo conocido como milpa (Zizumbo-Villarreal & Colunga-García Marín 2010). Los parientes silvestres del maíz, el frijol y la calabaza se distribuyen en el territorio mexicano y en algunas regiones también son simpátricos con su contraparte domesticada (Sánchez González et al. 2018; Freytag & Debouck 2002), lo cual ha promovido el flujo genético entre domesticados y silvestres (Hufford et al. 2013; Guerra-García et al. 2017).

El proceso de domesticación del maíz dio origen a la impresionante diversidad fenotípica y genética de las variedades tradicionales (VTs) (Bellon et al. 2018; Arteaga et al. 2016). Las VTs son cultivares de polinización abierta, adaptados al manejo tradicional practicado por agricultores que cultivan en baja escala y a una amplia variedad de condiciones ambientales, derivadas de la heterogeneidad geográfica y ambiental del territorio mexicano (Bellon et al. 2018). Este conjunto de VTs se ha clasificado en 59 razas nativas (RNs) de México (CONABIO 2011).

### 1.3 Flujo genético en plantas

El papel de la hibridación ha sido ampliamente debatido desde que Linneo propuso, hace más de 300 años, que la hibridación podría ser un mecanismo para el surgimiento de nuevos linajes de plantas (Rieseberg 2019). Aunque este debate ha persistido hasta el siglo pasado, la noción de que la hibridación y el flujo genético tiene un papel fundamental en la biología evolutiva teórica y aplicada cada vez es más aceptada (Ellstrand 2014). El flujo genético en plantas es el movimiento exitoso de genes entre poblaciones por entrecruzamiento, migración de semillas o de propágulos vegetativos (Ellstrand 1992; Slatkin 1987).

El intercambio genético, aún cuando suceda en baja frecuencia, es capaz de contraponerse a los efectos de otras fuerzas como la deriva genética, la mutación o la selección (Ellstrand 2014). La incorporación permanente de un conjunto de alelos o genes provenientes de una

población (especie, subespecie, raza, etc.) diferente, se conoce como introgresión (Rhymer & Simberloff 1996; Stewart, Halfhill, & Warwick 2003). Cuando el flujo sucede de una población de gran tamaño hacia una población de tamaño reducido, puede llevar a la extinción de la segunda (Ellstrand & Elam 1993).

## 1.4 Flujo de transgenes hacia variedades nativas y parientes silvestres

Se ha reportado que el uso de cultivos genéticamente modificados (GM) ha generado beneficios como la disminución en el uso de pesticidas en el cultivo de berenjena, y algodón (Agostini et al. 2020; Ahmed et al. 2019; Rocha-Munive et al. 2018); y la producción de cultivos con mayor contenido nutrimental como el llamado arroz dorado, rico en beta caroteno (provitamina A), que fue diseñado para prevenir la deficiencia de vitamina A que está asociada a ceguera (Potrykus 2001; Demont & Stein 2013). Así mismo, de acuerdo con agricultores mexicanos entrevistados por Rocha-Munive et al. 2018 el algodón GM cultivado en México tiene un mayor rendimiento, y requiere de una menor inversión económica en pesticidas y manejo agrícola; aunque la disminución en el uso de pesticidas va acompañada de un incremento en el uso de herbicidas, como el glifosato (Rocha-Munive et al. 2018; Benbrook 2012).

Del mismo modo, el surgimiento de plagas resistentes a las proteínas insecticidas, como el gusano de la raíz del maíz en USA (Gassmann et al. 2014), es uno de los mayores retos para el desarrollo de cultivos GM (Gassmann et al. 2014; Anderson et al. 2019). Además, puede ocurrir la transferencia y el establecimiento de transgenes de resistencia al glifosato en parientes silvestres, que ha generado malezas resistentes a este herbicida, como *Agrostis stolonifera* L. un pasto cosmopolita empleado como césped en los campos de golf (Zapiola et al. 2008). Otro caso bien documentado es la transferencia de genes de resistencia a herbicidas desde *Brassica napus* L. (canola) GM hacia su pariente silvestre *Brassica rapa* en plantaciones experimentales en Canadá (Warwick et al. 2003), donde los transgenes han permanecido hasta 6 años después de su detección sin la presión selectiva del glifosato (Warwick et al. 2007). Lo mismo se ha reportado para poblaciones ruderales de *B. rapa* en Argentina, donde la

persistencia de los transgenes se ha confirmado por más de cuatro años, a pesar de que el cultivo de *B. napus* GM está prohibido (Pandolfo et al. 2018). Por otra parte, la presión selectiva ejercida por el incremento en el uso de herbicidas como parte del paquete tecnológico de los organismos GM, ha promovido el surgimiento de malezas resistentes a glifosato aún cuando no hay transferencia de transgenes; este es el caso de *Amaranthus palmeri* una maleza que merma la producción de algodón, y para la cual se han reportado poblaciones resistentes en EUA (Norsworthy, Schwartz, & Barber 2016) y en México (Domínguez-Valenzuela et al. 2017).

A pesar de que la era de los cultivos GM inició en EUA en 1995 (Beckie & Hall 2014), hay poca evidencia que de cuenta de los impactos ecológicos y las interacciones con otros factores bióticos. Un ejemplo es el estudio realizado por Hernández-Terán et al. 2019 que realizó la comparación *in vitro* de individuos silvestres con y sin transgenes de *Gossypium hirsutum* (algodón), especie para la cual se ha reportado la presencia de transgenes en poblaciones silvestres (Wegier et al. 2011). Los resultados del análisis *in vitro* observaron diferencias estadísticamente significativas para la tasas de incremento de la altura y propagación, que resultaron menores para las plantas con transgenes; mientras que el crecimiento microbiano en el tejido resultó mayor para el mismo grupo respecto al control (Hernández-Terán et al. 2019).

El caso del maíz en México no ha estado libre de controversias (Vargas-Parada 2014), en 1998 el Comité Nacional de Bioseguridad Agrícola (CNBA) estableció una moratoria de *facto* para el cultivo de maíz transgénico que fue implementada por la Secretaría de Agricultura y Desarrollo Rural (SAGARPA) en 1999. A pesar de la moratoria, desde el 2001 diferentes investigaciones han detectado la presencia y persistencia de transgenes en VTs (Quist & Chapela 2001; Piñeyro-Nelson et al. 2009; Agapito-Tenfen et al. 2017). Recientemente, un estudio realizado por Rendón-Aguilar et al. 2019 en el estado Oaxaca sugirió que la aparente aparición (Quist & Chapela 2001; Piñeyro-Nelson et al. 2009; Agapito-Tenfen et al. 2017) y desaparición (Ortiz-García et al. 2005) de los transgenes indica un reemplazo espacio-temporal asociada a la incorporación germoplasma foráneo.

Cabe señalar, que a pesar de la persistencia de transgenes en VTs

de maíz, no se ha reportado la transferencia hacia parientes silvestres (PSs); y el único estudio que ha evaluado los potenciales impactos fenotípicos y ecológicos del flujo genético de maíz GM a VTs, fue realizado en condiciones experimentales, mediante retrocruzas controladas para las razas Tuxpeño y Tabloncillo. En ninguno de los casos se observaron diferencias entre las razas libres de transgenes y las introgresadas, y se concluyó que los transgenes introgresados siguen un patrón de segregación mendeliana (Duncan et al. 2019). No obstante, este marco experimental deja de lado el intercambio de semillas realizado por los agricultores y la heterogeneidad ambiental en la que son cultivadas las VTs de maíz.

Dada la complejidad de las interacciones que recientemente han comenzado a evaluarse (Duncan et al. 2019), resulta complicado predecir el impacto de la transferencia de transgenes (Ryffel 2014), por tanto, apearse a un principio precautorio tiene sentido cuando se trata de los centros de origen, donde se distribuyen los PSs de las plantas cultivadas, los cuales son un recurso genético valioso por su uso potencial en el mejoramiento (Acevedo et al. 2011; Myhr & Traavik 2002).

## 1.5 Flujo entre domesticados no transgénicos y parientes silvestres

El flujo genético entre poblaciones domesticadas y silvestres puede ocurrir de manera unidireccional y/o bidireccional. Sin que sea necesaria la presencia de transgenes, el flujo de genes de organismos domesticados hacia silvestres se ha asociado con la generación de nuevas malezas (Ellstrand et al. 2013) o con el origen de adaptaciones que incrementan el rango de distribución (Rieseberg et al. 2007).

En contraparte, el flujo de silvestres hacia domesticados puede producir adaptaciones que amplían el rango de distribución de los grupos domesticados (Hufford et al. 2013), un carácter deseable en plantas domesticadas. Tal es el caso de la introgresión de los PSs de maíz, también conocidos como teocintles, hacia poblaciones domesticadas. El flujo de *Zea mays* ssp. *mexicana* (*mexicana*) hacia razas de maíz resultó en la generación de variedades que actualmente se distribuyen en las regiones altas del Valle de México (Hufford et al. 2013). Aunque este tipo de flujo



también está asociado a fenotipos no deseables en los domesticados híbridos, como granos de menor tamaño (Kato 1997) o con mayor dureza (comunicación personal Adelaida Barrera, Santiago Analco, Estado de México), que son seleccionados negativamente por el agricultor.

La domesticación es un proceso histórico y continuo dirigido por la selección artificial (Harlan 1992). En el caso del maíz, las VTs de México han coexistido desde su origen con sus PSs, y hasta la fecha es común encontrar en los campos de cultivo de la Meseta Central de México y en el suroeste de la cuenca del Balsas, a los teocintles creciendo alrededor y dentro de las parcelas. Sin embargo, este paisaje fue modificado durante la primera mitad del siglo XX por los avances en genética y en tecnología agrícola, que impulsaron en algunas regiones del país el cultivo de variedades modernas de maíz (VMs) derivadas por un proceso de mejoramiento con cruza controladas (Wellhausen 1961).

Las primeras VMs de maíz se generaron en Estados Unidos de América (EUA) al final de la década de 1920 e inicios de la década de 1930. Su introducción fue exitosa debido al incremento en la producción de grano y en menos de diez años su adopción fue casi total en Iowa y en otros estados del cinturón maicero de EUA. La primera consecuencia de la adopción de VMs, fue la erosión genética a nivel de variedades, pues su uso llevó al abandono de las variedades de polinización abierta de EUA (Crow & Dove 1998).

El caso de México ha sido diferente, debido a que el sistema agrícola mexicano difiere sustancialmente del modelo de EUA, esto se debe a la heterogeneidad ambiental y cultural. En México el 57.94 % de las unidades agropecuarias en México son del tipo familiar (FIDA 2014), las cuales se caracterizan por un sistema de organización familiar, en donde la producción es usada de manera parcial o total para el autoconsumo, el uso de VTs es frecuente y suelen ser sistemas de policultivo (Bellon et al. 2018). Estas características, en conjunto con la adopción gradual de las VMs, el reciclaje y la adaptación de la semillas mejoradas al manejo tradicional -- una práctica conocida como acriollamiento (Bellon & Brush 1994) -- implicó la llegada de semillas migrantes y el probable intercambio genético entre VMs, VTs e incluso PSs de maíz.

El estudio del flujo genético de VMs hacia RNs y PSs ha sido menos

exhaustivo a nivel genético, la mayor parte de la evidencia se basa en caracteres morfológicos (Ortega Paczka 1973). Detectar la presencia de transgenes es relativamente sencillo debido a que las secuencias de DNA o proteínas evaluadas provienen de especies diferentes (Piñeyro-Nelson et al. 2009; Quist & Chapela 2001). En contraste, evaluar el flujo de VMs hacia VTs implica poner a prueba la transferencia de alelos entre dos poblaciones que han divergido hace decenas de generaciones, por un proceso de mejoramiento, que implica un efecto fundador asociado a un proceso de selección artificial (Yamasaki et al. 2005, Hufford et al. 2012), estas condiciones dificultan la distinción entre ancestría y flujo genético o introgresión recientes. Sin embargo, el análisis de germoplasma antiguo resguardado en los bancos de germoplasma ha sido empleado para establecer una referencia histórica, que permite analizar los cambios contemporáneos en especies domesticadas cuyo acervo genético ha sido moldeado por el desplazamiento de VTs (Steele et al. 2009), y/o el flujo con VMs (Bitocchi et al. 2009).

## 1.6 Barreras al flujo genético

De acuerdo con el modelo de aislamiento por distancia, debería existir una relación proporcional entre la tasa de migración y la distancia geográfica entre dos poblaciones (Wright 1946, 1943). Bajo esta perspectiva, se ha propuesto con base en evidencia experimental que el aislamiento espacial podría funcionar como estrategia para minimizar el flujo entre maíces GM y RNs, pues cuando la distancia entre estos cultivos es mayor a 20 m, la tasa de entrecruzamiento por polen y viento cae por debajo del 1% (Baltazar et al. 2015). Sin embargo, estas perspectivas pasan por alto la dispersión de semillas mediada por humanos (Wichmann et al. 2009); un sistema de distribución informal que consiste en el intercambio, entrecruzamiento, manejo y reemplazo de semillas (Thiele 1999). Un caso que refleja la importancia del sistema informal de semillas es un estudio realizado en el estado de Oaxaca, el cual no encontró aislamiento por distancia entre las razas de maíz estudiadas (Pressoir & Berthaud 2004).

El flujo genético entre organismos domesticados y silvestres puede ocurrir de manera asimétrica, debido al tamaño censal de las poblaciones, cuando el número de individuos silvestres que intercambian polen con los

domesticados se limita a aquellos que se encuentran dentro o alrededor de un campo de cultivo (Papa & Gepts 2003). Otra barrera al flujo genético e introgresión es la selección artificial ejercida por los agricultores; después de un evento de hibridación entre un domesticado y un silvestre, los descendientes de la  $F_1$  pueden presentar fenotipos intermedios, como un tamaño de semilla reducido o una coloración diferente (Kato 1997), lo que facilita su reconocimiento por los agricultores y disminuye las probabilidades de introgresión. Por tanto, la selección artificial negativa en contra de los híbridos  $F_1$  difiere de la selección ejercida por el ambiente, pues afecta a todo el genoma cuando los híbridos no son sembrados en el ciclo siguiente, por tanto, este tipo de mecanismo resulta similar al de una barrera postcigótica (Papa & Gepts 2003).

## 1.7 Sistemas de incompatibilidad

Los mecanismos de incompatibilidad genética, en específico los de auto-incompatibilidad, parecen haber surgido como una estrategia que promueve la fecundación cruzada en las angiospermas (Newbigin, Anderson, & Clarke 1993; Muthuselvi & Praneetha 2019; Brugière, Rothstein, & Cui 2000). Un mecanismo adicional es la incompatibilidad cruzada (IC) que se define como la interacción (o la ausencia de ella) entre el polen y el pistilo que previenen la formación de un cigoto híbrido entre dos especies fértiles (de Nettancourt 2001; Maune, Camadro, & Erazzú 2018). La IC puede ser unidireccional o bidireccional dependiendo de la combinación de los gametos (Maune, Camadro, & Erazzú 2018). La IC es una forma de aislamiento reproductivo que se ha propuesto como estrategia para prevenir el intercambio genético no deseado durante la producción de semillas mejoradas de maíz (Nelson 1953; Evans & Kermicle 2001).

*Zea mays* es una especie monoica, con flores masculinas y femeninas en inflorescencias diferentes, y los individuos pueden autofertilizarse o ser fecundados con polen de otras plantas (Wang et al. 2018). En *Zea mays*, se han descrito tres sistemas de incompatibilidad genética: *Teosinte crossing barrier 1-strong* (Tcb1-s), *Gametophyte factor 1-strong* (Ga1-s) y *Ga2-s*. Los tres sistemas se encuentran en las VTs de maíz y en teocintles (Evans & Kermicle 2001; Kermicle et al. 1990; Kermicle & Evans 2010; Padilla García

et al. 2012).

Los pistilos de las VTs de maíz y teocintles que tienen el alelo Ga1-S pueden ser fertilizados por polen con genotipo Ga-1S o Ga1-m, pero no aceptan polen con el alelo recesivo ga1; sin embargo, el polen Ga1-S/- pueden fecundar una planta ga1/ga1 (Kermicle 2006; Padilla García et al. 2012). En el caso del sistema *tcb*, el alelo Tcb1 es capaz de polinizar a los estigmas tcb-1, pero la fecundación en el otro sentido no ocurre. En un principio se pensó que el sistema *tcb* era exclusivo de los teocintles y que actuaba como un mecanismo de aislamiento entre el maíz y sus PSs (Evans & Kermicle 2001), aunque posteriormente su presencia fue reportada en razas de maíz (Padilla García et al. 2012).

El comportamiento del *locus* ga-2 es parecido a los dos sistemas anteriores, se han descrito cuatro alelos, Ga2-w que presenta una barrera débil del pistilo, Ga2-s que es una barrera fuerte del pistilo, Ga2-m que se comporta de manera neutral y sólo se ha reportado en teocintles, y el último alelo ga2, que se ha descrito como un alelo nulo (Kermicle & Evans 2010; Padilla García 2011).

Recientemente, se ha evaluado la presencia de los alelos, Ga1-s y Tcb1, responsables de la compatibilidad no recíproca en 84 variedades híbridas de maíz. Ninguno de los alelos dominantes fue detectado; y el alelo neutral Ga-m se halló en el 63 % de las muestras evaluadas, lo que indica que la incompatibilidad genética no es suficiente para mantener aisladas a las VMs de las RNs y de sus PSs (Padilla-García et al. 2016).

## 1.8 Métodos para la detección de flujo genético

La introducción de las isoenzimas en la década de 1960 permitió explorar la variación presente en las poblaciones naturales. Los análisis realizados con isoenzimas se basaron en la existencia de gradientes de frecuencias alélicas para proponer oleadas de migración, donde la frecuencia de los alelos migrantes disminuye gradualmente conforme aumenta la distancia entre poblaciones (Schaffer & Johnson 1974). Otras estrategias empleadas para estimar las tasas de migración fueron el coeficiente de divergencia  $F_{ST}$  introducido por Wright (1931), el coeficiente  $G_{ST}$  introducido por Nei (1973) -- considerado una extensión del coeficiente de Wright que se aplica cuando existen múltiples alelos (Takahata & Nei 1984)-- y el modelo de

islas de Wright (Neigel 1997; Slatkin 1985). Los métodos anteriores se consideran indirectos, por requerir modelos genéticos con parámetros poblacionales previamente definidos para inferir la magnitud del flujo; siendo esa su principal desventaja; por la dificultad que implica estimar de manera certera estadísticos como el tamaño efectivo de la población ( $N_e$ ) (Neigel 1997). No obstante, la confianza en estos se debe a la robustez que poseen para estimar flujo genético en amplias escalas espaciales y temporales (Neigel 1997; Slatkin 1985). Además, de la incertidumbre asociada a la estimación de los parámetros poblacionales requeridos para simular modelos demográficos complejos, otra desventaja era el tiempo de cómputo requerido para su simulación.

La posibilidad de obtener secuencias de DNA llevó al desarrollo de modelos genealógicos o de coalescencia. La inferencia de flujo genético empleando coalescencia se ha basado en: 1) La longitud de las ramas que conectan las secuencias en el árbol genealógico, que indican el tiempo a la coalescencia, y que son empleadas como la distancia entre las secuencias; y 2) El orden de las ramas en el árbol, que muestran el orden de los eventos de coalescencia (Neigel 1997). Métodos como *CoalHMM* y *fastSimCoal2* (Excoffier & Foll 2011) se basan en el uso de modelos evolutivos a priori, y al igual que los modelos genéticos empleados para isoenzimas, su principal inconveniente es la precisión con que se calculan los parámetros para alimentar el modelo (Zheng & Janke 2018). Brevemente, *fastSimCoal2* es un método de probabilidad compuesta que emplea el espectro de frecuencias alélicas (SFS, por sus siglas en inglés) y tiene un buen desempeño para estimar la tasa de migración ( $Nm$ ) cuando se exploran modelos evolutivos de complejidad arbitraria con más de tres poblaciones (Excoffier et al. 2011); aunque su desempeño es menos eficaz cuando se desconoce a la población donadora de alelos migrantes (Jenkins, Song, & Brem 2012; Excoffier et al. 2013; Kraft et al. 2014). En contra parte, los métodos basados en genealogías que evalúan la presencia de haplotipos no fijados, han resultado una buena estrategia cuando la fuente de migrantes se desconoce (Jenkins, Song, & Brem 2012).

Por otro lado, el estudio de procesos evolutivos contemporáneos, que están en el orden de decenas generaciones (Franks et al. 2016) donde las prácticas humanas y hechos contemporáneos son factores clave,

permite la inclusión de evidencia histórica y cultural (Brown et al. 2013; Larson et al. 2012; Tixier-Boichard, Bed'hom, & Rognon 2011). Por tanto, la suposición de los escenarios evaluados se basa en eventos conocidos y en relaciones filogenéticas fiables.

El estadístico D también conocido como prueba de ABBA-BABA, ha mostrado ser robusto cuando se cuenta con evidencia suficiente sobre las relaciones filogenéticas entre las poblaciones con intercambio genético (Zheng & Janke 2018). Este método fue desarrollado por Green et al. 2010, para evaluar la introgresión de *Homo neanderthalensis* en grupos humanos de Eurasia. Desde su publicación ha sido ampliamente usado para poner a prueba la existencia de flujo genético entre especies cercanas de osos (Kumar et al. 2017), equinos (Kumar et al. 2017; Jónsson et al. 2014), mariposas (The Heliconius Genome Consortium 2012), plantas (Escudero et al. 2014; Owens, Baute, & Rieseberg 2016) y patógenos microbianos (Martin et al. 2016).

El análisis de ABBA-BABA se basa en la comparación de 4 poblaciones, cada una situada en la punta de un árbol que representa su relación filogenética; donde H1 y H2 corresponden a un mismo clado que ha divergido del grupo H3; y H4 es un grupo ancestral externo. Este método se limita a evaluar los sitios bialélicos donde se observan dos copias para el mismo alelo por genotipo, lo que reduce la probabilidad de que la substitución observada se deba a errores de secuenciación. Por tanto, el análisis considera aquellos *loci* en los que H1 y H2 difieren, el alelo ancestral es denotado con la letra "A" y el derivado con la letra "B". Los alelos que cumplen con un patrón ABBA o BABA son los que se emplean para estimar el estadístico D. En ausencia de flujo genético se esperaría que ambos estados ocurran con la misma frecuencia, sin embargo, su proporción puede variar de manera significativa si ha ocurrido flujo genético de H3 hacia H2 (ABBA) o H3 hacia H1 (BABA). Por otra parte, si la relación filogenética establecida no es correcta, es decir si las poblaciones H1 y H2 no forman parte de un mismo clado o la tasa de mutación difiere entre ellas, el análisis puede arrojar falsos-positivos (Green et al. 2010). Así mismo, se ha evaluado la robustez del análisis variando parámetros como el tiempo de divergencia entre H1 y H2, el tamaño poblacional y la distancia al grupo externo, de lo anterior se

concluyó que el principal factor asociado a la deficiencia para detectar flujo genético es la separación incompleta de linajes y el tamaño relativo de las poblaciones respecto al tiempo de divergencia entre H1 y H2 (Zheng & Janke 2018).

Los métodos y estrategias enumerados en esta sección pueden combinarse, para evaluar la robustez de los resultados obtenidos. Basados en lo anterior, una estrategia factible es iniciar con una aproximación sencilla basada, por ejemplo, evaluar la existencia de flujo genético por medio de caracteres fenotípicos, y posteriormente escalar a modelos que incluyan un mayor número de parámetros y poblaciones.

### 1.9 Planteamiento del problema: panorama socioeconómico, político y tecnológico

En la década de 1940 las primeras VMs de maíz basadas en variedades nativas comenzaron a producirse a baja escala en México, reconocido como centro de origen del maíz (Matsuoka et al. 2002; Piperno et al. 2009), lo que llevó a la simpatría con VTs y sus PSs, y probablemente al intercambio genético entre estos grupos (Ortega Paczka 1973; Perales, Brush & Qualset 2003). La introducción de VMs de maíz no causó la misma oposición que la solicitud de uso de cultivos GM realizada décadas más tarde. No obstante, el uso de VMs en el centro de origen, sí dio pie a varios cuestionamientos sobre las consecuencias del flujo genético entre VMs, VTs y PSs (Serratos 1997).

La adopción de VMs suele estar asociada a un incremento en la productividad de los cultivos, aunque también lo está con la erosión genética por el desplazamiento parcial o completo de VTs, o por la pérdida de alelos asociados a adaptaciones locales y a ambientes heterogéneos (Serratos 1997; van de Wouw et al. 2010; Brush 1995). Este fue el caso de la rápida sustitución de las variedades de maíz de polinización abierta por las VMs introducidas en el oeste medio de EUA en la década de 1930. En 1935 la proporción de maíz híbrido cultivado era menor al 10 % y en solo 4 años incrementó al 90 % (Crow & Dove 1998).

En contraste con lo anterior, se ha reportado que cuando el desplazamiento de las VTs es parcial, la simpatría con VMs puede generar un incremento de la diversidad genética debido a la introgresión de alelos nuevos procedentes de las VMs (Steele et al. 2009; Bitocchi et al. 2009;



Ortega Paczka 1973). Este incremento es seguido por una caída en la diversidad genética en las VTs (Brush, Taylor, & Bellon 1992; Bellon & Brush 1994), que se estabiliza después de la sustitución parcial o completa de las VTs por las VMs (van de Wouw et al. 2010). Lo anterior implica que para las plantas cultivadas las prácticas humanas constituyen el móvil principal para amortiguar o prevenir la pérdida de diversidad genética (Brush 1995; Bellon & Brush 1994).

Las primeras VMs de maíz generadas en México fueron variedades de polinización abierta bajo un par de ciclos de selección masal y fueron liberadas en 1946. Estas VMs fueron producidas por la Fundación Rockefeller en asociación con el Secretaría de Agricultura de México. Inicialmente, los esfuerzos para producir maíces mejorados se enfocaron en las regiones de la Meseta Central, el noroeste y suroeste de México (Wellhausen 1961).

La década de 1960 es reconocida como el inicio de la revolución verde, que llevó al incremento en la producción de granos a nivel mundial (Evenson & Gollin 2003). Esta fue el resultado de la industrialización de la agricultura, un proceso que consistió en la adopción de maquinaria agrícola como el tractor, el uso de pesticidas, fertilizantes, y la aplicación de las técnicas de mejoramiento para generar variedades con mayor rendimiento de grano (Sonnenfeld 1992).

México ha sido asociado a la revolución verde debido al exitoso programa de mejoramiento de trigo liderado por el agrónomo Norman Borlaug (Sonnenfeld 1992). En 1956 México se volvió autosuficiente para la producción de trigo y para 1960 el 90% de la producción de trigo provenía de las VMs (CIMMYT 1976). En contraste el caso del maíz fue diferente. La aceptación de las VMs de maíz fue limitada y para 1976 menos del 15% del área cultivada con maíz era sembrada con VMs (CIMMYT 1976). A pesar de lo anterior, reportes posteriores señalan que en el sureste de México las semillas de algunas de las VMs fueron guardadas y sembradas en ciclos posteriores como variedades de polinización abierta. Este proceso conocido como acriollamiento llevó a la disminución del área sembrada con VTs (Jones 1977; Ortega Paczka 1973; Bellon & Risopoulos 2001; Hernández Xolocotzi & Ortega Paczka 1987; Bellon & Brush 1994).

En la década de 1960 fueron creadas dos instituciones en México: el Instituto Nacional de Investigaciones Agrícolas (INIA) y el Centro Internacional de Mejoramiento de Maíz y Trigo (CIMMYT). En 1961 La Ley de Semillas designó la Productora Nacional de Semillas (PRONASE) como la entidad encargada de la producción y comercio de las VMs desarrolladas por el INIA. Así, en el periodo que abarca de 1960-1990, la producción y comercio de las VMs estuvo controlada por el gobierno (Jones 1977; Ortega Paczka 1973; Bellon & Risopoulos 2001; Hernández Xolocotzi & Ortega Paczka 1987; Bellon & Brush 1994). En 1991 el Acta de Semillas, fue modificada y permitió la participación del sector privado en la investigación, la producción y el comercio de las VMs. Para finales de los 90s, el mercado fue transferido a las semilleras privadas. En 2007 se decretó la desaparición de la PRONASE, lo que incentivó la introducción de germoplasma no nativo. Tan solo dos años más tarde, en el 2009, el 95% de la semilla híbrida de maíz fue producido por dos empresas transnacionales: Monsanto y Pioneer, lo que implica que el mercado era controlado casi en su totalidad por el sector privado (Luna Mena et al. 2012; Espinosa-Calderón et al. 2014).

La información actual sobre el territorio en México plantado con VMs de maíz es heterogénea, los estimados registrados varían del 42.5% (García-Salazar & Ramírez-Jaspeado 2014) al 58% (SIAP 2016). Sin embargo, los cálculos basados en la penetración del mercado de las VMs (Donnet et al. 2012), indican que las estadísticas oficiales pueden estar sobreestimadas (Bellon et al. 2018). A pesar de lo anterior, todas las fuentes concuerdan en que el uso de VMs es mayor en el noroeste y el occidente central de México, en contraste con el resto del país (García-Salazar & Ramírez-Jaspeado 2014).

La adopción parcial de las VMs puede explicarse por valor agregado que poseen las RNs, por ejemplo los atributos sensoriales, estéticos y una mayor calidad para preparar platillos locales (Perales, Brush & Qualset 2003; Bellon & Taylor 1993; Zarazua-Escobar, Almaguer-Vargas, & Gustavo Ocampo-Ledesma 2011), además, de su adaptación a condiciones locales, y a ambientes marginales y suelos pobres (Bellon & Taylor 1993; Bellon & Brush 1994). Las VTs de maíz son la base de la cocina tradicional mexicana (Perales, Brush & Qualset 2003; Miramontes-Piña 2006),

mientras que las VMs son usadas para la industria de la tortilla, la producción de harinas o para alimentar ganado (García-Salazar & Ramírez-Jaspeado 2014; Miramontes-Piña 2006).

La adopción parcial de las VMs y el acriollamiento de las mismas ha propiciado la simpatría de las VMs con las VTs y sus PSs, las cuales suelen crecer en campos vecinos o en la misma localidad, lo que facilita el flujo genético de las VMs hacia las VTs, el cual en ocasiones es promovido por los agricultores, a la fecha sólo se cuenta con evidencia fenotípica de este intercambio (Ortega Paczka 1973; Perales, Brush & Qualset 2003; Bellon & Berthaud 2004), pero no se han documentado los cambios a nivel del genoma o de estructura poblacional.

### 1.10 Sustitución de variedades tradicionales y/o flujo genético con variedades modernas

La introgresión de alelos de VMs en VTs de plantas cultivadas es un evento predecible si consideramos que se trata de poblaciones de una misma especie que han divergido a través de un proceso de selección artificial. Existen evidencias de flujo genético de VM hacia VTs de arroz en Nepal (Steele et al. 2009), de cebada (Bellucci et al. 2013) y maíz (Bitocchi et al. 2009, 2015), ambos en Italia.

El flujo VMs hacia VTs de arroz (*Oryza sativa* L.) fue documentado desde su inicio, en el del distrito de Kaski, Nepal, ubicado en un rango altitudinal de 1200- 1600 msnm. Las VMs de arroz fueron generadas a través de un modelo de mejoramiento participativo (Morris & Bellon 2004), y las primeras VMs fueron liberadas entre 1996 y 1999. El área dedicada al cultivo de las VTs más comunes disminuyó, y algunas VTs fueron abandonadas completamente; aunque las variedades “raras” fueron conservadas por los agricultores por sus adaptaciones y usos específicos (Steele et al. 2009). En menos de una década fue posible detectar para seis VTs la presencia de alelos de microsatélites que no estaban presentes anteriormente en variedades locales, y que produjeron un incremento en la diversidad genética. Los autores predijeron con un modelo teórico que sólo cuando la sustitución de VTs sea mayor al 65 %, la variación genética comenzará a disminuir (Steele et al. 2009). Este modelo coincide con el propuesto por van de Wouw et al. (2010), en el que

las VMs sustituyen parcialmente a las VTs; luego, las VTs ganan diversidad genética por la incorporación de nuevos alelos, para después de un periodo de estabilización en el cual la mayoría de las VTs son sustituidas, la diversidad genética disminuye.

Otro ejemplo es el de la papa (*Solanum tuberosum* L.). La papa es el tercer cultivo en importancia para la dieta humana y el primero cuando no se consideran los cereales (Monte et al. 2018). Su centro de origen se localiza en Perú (Spooner et al. 2005) y se tiene registro de más de 4000 variedades nativas (CIP, 2019). Los cultivos de papa en los Andes no experimentaron la revolución verde con la misma intensidad que los cereales; sin embargo, se introdujeron nuevos insumos y las primeras variedades mejoradas fueron liberadas en Perú en 1950. A pesar de que las variedades nativas eran percibidas como superiores en términos culinarios, agrónomicamente se consideraron como inferiores ante las nuevas variedades (Brush et al., 1992).

La hipótesis de erosión genética de Hawkes (1983) propone que la erosión genética es el resultado parcial del éxito del mejoramiento, que lleva a un desplazamiento gradual de las formas primitivas y de las VTs en regiones con alta diversidad. Sin embargo, en el caso de la papa en Perú, cuando se consideró el número de variedades nativas de papa y el área de cultivo dedicada a éstas en comparación con las VMs, el área dedicada a las VTs disminuyó, pero los agricultores conservaron un mayor número de variedades en un área menor, lo que sugiere que la pérdida de genotipos puede ser menor a lo predicho por Hawkes (Brush 1992; Brush, Taylor, & Bellon 1992). En un trabajo posterior (Brush et al. 1995) con isoenzimas para 610 tubérculos colectados en los campos de cultivo de dos microrregiones; y 503 tubérculos colectados en los mercados de siete provincias alrededor de Cusco, Perú; se identificaron 112 genotipos, con un alto grado de endemismo, y un número reducido de variedades que se comparten entre provincias. Esto indica que son los genotipos y no las variedades la unidad de conservación. Por tanto, la erosión genética en términos de la pérdida de genotipos únicos es un riesgo real, aunque la conservación de los alelos en terrenos pequeños parezca ser una estrategia exitosa de conservación *in situ* (Brush et al. 1995). A pesar de lo anterior, nuestra capacidad para medir el impacto de la tecnología en la

agrobiodiversidad está limitado por el gran número de factores a considerar, como la colecta de los datos, y las variables sociales, económicas, políticas, culturales, y ecológicas involucradas (Brush 1992). En la misma línea, las estrategias de modernización de la agricultura que desincentivan la conservación de las razas, aunada a la reducción del área de cultivo destinada a las VTs ponen en peligro la conservación *in situ* y los servicios evolutivos que prestan los agricultores a baja escala.

### 1.11 Flujo genético de variedades modernas hacia razas de maíz

En México los primeros trabajos que documentaron el flujo genético de VMs hacia VTs se valieron de la “xenia”, que se define como el efecto de los genes del polen en el desarrollo de los frutos o las semillas (Bulant et al. 2000). Caracteres como el color y el tipo del endospermo fueron utilizados para demostrar la existencia de flujo genético entre diferentes tipos de maíz con floración simultánea, y proximidad geográfica, principalmente para individuos que se encontraron en los bordes de las parcelas a unos 15 metros de la fuente del polen (Ortiz Torres, 1993, en Serratos 1997). Ortega Paczka (1973) comparó colectas de la raza *Tuxpeño* realizadas en 1971 en el estado de Chiapas, contra las colectas realizadas por Hernández Xolocotzi, hechas en el mismo estado en 1946. De acuerdo con el análisis fenotípico, un tercio de las accesiones colectadas en 1971 presentaron introgresión de VMs. En un trabajo posterior se reportó que el rendimiento de las VTs incrementó un 25 % y se propuso que tal incremento estaba asociado a la introgresión de las VMs en las VTs, o a la selección masal continua que ejercen los agricultores (Vega, 1973 en Serratos 1997) .

La introgresión y el flujo genético pueden ocurrir de manera espontánea o intencional. En el estado de Morelos, las VMs ocupan un tercio del área sembrada y los agricultores han promovido el intercambio genético de la VT *delgado* (nombre de local que recibe la raza *Pepitilla*) con variedades de polinización abierta o la F<sub>1</sub> de VMs. Algunas de las características favorables de este entrecruzamiento son una mayor tolerancia a la sequía, resistencia al acame y un incremento en la productividad (Perales, Brush & Qualset 2003); la última condición apoya

la hipótesis de que el intercambio entre RNs y VMs puede llevar a incremento en la producción de grano.

En ausencia de datos genéticos, es difícil establecer los cambios a nivel de diversidad genética, la pérdida de alelos o los cambios en la estructura poblacional; esta última es una consecuencia que ha sido sugerida anteriormente (Brush 1992). A pesar de lo anterior, la evidencia recolectada después de la introducción de VMs en el centro de origen del maíz apunta hacia un proceso evolutivo influenciado por el intercambio genético entre RNs tradicionales y las VMs introducidas en la segunda mitad del siglo pasado.

Anteriormente, el trabajo realizado por van Heerwaarden (2007) en México, exploró el flujo genético entre VTs en un gradiente altitudinal que incluyó 60 colectas de VTs para las cuales se amplificaron 11 microsatélites. Sus resultados mostraron que el flujo genético está restringido por la distribución altitudinal de las VTs, y no por la proximidad geográfica. Aunque, cuando el flujo genético ocurre suele darse en mayor proporción de VTs de tierras bajas hacia VTs de tierras altas. Sin embargo, este trabajo no exploró el flujo de VMs hacia VTs.

Un ejemplo más reciente, realizado con microsatélites y polimorfismos en la longitud de fragmentos amplificados (AFLPs, por sus siglas en inglés), demostró el cambio en la estructura poblacional asociada a la introgresión de VMs en VTs de maíz en el norte de Italia. Debido a que el maíz es un cultivo anual, en el lapso de unas decenas de generaciones fue posible identificar cambios en la estructura genética de las poblaciones. Al comparar los genotipos de VTs colectadas en la década de 1950 con las VTs colectadas en el año 2000, se detectó que la introgresión de alelos proveniente de las VMs en las VTs locales fue favorecida por la selección artificial (Bitocchi et al. 2015).

Como se ha mencionado anteriormente, algunas de las consecuencias asociadas a este flujo constante, son la pérdida eventual de alelos locales, que son reemplazados por alelos provenientes de las VMs, un cambio en la estructura poblacional y el desplazamiento gradual de las variedades locales por VMs. Este conjunto de eventos pueden interpretarse como un cuello de botella asociado a la modernización, dividido en dos etapas: la primera está caracterizada por una reducción de

la diversidad genética, debido al reemplazo de VTs por las VMs, y la segunda por la estabilización en la pérdida de diversidad genética, que se alcanza cuando el reemplazo de las VTs se ha completado (van de Wouw et al. 2010).

## 1.12 Acervo de germoplasma de maíz en México

El Proyecto Global de Maíces Nativos es una iniciativa liderada por el Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO), cuyo objetivo fue actualizar la información de maíces y sus PSs en México para determinar los centros de diversidad genética en el país. Como parte de este proyecto se realizaron colectas de las VTs de maíz distribuidas en todo el territorio mexicano. Este proyecto clasificó a las VTs en las 59 razas nativas (RNs) reconocidas a nivel autonómico (CONABIO, 2011). La información resultante de este esfuerzo de colecta se combinó con la información de las colectas que se han realizado desde la década de 1940. La base de datos del proyecto global de maíces nativos cuenta con 22 931 registros de accesiones resguardadas en diferentes bancos de germoplasma localizados en México. La colección está clasificada en tres periodos de colecta: 1) 1940-1965, 2) 1966-1990 y 3) 1991-2010 (“Proyecto Global de Maíces Nativos” 2011).

En este trabajo evaluamos si existe el flujo genético de las VMs hacia las VTs agrupadas en 7 RNs de maíz y hacía dos de sus PSs: *Zea mays* ssp. *parviglumis* (*parviglumis*), reconocido como el ancestro silvestre (Matsuoka et al. 2002), y *Zea mays* ssp. *mexicana* (*mexicana*). Las dos subespecies silvestres crecen en simpatría con las VMs, en las regiones de la Mesa Central Mexicana y en la cuenca del Río Balsas, que abarca el centro y suroeste de México. Esta última es reconocida como el centro de domesticación del maíz (Matsuoka et al. 2002; Piperno et al. 2009; Ranere et al. 2009). Basados en este contexto, hipotetizamos que existe un flujo genético continuo de las VMs hacia las RNs y sus PSs distribuidos en la Mesa Central y en la cuenca del Río Balsas.

Para evaluar los niveles de introgresión de las VMs en VTs y en sus PSs comparamos los genotipos de las siete RNs y de dos subespecies de teocintles: *mexicana* y *parviglumis*, agrupados en tres periodos, de acuerdo con el año de colecta: antes de 1960, entre 1960-1980 y posterior al 2000.



En el caso de *parviglumis* y *mexicana*, las accesiones se dividieron en dos grupos, las más antiguas fueron colectadas en las décadas de 1970 y 1980; y las recientes colectadas después del año 2000.

También se incluyeron a las VMs encontradas en simpatría con las VTs y los PSs al momento de colecta, y un conjunto de poblaciones alopátricas de PSs que se emplearon como control negativo del flujo genético por simpatría. Los genotipos fueron obtenidos por el método de Genotipificación por Secuenciación (GBS, por sus siglas en inglés) (Elshire et al. 2011), y se compararon de acuerdo con el periodo de muestreo. Las comparaciones entre muestras colectadas en diferentes periodos de tiempo nos permitieron evidenciar la existencia de flujo genético de las VMs hacia RNs y *mexicana*, detectar un cambio reciente en la estructura poblacional de las RNs y un incremento de la diversidad nucleotídica de las RNs y su PSs.

## Hipótesis

La introducción de variedades modernas de maíz ha promovido el flujo genético de éstas hacia razas nativas de maíz y sus parientes silvestres, particularmente en regiones donde crecen de manera simpátrica.

## Objetivo general

Evaluar el flujo genético de variedades modernas hacia razas nativas y su parientes silvestres, estimar su extensión y determinar los cambios en el genoma de razas de maíz y sus parientes silvestres.

## Objetivos particulares

1. Evaluar la estructura poblacional de siete razas de maíz antes y después de la adopción de variedades modernas.
2. Evaluar la existencia de flujo genético de variedades modernas hacia razas y parientes silvestres de maíz.
3. Identificar si las VMs cultivadas en simpatría con razas y parientes silvestres están basadas en germoplasma nativo o exótico.
4. Comparar la divergencia poblacional ( $F_{ST}$ ) entre razas y variedades modernas; razas y parientes silvestres, antes y después de la introducción de variedades modernas.
5. Determinar si el flujo genético ha modificado la diversidad genética en razas de maíz y sus parientes silvestres.

## 2. Materiales y Métodos

### 2.1 Muestras y accesiones

El trabajo incluyó muestras de VTs clasificadas en siete RNs que actualmente son cultivadas en simpatría con sus PS y/o VMs. Las RNs y los PSs colectados después del año 2000 fueron comparadas con muestras de las mismas RNs y PSs colectadas en periodos anteriores, para evaluar el flujo genético con VMs.

Para seleccionar las RNs examinamos la base de datos del Proyecto de Maíces Nativos (CONABIO 2011), y seleccionamos a aquellas muestreadas en las regiones de la Mesa Central, la cuenca del Río Balsas y la región de Oaxaca (Fukunaga et al. 2005). Asimismo, definimos tres periodos para clasificar las muestras de acuerdo con el año de colecta: 1) Antes de 1960, previo a la adopción de VMs, 2) Entre 1960-1980, y 3) Después del 2000.

Además de considerar el año de colecta, seleccionamos a aquellas RNs que fueron muestreadas dentro de la distribución potencial de los PSs y en proximidad geográfica con las regiones donde actualmente se siembran en simpatría RNs y PSs. Basado en lo anterior supusimos que las accesiones antiguas estuvieron en simpatría con *mexicana* y *parviglumis* (Fig. 1).

Siete localidades y RNs cumplieron con los criterios establecidos: Chalqueño, Zamorano, Mushito, Pepitilla, Tabloncillo, Bolita y Vandeño. Cada RN representa una unidad simpátrica integrada por una RNs, un PSs y una o más o VMs. Por otra parte, con el fin de analizar los datos a nivel global las colectas de RNs, VMs y PSs fueron agrupados como se detalla en la Tabla 1.

Posterior a la selección de las RNs y sus PSs, visitamos las localidades donde se ha reportado que las siete RNs son simpátricas con los PSs: *mexicana* o *parviglumis*. Cuatro de las siete RNs y sus PSs simpátricos, fueron colectados en proximidad con una o más VMs. Las tres RNs restantes y sus PSs simpátricos, no se encontraron en la vecindad de cultivos de VMs al momento de ser colectadas, aunque fueron incluidas para evaluar si la simpatría al momento del muestreo era necesaria para detectar flujo genético.

Todas las VMs incluidas en este trabajo corresponden a la generación F1 de VMs comerciales colectadas entre 2014 y 2015. Como control en la evaluación del flujo entre PSs y VMs simpátricas, incluimos poblaciones alopátricas de *mexicana* y *parviglumis*.

Para las colectas realizadas entre 2014-2015, se muestrearon al menos 5 mazorcas completas de madres diferentes, con excepción de Vandeño la cual fue colectada en Huetamo, Michoacán y para la cual los agricultores nos proporcionaron las semillas desgranadas. La identidad racial de las VTs colectadas en 2014 fue determinada por el Dr. Juan Manuel Hernández Casillas, del INIFAP, y las mismas localidades fueron visitadas en 2015. En el caso de los teocintles, recolectamos las semillas de al menos 5 plantas, éstas fueron mezcladas y 5 semillas por localidad se seleccionaron al azar para germinar y genotipificar.

Tabla 1. Asignación temporal de las mazorcas colectadas y las accesiones de maíz solicitadas a los bancos de germoplasma: razas nativas (RNs), variedades modernas (VMs) y parientes silvestres (PSs): *Zea mays* ssp. *mexicana* (Zmx), *Zea mays* ssp. *parviglumis* (Zpr) y *Zea diploperennis*.

<b>Población</b>	<b>Abreviatura por población</b>	<b>N</b>
RNs colectadas entre 1960-1980	RN 1960-1980	35
RNs colectadas entre 2000-2015	RN >2000	106
VM colectadas entre 2014-2015	VMs	47
<i>Zea mays</i> ssp. <i>mexicana</i> colectadas 1984	Zmx1980	12
<i>Zea mays</i> ssp. <i>mexicana</i> colectadas entre 2000-2015	Zmx2000	48
<i>Zea mays</i> ssp <i>mexicana</i> alopátrica colectadas entre 2002-2015	mexicana alopátrica	11
<i>Zea mays</i> ssp. <i>parviglumis</i> colectadas entre 2000-2015	Zpr2000	44
<i>Zea diploperennis</i>	Zd	12

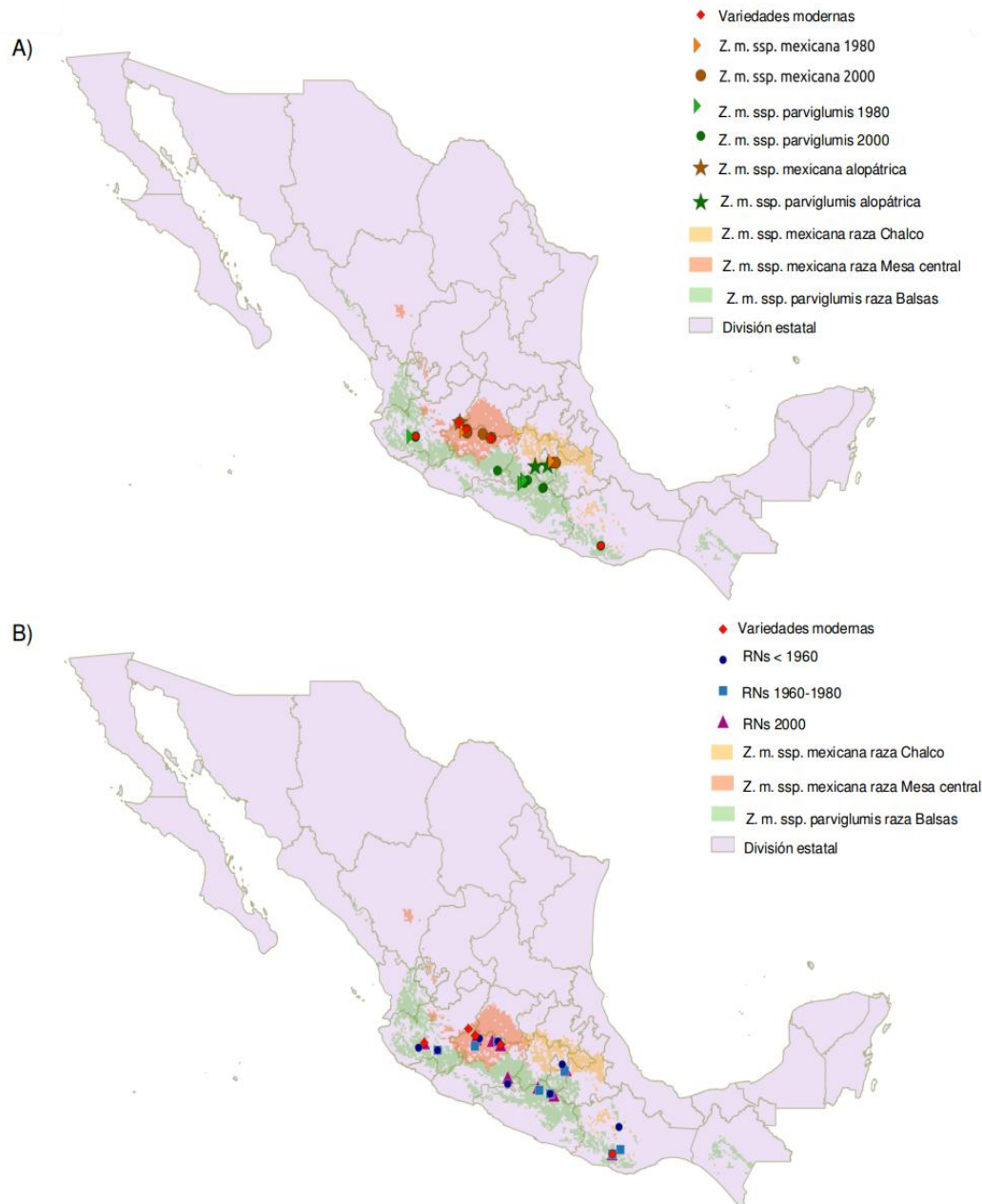


Figura 1. Distribución de las poblaciones y accesiones de razas (RNs), variedades modernas (VMs) y parientes silvestres (PSs) incluidas en este trabajo. Las áreas sombreadas indican la distribución potencial de las dos subespecies de PSs incluidos en el muestreo (Rojas-Barrera et al.,2019). A) El panel superior muestra las accesiones y colectas de PSs colectados en simpatria con VMs. B) El panel inferior muestra las colectadas y accesiones de RNs colectadas en simpatria con VMs. Los colores y figuras indican la agrupación temporal de las poblaciones y accesiones para ambos paneles.

Las accesiones colectadas antes de 1980 fueron solicitadas a los bancos de germoplasma del CIMMYT, del Instituto Nacional de Investigaciones Forestales y Agrícolas (INIFAP) y de la Universidad de Guadalajara. La información completa de las accesiones está disponible en el Anexo 1.

## 2.2 Extracción de DNA y genotipificación

Las semillas de maíz y teocintles fueron germinadas en una cámara ambiental CONVIRON a 25°C con un periodo de luz/oscuridad de 12h/12h, respectivamente. La extracción de DNA se realizó con el kit de extracción de DNA DNeasy Plant Mini Kit (QIAGEN Inc., Valencia, CA, USA). Las muestras de DNA se enviaron al Instituto de Recursos Biotecnológicos de la Universidad de Cornell, para obtener los genotipos por secuenciación (GBS, por sus siglas en inglés) de acuerdo con el protocolo de Elshire et al. (2011) optimizado para maíz. Brevemente, las muestras de DNA fueron digeridas con la enzima de restricción *ApeKI*, posteriormente se ligaron oligonucleótidos con un identificador para cada muestra, las librerías se enriquecieron por PCR y se realizó una selección de fragmentos de un tamaño entre 170-350 pb. Entonces, las muestras fueron secuenciadas en masa, cargando 96 muestras por línea de secuenciación en la plataforma HiSeq2000 (secuenciación de cadena sencilla y 100 pb). El número final de muestras incluidas en este trabajo fue de 385 individuos: 190 RNs colectas entre 1943-2015, 47 VMs colectadas en 2014-2015, 71 individuos de *mexicana*, 65 de *parviglumis*, ambos PSs colectados entre 1978-2015, y 12 individuos de la especie *Zea diploperennis*, estos últimos se emplearon como grupo externo en el análisis de flujo genético (Tabla 1).

## 2.3 Descubrimiento de variantes

Los archivos Fastq se demultiplexearon con GBSx (Herten et al. 2015) y las secuencias fueron podadas con Trimmomatic 0.36 (Bolger, Lohse, & Usadel 2014). El alineamiento de las secuencias se realizó con Nextgenmap 0.5.3 (Sedlazeck, Rescheneder, & von Haeseler 2013) utilizando como referencia el genoma de B73 (AGPv4) (<https://www.maizegdb.org/assembly>); posteriormente, los alineamientos se convirtieron a archivos binarios con samtools 1.5 (Heng et al. 2009). Las variantes fueron descubiertas con el módulo HaplotypeCaller, y los

genotipos individuales se fusionaron con GenotypeGVCFs; ambas herramientas forman parte del Genome Analysis Toolkit (GATK 3.8.0) (McKenna et al. 2010). El número final de polimorfismos de un sólo nucleótido (SNPs, por sus siglas en inglés) descubiertos fue de 4736306.

Los SNPs fueron filtrados con vcftools 0.1.15 (Danecek et al. 2011), considerando una frecuencia alélica mínima (MAF) de 0.01, un máximo de datos faltantes de 0.2 por SNP y una profundidad promedio mínima de 2X. Posterior a los filtrados calculamos la profundidad promedio por sitio, la profundidad promedio por individuo y la profundidad promedio acumulada por sitio (Tabla 2). Al final de este proceso conservamos un conjunto final de 316,294 SNPs distribuidos en los 10 cromosomas y en los contigs del genoma de referencia, los scripts de procesamiento de datos están disponibles en github (<https://goo.gl/p6pjNC>). Con los SNPs filtrados evaluamos el desequilibrio de ligamiento (LD) en ventanas de 100 kb con vcftools 0.1.15 (Danecek et al. 2011).

## 2.4 Análisis de ancestría

La ancestría de los genotipos se infirió con el paquete Admixture 1.3.0 (Alexander, Novembre, & Lange 2009). El análisis incluyó las 385 muestras reportadas en este trabajo, y evaluamos valores de grupos genéticos ( $K$ ) desde 1 hasta 8, y se seleccionó la  $K$  con el error de validación cruzada más bajo (EVC). El resultado del análisis de ancestría se usó como criterio para remover las accesiones de TbTp2014, TbTp2015 colectadas como *parviglumis* en la región de Jalisco, pero asignadas al grupo de *mexicana*, también se descartaron dos individuos que inicialmente se clasificaron como la RN Vandéño, pero que se agrupan con las VMs, y se descarto el uso de las colectas de Amatlán de Quetzalcóatl y Malinalco como poblaciones alopátricas de *parviglumis*, pues presentaron un alto grado de hibridación o ancestría compartida con *mexicana* (Fig. 4).

Para evaluar el origen de las VMs colectadas en simpatria con RNs y VMs, realizamos un análisis de ancestría con 463 genotipos incluidos en (Romay et al. 2013), considerando  $K$  desde 1 hasta 20. En este trabajo Romay y colaboradores analizaron los perfiles genéticos, obtenidos por GBS, de 2815 accesiones de líneas endogámicas generadas por programas de mejoramiento alrededor de todo el mundo.

## 2.5 Evaluación del flujo genético (ABBA-BABA)

Para evaluar el flujo genético de las VMs hacia las RNs y sus PSs, empleamos la prueba de ABBA-BABA (Green et al. 2010; Durand et al. 2011). Los archivos de entrada para realizar esta prueba fueron los archivos bam de los genotipos individuales, a diferencia de los archivos VCF que sólo contienen los sitios variables, los archivos bam están conformados por las secuencias de GBS alineados al genoma de referencia en formato binario.

El análisis de ABBA-BABA se basa en una filogenia resuelta entre cuatro taxa (((H1,H2),H3),H4) y evalúa si la proporción de estados derivados está influenciada por el flujo genético (Green et al. 2010). El análisis se corrió con los archivos bam y el módulo multipop ABBA BABA del paquete ANGSD (Analysis of Next Generation Sequencing Data; <https://github.com/ANGSD/angsd>) (Korneliussen, Albrechtsen, & Nielsen 2014); los parámetros usados están disponibles en <https://goo.gl/p6pjNC>. Los resultados se consideraron significativos para  $p < 0.05$  después de efectuar la corrección de Bonferroni.

En el modelo utilizado, las RNs colectadas en diferentes periodos de tiempo se asignaron a las posiciones H1 y H2, para evaluar si las muestras de una misma raza muestreada en diferentes periodos compartían un mayor número de alelos derivados en comparación con las VMs; las cuales se asignaron a la posición H3. *Zea diploperennis* se asignó al grupo H4 y se consideró como grupo externo. En presencia de flujo genético, el árbol evaluado presenta un exceso de patrones: ABBA o BABA. Un exceso de estados BABA en el genoma indica flujo de H3 a H1, mientras que el exceso de patrones ABBA indica flujo de H3 a H2 (Durand et al. 2011; Kulathinal, Stevison, & Noor 2009; Green et al. 2010).

Utilizamos una estrategia similar para los PSs, *mexicana* y *parviglumis*, los individuos colectados en diferentes periodos se asignaron en las posiciones H1 y H2, y las VMs y RNs simpátricas fueron asignadas en las posiciones H3. Las poblaciones alopátricas de los PSs se utilizaron como control para evaluar el efecto de la simpatría. Adicionalmente, realizamos análisis para cada unidad experimental representada por una RN, un PS y una o más VM, todos ellos en simpatría. Para las tres RNs que no fueron encontradas en simpatría con VMs, empleamos como grupo H3 a las



accesiones ZmH12015 para evaluar flujo genético del grupo 1 de variedades mejoradas (VM1) y la accesión SJH12015 para evaluar el flujo genético del grupo 2 (VM2).

## 2.6 Análisis de agrupamiento

Para explorar el origen del germoplasma usado para generar las VMs muestreadas en simpatria con RNs y los PSs, realizamos un llamado de variantes con Tassel 5 (Bradbury et al. 2007) y el mapa de variantes Tags On Physical Map v2.7 (<http://www.cyverse.org/>), que utiliza la versión 2 del genoma de B73 (AGPv2). Los genotipos generados en este trabajo fueron fusionados y comparados con un conjunto de 2578 genotipos obtenidos por GBS disponibles en la versión AGPv2, que corresponden a la colección de líneas endogámicas del Banco de semillas de Estados Unidos (Romay et al. 2013). Los datos fueron filtrados con un MAF 0.01 y un máximo de datos faltantes igual al 20% por SNP, el número final de variantes fue de 13953 SNPs.

Realizamos un primer análisis de componentes principales (PCA) usando el paquete de R SNPRelate v1.12.0 (Zheng et al. 2012) con las RNs y las VMs, y el primer conjunto de SNPs descrito en la sección 2.3. Posteriormente, para explorar las relaciones genéticas entre las VMs incluidas en este trabajo y las VMs generadas por otros programas de mejoramiento, realizamos tres PCAs adicionales con el conjunto de 13953 SNPs (resultado de la fusión de los genotipos incluidos en este trabajo con los generados por Romay et al. 2013).

Adicionalmente, construimos dos árboles de máxima verosimilitud basados en un modelo de evolución molecular GTR (Generalised Time-Reversible), con 1002 muestras y los 13 953 SNPs usando el paquete FastTree 2.1.9 (Price, Dehal, & Arkin 2009). Los árboles fueron graficados con Figtree v1.4.2. El primer árbol incluyó a los grupos VM1, VM2 y aquellos genotipos que estaban asignados a un programa de mejoramiento (Romay et al. 2013). El segundo árbol se realizó con 463 genotipos, que incluyeron a VM1, VM2 y los acervos tropicales generados por los programas de mejoramiento de México, Nigeria y Camerún, ambos árboles se enraízan con *parviglumis*.

## 2.7 Divergencia ( $F_{ST}$ ) a lo largo del genoma de RNs, VMs y PSs

Para evaluar la divergencia poblacional entre VMs, RNs y PSs, calculamos el índice  $F_{ST}$ , con los SNPs bialélicos, utilizando scripts personalizados disponibles en <https://goo.gl/Fuuho4>, <https://goo.gl/bnQxcc>. Los promedios de los valores de  $F_{ST}$  fueron calculados en ventanas de 5 kb; el tamaño de las ventanas se basó en el LD reportado para maíz (Yan et al. 2009). Posteriormente, para evaluar los cambios en el índice de divergencia ( $F_{ST}$ ) a través del tiempo, realizamos dos pruebas no paramétricas Kruskal-Wallis, seguidas de una prueba pareada U de Mann-Whitney. Finalmente, se graficaron los valores de  $F_{ST}$  en ventanas de 10 Mb por cromosoma para facilitar la visualización (Fig.3).

## 2.8 Diversidad genética

Los estimados de  $\pi$  y  $\theta_w$  se realizaron con los archivos bam, que contienen las secuencias alineadas al genoma de referencia en estado binario.

Las RNs y los PSs se agruparon en tres periodos, y para cada uno se calculó la diversidad nucleotídica ( $\pi$ ), la theta de Waterson ( $\theta_w$ ) y la D de Tajima con ANGSD (Korneliussen, Albrechtsen, & Nielsen 2014), en ventanas de 20, 50 y 100 kb y pasos de 5, 10 y 25 kb, respectivamente (Tabla S3). Las distribuciones de  $\pi$  y  $\theta_w$ , fueron comparadas para evaluar su cambio a través del tiempo. Posteriormente, realizamos una prueba no-paramétrica de Kruskal-Wallis, seguida de una prueba pareada U de Mann-Whitney. La significancia se estableció con una  $p < 0.05$  seguida de la corrección de Bonferroni. La distribución de las frecuencias calculadas en ventanas de 100 kb y pasos de 25 kb fue graficada con ggplot2\_3.1.1 (Wickham 2016). Adicionalmente, calculamos los valores de heterocigosis para cada una de las accesiones con el paquete hierfstat versión 0.04-22 de Rstudio (Anexo 3).

### 3 Resultados

#### 3.1 El conjunto de polimorfismos se distribuye a lo largo de los diez cromosomas de las RNs, VMs y PSs

Respecto al llamado de variantes, el número final de SNPs antes del filtrado fue de 4 736 306. Este conjunto de SNPs se filtró y se conservaron aquellas variantes con un maf= 0.01, considerando un número final de 385 genotipos diploides, un SNP con un maf equivalente a 0.01 tendrá al menos 7 copias. La proporción de datos faltante por SNP fue de 0.2, por tanto, las variantes incluidas estuvieron presentes en al menos el 80% de los genotipos. La profundidad promedio por sitio antes del filtrado fue 2.465 x, basado en este valor, el umbral de corte para la profundidad mínima por sitio se fijó en 2x (cuantil 70) lo que excluyó ~70 de SNPs descubiertos. Posterior a los filtrados, el conjunto de datos se redujo a 316 294 SNPs distribuidos en los 10 cromosomas y contigs. La profundidad promedio por individuo después del filtrado fue de  $8.53 \pm 3.34$ ; la profundidad promedio acumulada por sitio (que incluye las lecturas registradas en todos los genotipos) fue de  $3266.3 \pm 1669.68$  y la profundidad promedio por sitio fue de  $8.53 \pm 4.56$  (Tabla 2). Dado que el conjunto de SNPs obtenidos se obtuvo por medio de secuenciación, donde la profundidad de las lecturas es un parámetro que priorizar (Colin McEwan, 2016), omitimos el filtro de equilibrio Hardy-Weinberg.

Tabla 2. Parámetros promedio de profundidad por sitio e individuo para el conjunto de SNPs antes de aplicar los filtrados y después del filtrado.

	Datos crudos			Datos filtrados		
	Profundidad promedio por individuo	Profundidad acumulada por sitio	Profundidad promedio por sitio	Profundidad promedio por individuo	Profundidad acumulada por sitio	Profundidad promedio por sitio
Promedio	2.465	1074.174	2.465	8.538	3266.306	8.538
SD	0.996	1671.799	3.836	3.346	1669.685	4.560
Mediana	2.527	344.000	0.789	8.620	3030.000	7.943
SE	0.048	0.768	0.002	0.171	11.685	0.008
Cuartil 1	1.968	68.000	0.156	6.774	1988.000	5.236
Cuartil 3	3.065	1300.000	2.984	10.638	4296.000	11.208

La densidad de distribución de los SNPs en los diez cromosomas se muestra en la figura 2, los puntos rojos indican el punto medio del centrómero para cada uno de los diez cromosomas (Wolfgruber et al. 2009), donde la densidad de SNPs suele disminuir. El cromosoma 1 albergó el mayor número de SNPs: 50869 y el cromosoma 10 tuvo el menor número de SNPs: 20923, estos cromosomas son el de mayor y el de menor tamaño, respectivamente (Wei et al. 2007).

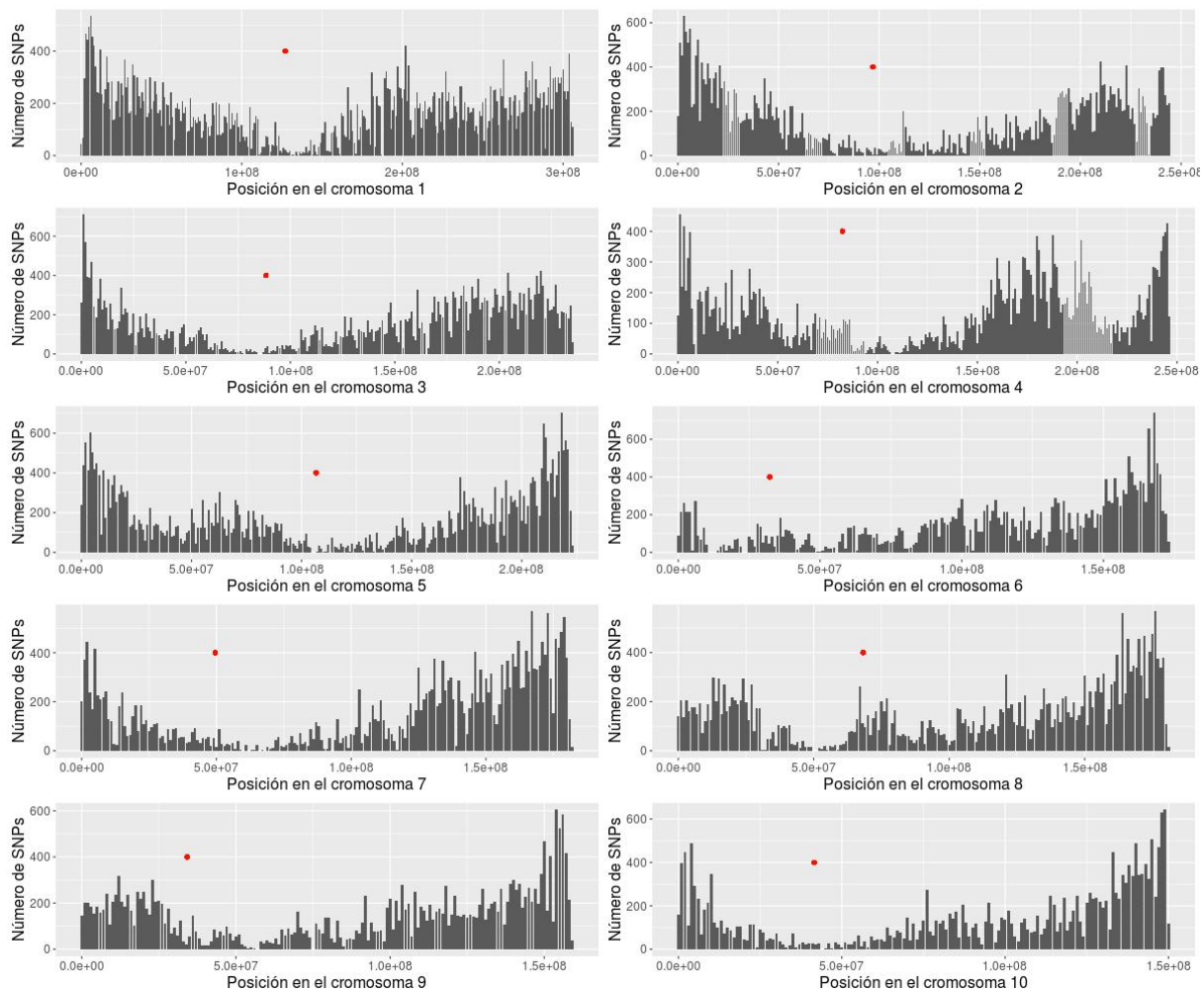


Figura 2. Densidad de los SNPs en ventanas de 1 Mb para los 10 cromosomas de maíz. El punto rojo muestra el punto medio del centrómero de cada cromosoma.

Posterior al filtrado, estimamos el desequilibrio de ligamiento (LD, por sus siglas en inglés) en ventanas de 100 kb, agrupando a las accesiones cultivadas en RNs, VMs; y a los PSs en *mexicana*, *parviglumis* y *Z. diploperennis* (Figura 3). El LD cayó por debajo de  $r^2=0.25$  para la mayoría de los marcadores antes de las 20 kbs para las RNs, *mexicana* y *parviglumis* y las VMs. En contraste, *Z. diploperennis* presentó un LD más alto en comparación con todos los grupos de *Z. mays*.

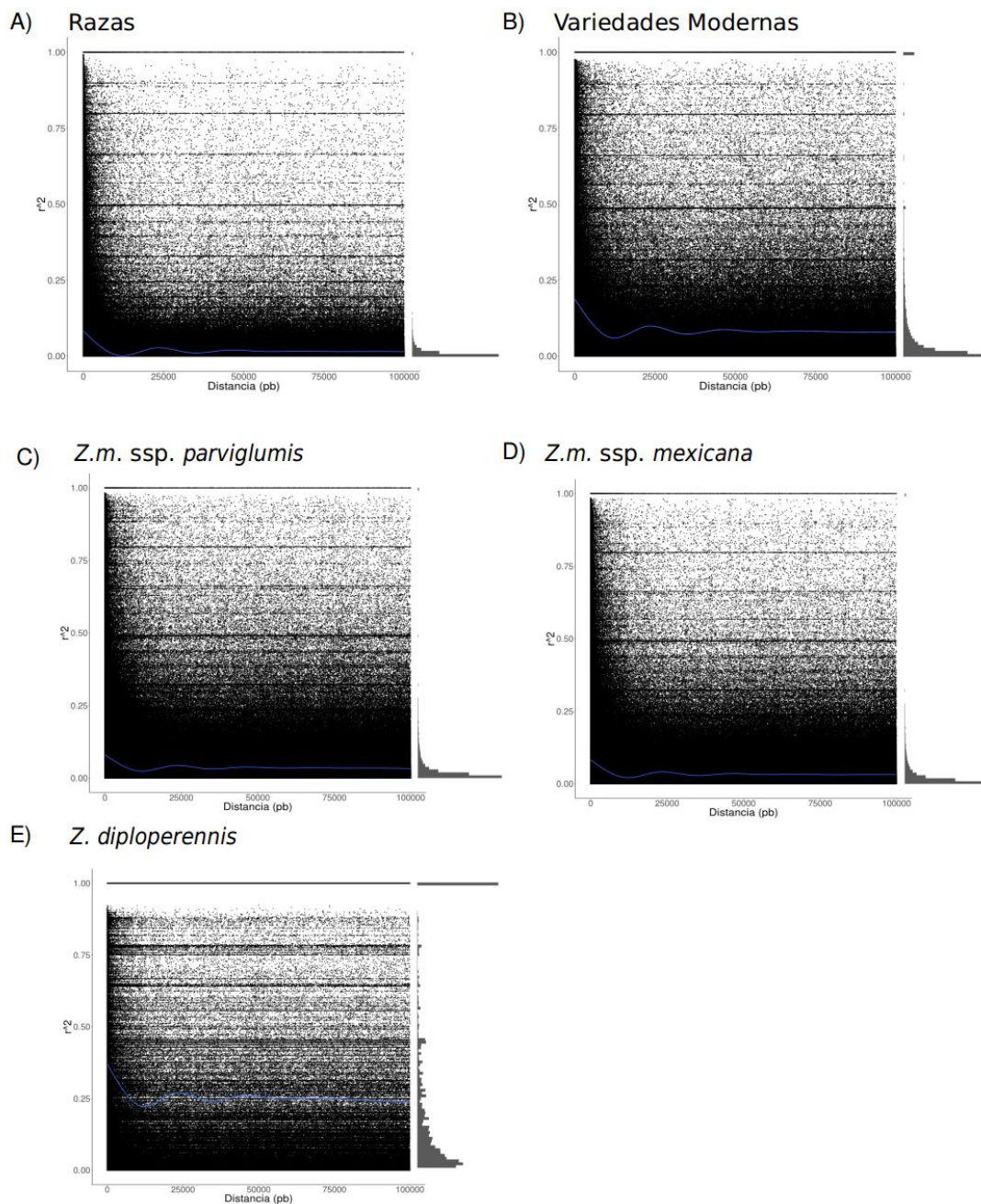


Figura 3. Desequilibrio de ligamiento ( $r^2$ ) calculado en ventanas de 100 kb para razas, variedades modernas, *parviglumis*, *mexicana* y *Z. diploperennis*.



### 3.2 La ancestría de los individuos cambia de acuerdo con la altitud y el periodo de colecta

El análisis de ancestría incluyó a las RN, VMs, y los PS (Fig. 4). El número de grupos genéticos con el error de validación cruzada (EVC) más bajo fue  $K=7$ . Las RNs y VMs se agruparon en tres ancestrías, la primera de ellas etiquetada como “Razas nativas” (azul, Fig. 4B) predominó en la raza de maíz Chalqueño que se distribuye en valles altos, usualmente por encima de los 2000 msnm, cabe señalar que Chalqueño no se encontró en simpatria con ninguna VM al momento del muestreo.

La segunda ancestría etiquetada como “Variedades modernas grupo 1” (rojo, Fig. 4) se compartió entre las RNs que distribuyen en regiones con una altitud menor a los 1900 msnm, y agrupó a 6 de las 9 VMs colectadas, que en adelante serán referidas como VM grupo 1 (VM1). Este grupo genético incluye individuos muestreados a lo largo de todo el rango temporal.

La tercer ancestría etiquetada como “Variedades modernas grupo 2” (amarillo, Fig. 4) incluye a las tres accesiones de VMs restantes, que en adelante serán referidas como VM grupo 2 (VM2). El grupo VM2 fue muestreado en las localidades de Churintzio, Michoacán y San Jerónimo, Jalisco. Su presencia es escasa en la raza Bolita de 1970 (0.06), pero después del 2000 a detectamos en 6 de las siete razas muestreadas (Fig. 5 B-C), en este grupo también encontramos a dos individuos colectados como raza Vandeño en el año 2015, por lo que fueron descartados de los análisis subsecuentes (Fig. 4).

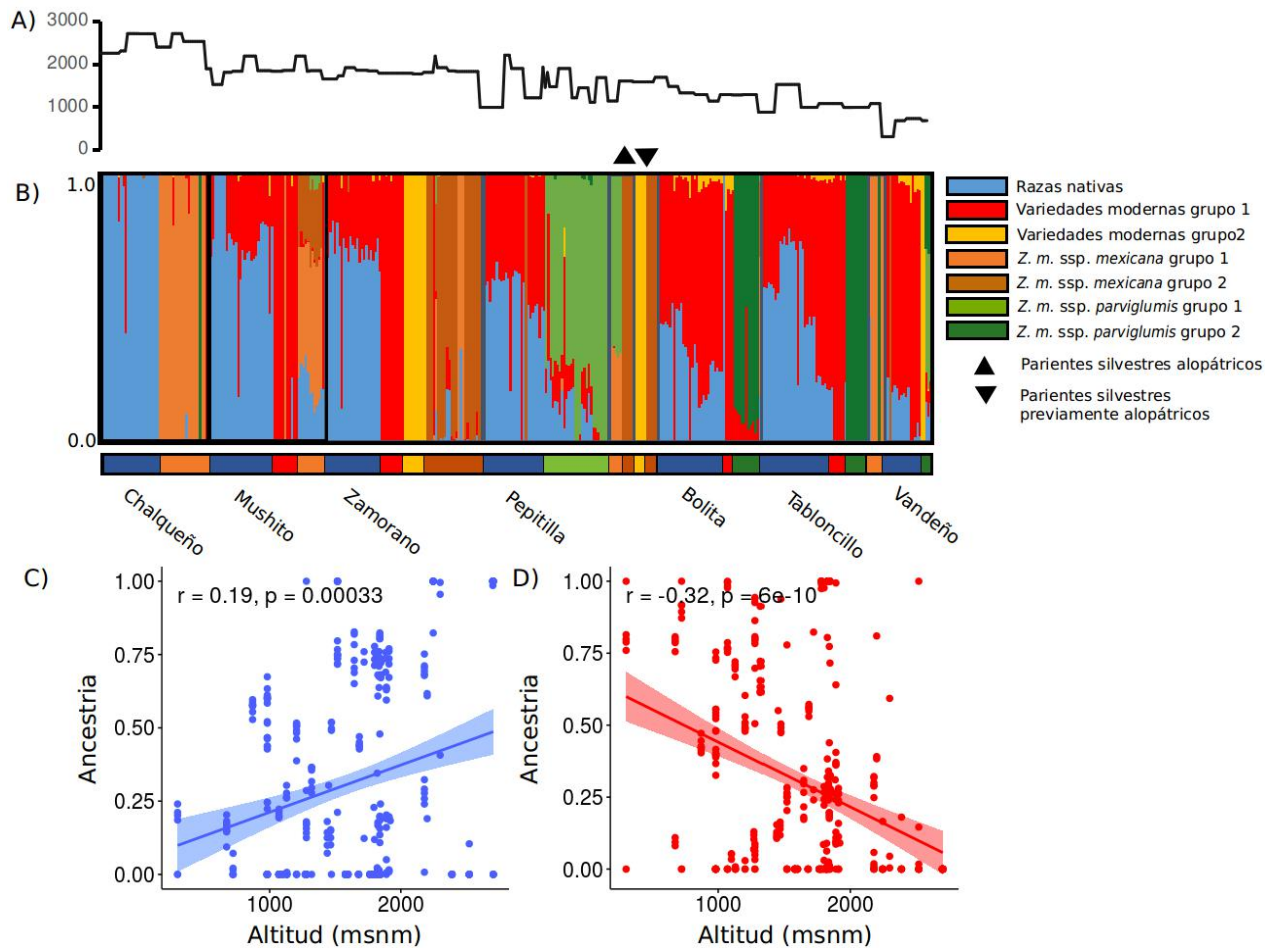


Figura 4. Estructura poblacional de las razas (RNs), variedades modernas (VMs) y parientes silvestres (PSs). A) Distribución altitudinal de las muestras incluidas en el análisis de estructura. B) Estructura poblacional de las siete unidades experimentales ordenadas de acuerdo con la altitud y el año de colecta de las más antiguas a las más recientes. Cada unidad de análisis fue nombrada de acuerdo con la raza colectada en simpatria con las variedades modernas y sus parientes silvestres, donde  $K=7$  tuvo el menor ECV. C) Correlación entre la ancestría predominante en Chalqueño. D) Correlación entre la ancestría predominante en razas de altitud media a baja y variedades modernas y la altitud (Rojas-Barrera et al. 2019).

Los teocintles *mexicana* y *parviglumis* se agruparon en 4 grupos genéticos (Fig. 4, dos tonos de naranja y dos tonos de dos tono de verde) cuya proporción concuerda con la distribución altitudinal reportada para *mexicana* y *parviglumis* (Sánchez González et al. 2018); excepto por los teocintles colectados en 2014 y 2015 en San Lorenzo, Jalisco, localizado a 1071 msnm. Estas colectas se agruparon con *Z. m. ssp. mexicana* grupo 1, a pesar de haber sido colectados en una región donde usualmente se distribuye *parviglumis* (Figura 4 y 5 G).

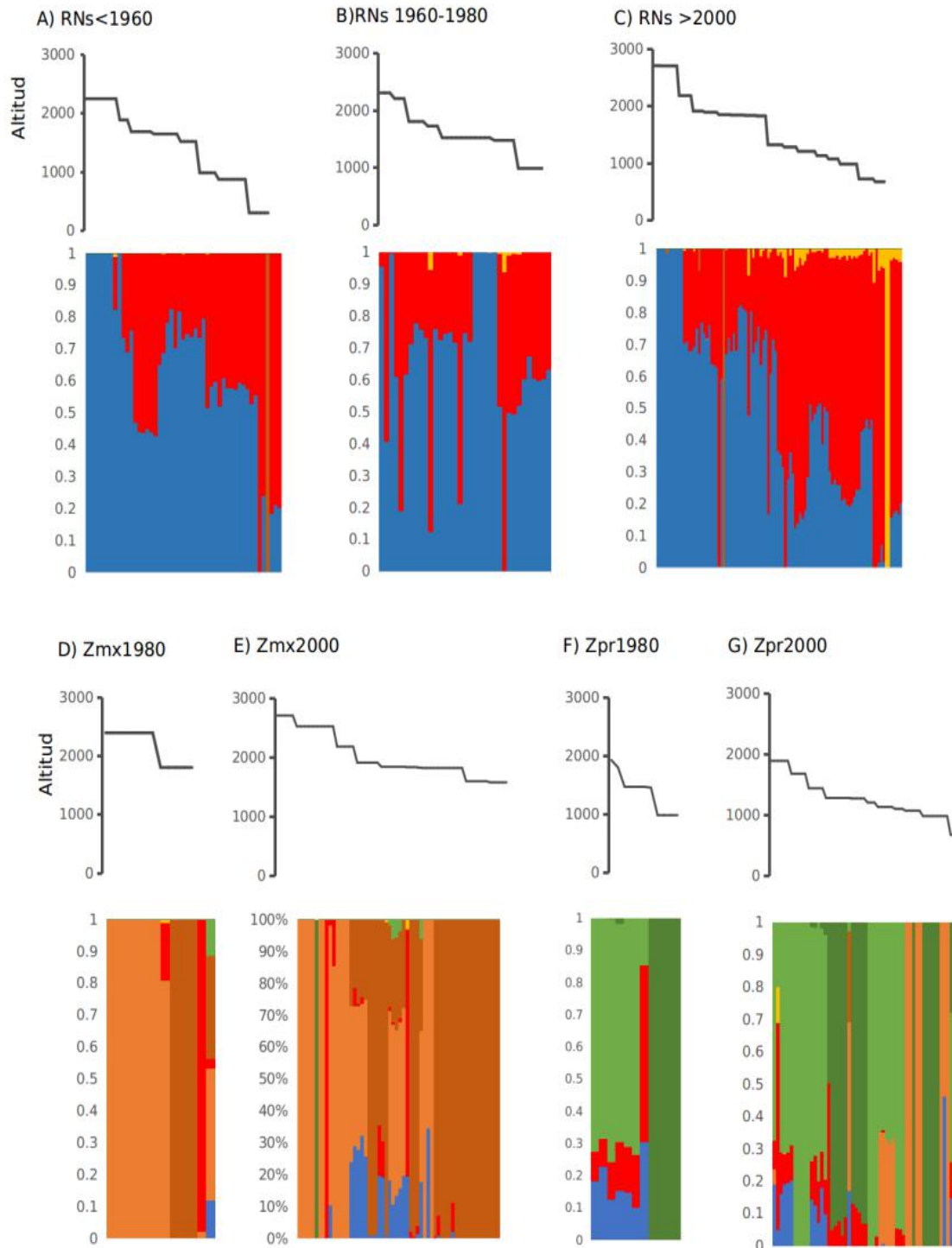


Figura 5. Rearreglo de los genotipos incluidos en el análisis de estructura poblacional. Las muestras están ordenadas de acuerdo con el periodo en que fueron colectados y a la clasificación a la que pertenecen: A) Razas nativas (RNs) colectadas antes de 1960, B) RNs colectadas entre 1960-1980, C) RNs colectadas después del 2000, D) *mexicana* colectados en 1980, E) *mexicana* colectados después del 2000, F) *parviglumis* colectado en 1980 y G) *parviglumis* colectado después del 2000.

La ancestría naranja etiquetada como *Z. m. ssp. mexicana* grupo 1 (Fig. 4) predominó en los individuos de *mexicana* muestreados en Chalco, Estado de México (>2400), y en San Agustín del Maíz, Michoacán (1800



msnm), aunque de manera inesperada, observamos que los individuos colectados como poblaciones alopátricas de parientes silvestres: *parviglumis* en las localidades de Malinalco, Edo. de México y Amatlán de Quetzalcóatl, Morelos; mostraron un alto grado de mestizaje, motivo por el cual decidimos descartar su uso como población alopátrica de *parviglumis* en los modelos de flujo genético. La ancestría etiquetada como *Z. m. ssp. mexicana* grupo 2 sólo fue detectada en teocintles colectados en la Mesa Central de México (Fig. 4B).

La ancestría etiquetada como *Z. m. ssp. parviglumis* grupo 1 agrupó a los teocintles *parviglumis* colectados en simpatria con Pepitilla, y también a las poblaciones de Malinalco, Edo. de México y Amatlán de Quetzalcóatl, Morelos (Fig. 4B). El grupo *Z. m. ssp. parviglumis* grupo 2 incluyó a los individuos de *parviglumis* colectados en San Lorenzo, Jalisco (~ 1000 msnm) antes del 2014 y a los muestreados en San Cristóbal Honduras, Oaxaca (1280 msnm). Al graficar las ancestrías etiquetadas como RNs y VM1 contra la altitud del sitio de colecta (4C-D), observamos una correlación significativa entre la ancestría y la altitud (Fig. 6). Esta correlación no está influenciada por un sesgo de muestreo de acuerdo con el periodo de colecta, pues el gradiente altitudinal está representado en todos los periodos de tiempo establecidos (Fig. 5).

Cuando reorganizamos los genotipos de acuerdo con los periodos de colecta descritos en la tabla 1 (Fig. 5). La ancestría asociada al grupo VM1 y RNs de altitudes medias y bajas (rojo) incrementa su proporción en el tiempo. Basados en la observación anterior, es probable que la ancestría asociada a VM2 (amarillo) siga la misma tendencia, como lo muestran las correlaciones entre los grupos genéticos y el año de colecta (Fig. 6). En contraste con las ancestrías asociadas a las VMs (rojo-2 y amarillo-3), observamos una correlación negativa entre la ancestría 1 (azul) y el periodo de colecta (Fig. 5), su proporción disminuyó en las RNs de altitudes media y baja colectadas después del año 2000. En el caso de *mexicana* y *parviglumis*, los grupos genéticos 1 (azul) y 2 (rojo) incrementan su proporción en las  $Z_{mx}>2000$  y en  $Z_{pr}>2000$  (Fig. 5D-G).

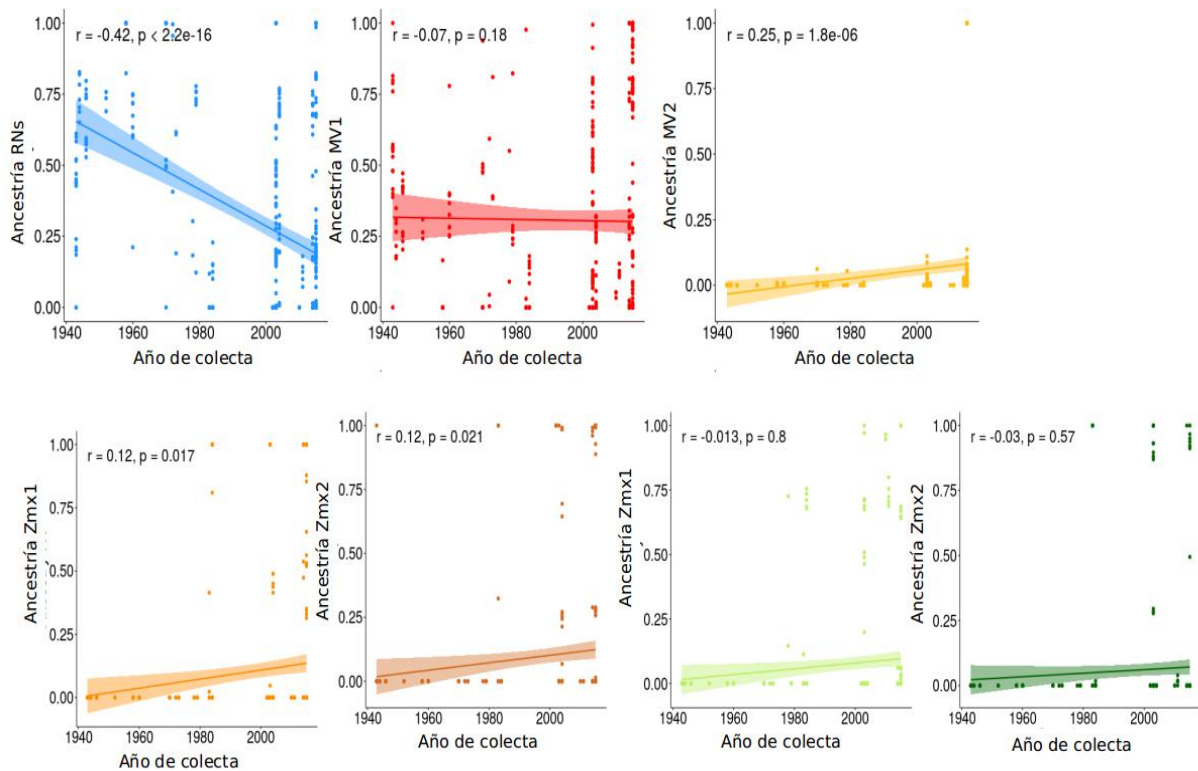


Figura 6. Correlación de los 7 grupos genéticos identificados para razas nativas (RNs), variedades modernas (VMs) y parientes silvestres.

### 3.2 Flujo genético de variedades mejoradas hacia razas nativas y parientes silvestres

Basados en el análisis de ancestría, que sugiere la presencia de flujo genético de las VMs hacia las RNs, evaluamos doce modelos para poner a prueba la presencia de flujo genético de los grupos VM1 y VM2 hacia RNs y PSs colectados después del 2000. Empleamos como referencia las accesiones de RNs colectadas antes de 1960 y los PSs colectados en la década de 1980 (Fig. 6). Los resultados obtenidos indican flujo genético de los dos grupos genéticos VM1/VM2 hacia las RNs colectadas después del 2000 (RNs >2000) ( $p < 0.0001$ ). Este resultado es independiente de la referencia usada: Razas colectadas antes de 1960 o Razas colectadas entre 1960-1980.

Los análisis de flujo genético realizados con los PSs también indicaron flujo de genes de VMs hacia *mexicana* y *parviglumis*. El valor del estadístico  $D$  de Patterson fue mayor para los individuos de *mexicana* colectados después del año 2000 (Zmx2000), aunque también es detectable el flujo de genes a los individuos de *mexicana* colectados en

1980 (Zmx1980); esto sucede cuando la referencia es una población alopátrica de *mexicana*. Para contrastar el flujo de VMs hacia *mexicana*, con el flujo de RNs hacia teocintles, las RNs fueron consideradas como el grupo donador (H3). El valor de  $D$  resultó mayor cuando el grupo donador son las RN>2000 en comparación con VM1 o VM2 (Anexo 2), lo que sugiere una interferencia con la señal de flujo genético de VMs hacia *mexicana*.

En el caso de *parviglumis*, observamos que la señal de flujo está dirigida a los individuos colectados en 1980 (Zpr1980) cuando son comparados con los colectados después del año 2000 (Zpr2000), a diferencia de *mexicana*, no contamos con una población alopátrica para *parviglumis* que nos permita discriminar entre las condiciones de alopatría y simpatría. Las poblaciones alopátricas de Amatlán, Morelos y Malinalco, Edo. de México mostraron una mezcla de ancestrias (Fig. 4B), lo que nos impidió asignarlas con certeza al grupo de *parviglumis*.

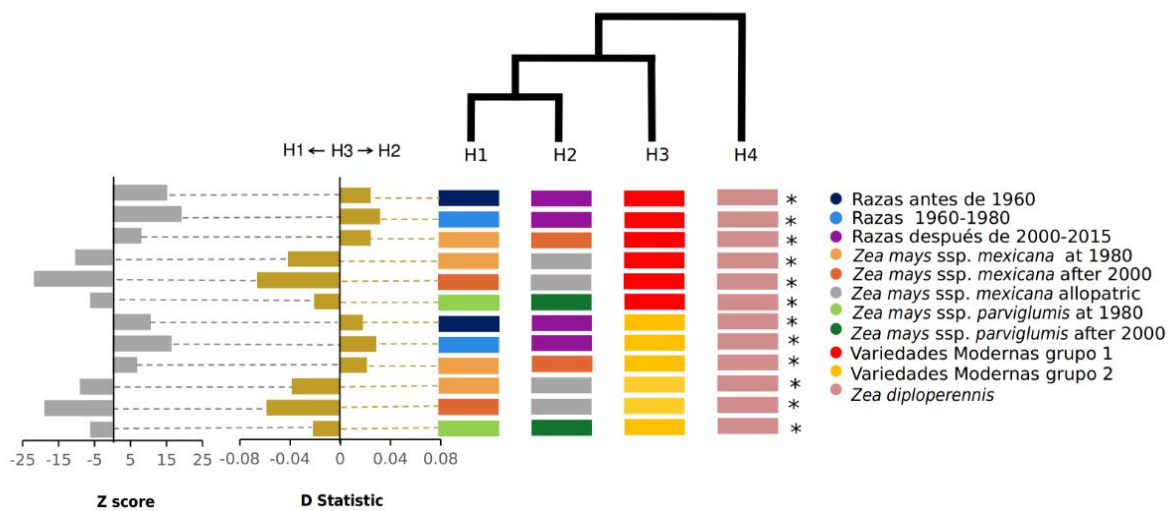


Figura 7. Modelos de flujo genético de variedades modernas hacia razas y parientes silvestres simpátricos, el \* denota significancia estadística ( $p < 0.0001$ ). Un valor positivo del estadístico  $D$  de Patterson indica flujo de H3 a H2; mientras que un valor negativo denota flujo de H3 a H1 (Rojas-Barrera et al. 2019).

Cuando realizamos los análisis por unidad experimental (una RN, un PS y una o más VMs), detectamos flujo ambos grupos de VMs hacia las seis RNs que se distribuyen por debajo de los 2000 msnm, la excepción fue la RN Chalqueño muestreada entre 2200 y 27000 msnm (Fig. 8, Anexo 2) para la que no observo introgresión.

En este trabajo nos enfocamos en un conjunto de RNs que son simpátricas con VMs, para incrementar la probabilidad de detectar flujo genético. Sin embargo, nuestros resultados indican que la simpatría al momento de colecta no es necesaria para observar introgresión. Encontramos flujo de las VMs hacia las RNs Vandeño y Pepitilla en la ausencia de VMs en las localidades muestreadas (Fig. 8, Anexo 2), lo que implica que el flujo genético detectado ocurrió en generaciones previas, o es consecuencia de la simpatría con VMs acriolladas, como es el caso de los dos individuos colectados como Vandeño, y que fueron genéticamente asignados al grupo VM2.

Al evaluar el flujo genético hacia los PSs por unidad experimental, la señal de flujo genético de VM1 y VM2 hacia *mexicana* y *parviglumis* varió de acuerdo con el año de colecta y a la distribución geográfica de las poblaciones de teocintles. Cuando la referencia fueron las localidades de *mexicana* de 1980 (H1), sólo detectamos flujo genético hacia los individuos de *mexicana* colectados en 2015 en la zona de Chalco (Fig. 7, Anexo 2). En cambio la comparación entre las poblaciones de *mexicana* 2004 (H1) y 2015 (H2), indicó flujo de ambos grupos de VMs hacia los individuos de *mexicana* de 2004 colectados en la Meseta Central (Anexo 2).

En el caso de *parviglumis*, cuando la referencia fueron las diferentes poblaciones de *parviglumis* 1980 (H1), el valor de  $D$  fue negativo para todas ellas, lo que concuerda con los modelos en los que agrupamos a todas las poblaciones en los tres periodos evaluados. Esto también apoya la introgresión de VMs hacia las poblaciones colectadas en la década de 1980 (Fig. 7, Anexo 2).

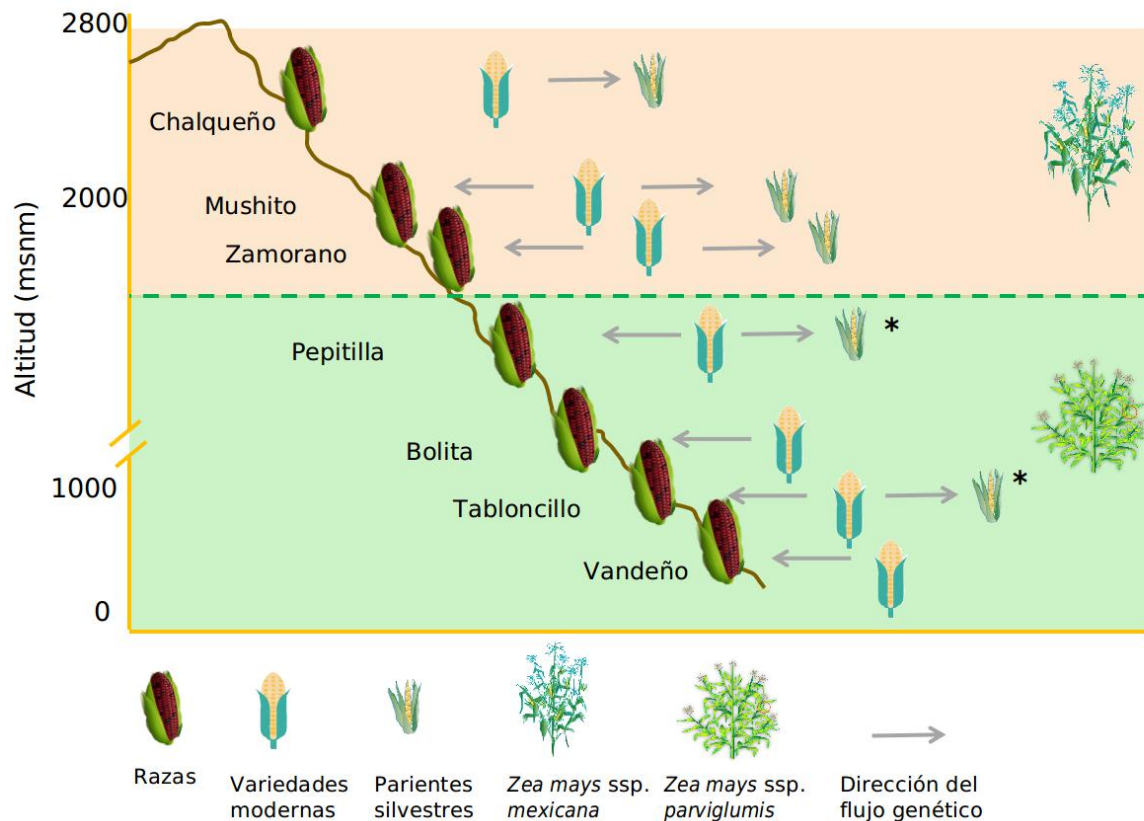


Figura 8. Resultados de los modelos de flujo genético para las siete unidades experimentales distribuidas en un gradiente altitudinal. Cuando evaluamos el flujo genético de las variedades modernas (VMs) hacia cada una de las siete razas (RNs), detectamos flujo genético para todas las razas, excepto para Chalqueño. En el caso de los parientes silvestres (PSs), detectamos flujo genético hacia las poblaciones de *Z. m. ssp. mexicana* colectadas después del año 2000. Para *Z. m. ssp. parviglumis* sólo detectamos flujo genético en las poblaciones de 1980 (\*). El fondo naranja claro indica que las RNs y las VMs son simpátricas con *Z. m. ssp. mexicana* y el fondo verde claro indica que son simpátricas con *Z. m. ssp. parviglumis*.

### 3.3 La divergencia ( $F_{ST}$ ) entre VMs con RNs y PSs decae en el tiempo

El índice de divergencia pareado ( $F_{ST}$ ) fue usado para evaluar la diferenciación a través del tiempo entre los dos grupos genéticos de VMs y las RNs, *mexicana* y *parviglumis* (Fig. 9). Para todas las estimaciones pareadas observamos que la diferenciación entre VM-RNs y VM-PSs disminuye de manera significativa en el periodo evaluado. Adicionalmente, las estimaciones de  $F_{ST}$  entre VM1-RNs, VM1-*mexicana* y VM1-*parviglumis* son significativamente menores que la divergencia con el grupo VM2; este resultado es consistente con la presencia temprana de VM1 y una mayor proporción de las líneas asociadas con esta ancestría (Fig. 4-5).



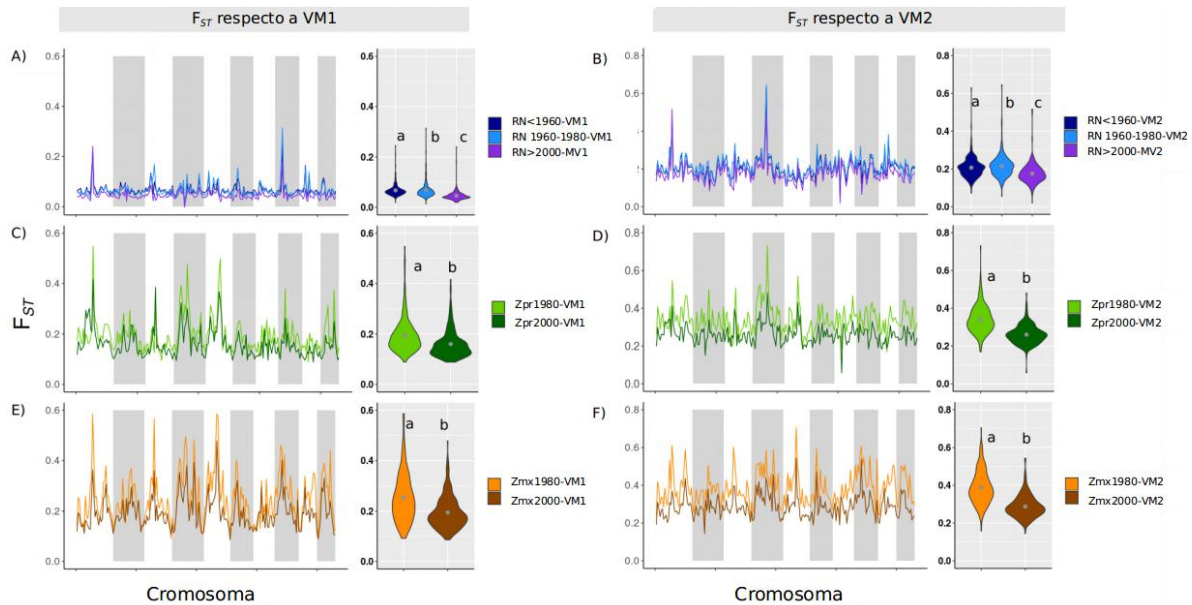


Figura 9. Divergencia poblacional ( $F_{ST}$ ) a lo largo del genoma entre razas (RNs), y parientes silvestres (PSs) colectados en diferentes periodos, comparados con los grupos genéticos VM1 y VM2 de las variedades modernas simpátricas. Las áreas sombreadas y no sombreadas alternadas representan los 10 cromosomas de maíz; y las letras minúsculas sobre los gráficos de violín denotan una diferencia con significancia estadística  $p < 0.01$ ) entre los grupos comparados por periodo de colecta (Rojas-Barrera et al. 2019).

### 3.4 MV1 y MV2 se agrupan con los acervos genéticos tropicales desarrollados en México y Camerún

Con el fin de identificar el origen de las VMs que fueron colectadas en simpatria con las RNs y los PSs y para evaluar si estas variedades están basadas en germoplasma nacional o exótico. En 2015 solicitamos información acerca del origen y el proceso de obtención de las semillas registradas ante el SNICS. Nuestra solicitud recibió una respuesta negativa, cito textualmente: *“dicha información está protegida por las Leyes Nacionales e Internacionales en materia de propiedad intelectual, así como en materia de Transparencia y Acceso a la Información Pública Gubernamental”*. En ausencia de esta información acerca de los pedigríes de las semillas, realizamos análisis de agrupamiento con un conjunto de genotipos públicos que corresponden a diferentes programas de mejoramiento (Romay et al. 2013), con el objetivo de darnos una idea sobre las ancestrías de las VMs colectadas en este trabajo.

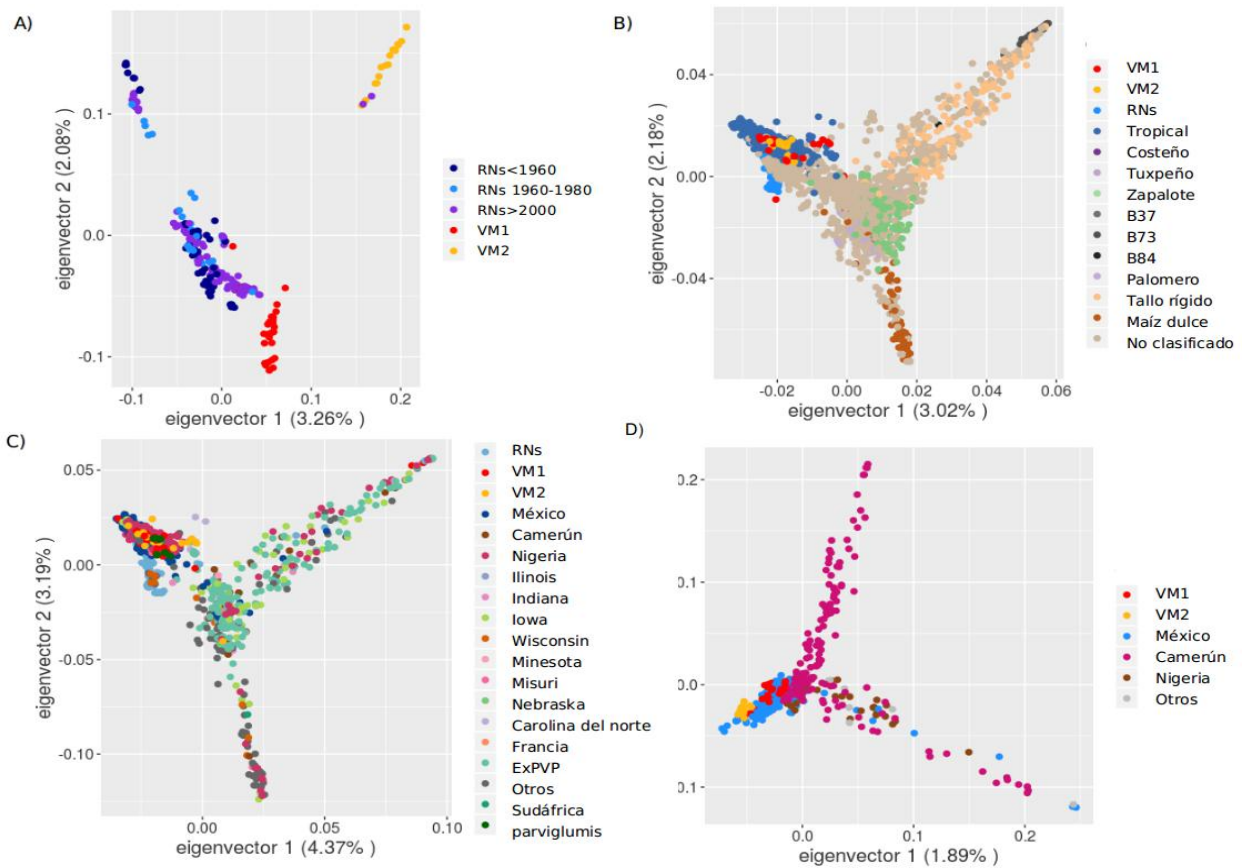


Figura 10. Análisis de componentes principales de las razas, variedades modernas y un subconjunto de genotipos del banco de germoplasma de Estados Unidos (Romay et al. 2013). A) Razas colectadas entre 1943-2015 en simpatria con variedades modernas (VMs); B) VMs simpátricas con razas y un conjunto de genotipos del Banco Nacional de Líneas Endogámicas de Maíz de EUA, coloreados de acuerdo con la colección a la que pertenecen; C) VMs simpátricas con razas y un subconjunto de genotipos coloreados de acuerdo con el programa de mejoramiento que lo generó; y D) VMs simpátricos con los genotipos tropicales coloreados por programa de mejoramiento (Rojas-Barrera et al. 2019).

En el primer análisis de componentes principales (PCA) el *eigenvector* 1 separó al grupo VM2 del resto de las muestras, lo que implica que la distancia entre VM1 y VM2 es mayor que la distancia de VM1 a las RNs, sin importar el periodo de colecta de las RNs (Fig. 9A). En los PCA posteriores graficamos la distribución de las RNs y VMs con los genotipos de Romay et al. 2013 que representan accesiones resguardadas en el banco de germoplasma de Estados Unidos de América (EUA). En este caso VM1 y VM2 cayeron dentro de la distribución de aquellos genotipos identificados como RNs y acervos tropicales (Fig. 9B).

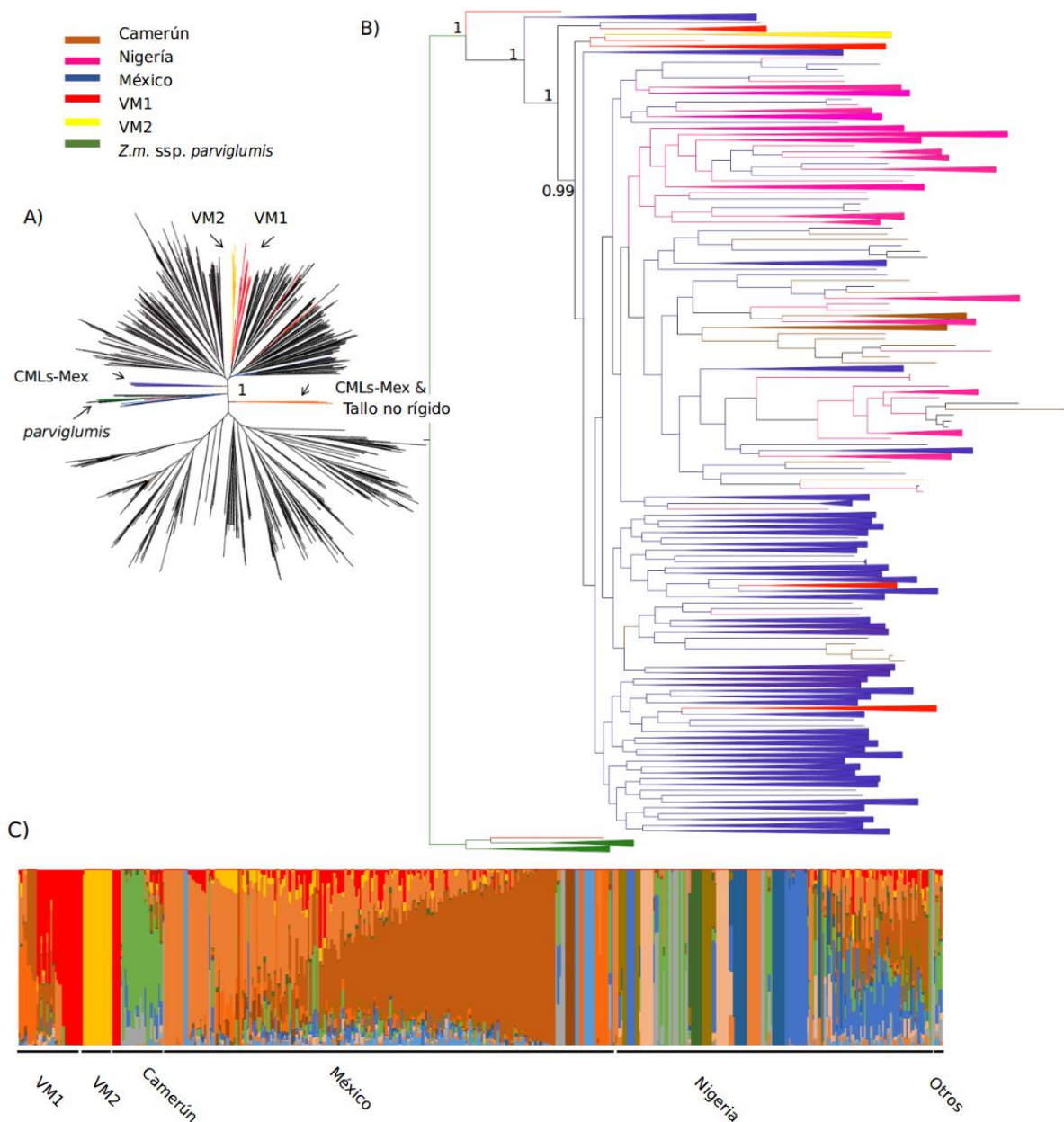


Figura 11. Análisis de ancestría de los grupos genéticos VM1 y VM2. A) Árbol de máxima verosimilitud con VM1, VM2 y los genotipos de Romay *et al.* (2013). B) Árbol de máxima verosimilitud con que incluye VM1, VM2 y a los genotipos agrupados en la rama superior de 7A donde los números indican el soporte estadístico de las ramas. C) Análisis de ancestría de VM1, VM2, y los genotipos de los programas de mejoramiento tropicales desarrollados en Camerún, México y Nigeria, entre otros (Rojas-Barrera *et al.* 2019).

Para mejorar la resolución del análisis, removimos los genotipos sin clasificación y conservamos aquellos que estaban asociados a un programa de mejoramiento; en esta ocasión el *eigenvector* 1 agrupó a las RNs, VM1 y VM2 con los genotipos de los programas de mejoramiento de México, Camerún y Nigeria (Romay *et al.* 2013) (Fig. 9C). Finalmente, tomamos el subconjunto de accesiones clasificadas como tropicales



generadas por algún programa de mejoramiento; en este caso VM1 y VM2 se agruparon principalmente con los genotipos del programa de mejoramiento mexicano, y con un subconjunto de individuos de Nigeria y Camerún (Fig. 9D).

Luego, para explorar la ancestría de VM1 y VM2, realizamos dos árboles de máxima verosimilitud basados en un modelo de sustitución GTR (Fig. 10). En ambos árboles *parviglumis* fue empleado como grupo externo. El primer árbol mostró, con un soporte de 1.0, que un conjunto de líneas de maíces del CIMMYT (CMLs) son basales a la sección que contiene a VM1 y VM2 (Fig. 10A). El segundo árbol, que sólo incluyó a los maíces de acervos genéticos tropicales de México, Nigeria y Camerún ubicó nuevamente a las CMLs (soporte = 1.0) como el grupo basal a los tres grupos genéticos (Fig. 10B). Después de la cuarta bifurcación el soporte es menor a 0.99, aunque en ramas interiores dos accesiones del grupo VM1 se agrupan con accesiones provenientes del programa de Nigeria.

Basados en lo anterior realizamos un análisis de ancestría con los genotipos presentes en el segundo árbol (Fig. 10C), donde  $K=15$  tuvo el menor EVC. Los genotipos incluidos fueron poblaciones muy estructuradas y con alto grado de mestizaje, VM2 formó un grupo genético homogéneo (amarillo, Fig. 10), mientras que VM1 presentó mayor grado de mezcla genética y estructura poblacional. Los grupos VM1 y VM2 comparten ancestría con el programa de mejoramiento genético de México, aunque un subconjunto de genotipos de Camerún se agrupó con VM1. Estos resultados, en conjunto con las evidencias anteriores muestran que las CMLs del programa de mejoramiento mexicano son ancestrales para VM1 y VM2.

### 3.5 La diversidad genética de las RNs y PSs es mayor para los genotipos colectados después del 2000

La diversidad nucleotídica ( $\pi$ ) de las RNs disminuyó en el periodo intermedio de 1960-1980 ( $1.84e-04$ ) en comparación con las RN 1960 ( $1.87e-04$ ), aunque incrementó de manera significativa para las RN > 2000 ( $2.03e-04$ ) (Fig. 11, Tabla 3). En el caso de *mexicana* y *parviglumis*, no tenemos genotipos colectados en la década de 1960, pero el valor de  $\pi$  resultó significativamente mayor para los individuos de ambas

subespecies colectados después del año 2000, en comparación con las colectas de 1980 (Tabla 3). Como era de esperarse, ambos parientes silvestres presentaron valores de diversidad nucleotídica más altos que las RNs. A su vez,  $\pi$  es mayor en las RNs que en las VMs, la diferencia es mayor respecto a VM2, cuyo promedio de diversidad es una orden de magnitud menor que el de las RNs y las VM1 (Tabla 3).

Tabla 3. Diversidad nucleotídica ( $\pi$ ) por pb, calculada en ventanas de 20, 50 y 100 kb, y pasos de 5, 10 y 25 kb, respectivamente; para las razas (LR), las subespecies silvestres estudiadas: *mexicana* (Zmx), *parviglumis* (Zpr), y las variedades modernas (MVs) de maíz. Nota: el valor de  $p$  para el panel inferior es menor (Rojas-Barrera et al. 2019).

Tamaño de ventana (kb), tamaño de paso (kb)	20,5 (Promedio)	20,5 (sd)	Significancia estadística ( $p < 0.001$ )	50,10 (Promedio)	50, 10 (sd)	Significancia estadística ( $p < 0.01$ )	100, 25 (Promedio)	100, 25 (sd)	Significancia estadística ( $p < 0.05$ )
LR <1960	1.87E-04	4.01E-04	a	1.87E-04	2.86E-04	a	1.87E-04	2.24E-04	a
LR 1960-1980	1.84E-04	3.92E-04	b	1.84E-04	2.80E-04	b	1.84E-04	2.19E-04	b
LR >2000	2.03E-04	4.30E-04	c	2.03E-04	3.08E-04	c	2.03E-04	2.41E-04	c

Tamaño de ventana (kb), tamaño de paso (kb)	20,5 (Promedio)	20,5 (sd)	Significancia estadística ( $p < 2e-16$ )	50,10 (Promedio)	50, 10 (sd)	Significancia estadística ( $p < 2e-16$ )	100, 25 (Promedio)	100, 25 (sd)	Significancia estadística ( $p < 2e-16$ )
Zpr 1980	1.94E-04	4.15E-04	a	1.94E-04	2.94E-04	a	1.94E-04	2.29E-04	a
Zpr >2000	2.45E-04	5.11E-04	b	2.45E-04	3.64E-04	b	2.45E-04	2.83E-04	b
Zmx 1980	2.05E-04	4.40E-04	a	2.05E-04	3.13E-04	a	2.05E-04	2.44E-04	a
Zmx >2000	2.45E-04	5.11E-04	b	2.45E-04	3.64E-04	b	2.45E-04	2.83E-04	b
MV1	1.44E-04	3.19E-04	a	1.44E-04	2.28E-04	a	1.44E-04	1.79E-04	a
MV2	8.08E-05	1.91E-04	b	8.08E-05	1.35E-04	b	8.08E-05	1.05E-04	b

Los valores del estimador de Waterson ( $\theta_w$ ), que se basa en el número de sitios segregantes únicos (*singletons*), aumentaron a lo largo del periodo evaluado (Tabla 4) para RNs, VMs y PSs. Al igual que sucedió con  $\pi$ ,  $\theta_w$  es mayor para los PSs, en comparación con los valores obtenidos para RNs y VMs. Nuevamente el promedio de  $\theta_w$  para VM2 fue una orden de magnitud menor en comparación con VM1.

Las tendencias observadas concuerdan con el promedio de la  $D$  de Tajima, que resultó positiva para las RN<1960 ( $0.105 \pm 0.505$ ), esto indica mayor proporción de polimorfismos de frecuencias intermedias ( $\pi$ ). Este valor de  $D$  de Tajima está asociado a selección balanceadora o una contracción poblacional reciente. Debido a que las VTs suelen crecer en gran variedad de ambientes y a lo largo de amplio gradiente altitudinal

(Figura 8), la selección balanceadora parece ser la interpretación más adecuada.

El estimado de la D de Tajima resultó negativo, para las RNs 1960-1980, RNs >2000 y los PSs agrupados por año de colecta (Tabla 5). Este patrón indica una mayor proporción de *singletons*, y se asocia a expansiones poblacionales y/o a selección negativa. La expansión poblacional no parece posible, pues las poblaciones de PSs han tendido a contraerse por la urbanización (Serratos et al. 1997), por lo que resultaría conveniente evaluar la hipótesis de selección negativa.

Los dos grupos de MVs tuvieron valores positivos de la D Tajima, dado que el mejoramiento es resultado de un segundo cuello de botella (Yamasaki et al. 2005), el escenario más probable es la contracción poblacional.

Tabla 4. Theta ( $\theta_w$ ) por pb calculada en ventanas de 20, 50 y 100 kb con pasos de 5, 10 y 25 kb, respectivamente; para razas (LR), *mexicana* (Zmx), *parviglumis* (Zpr) y variedades modernas (MV) (Rojas-Barrera et al. 2019).

Tamaño de ventana (kb), tamaño de ventana (kb)	20,5 (Promedio)	20,5 (sd)	Significancia estadística (p<0.00001)	50,10 (Promedio)	50, 10 (sd)	Significancia estadística(p<2e-16)	100, 25 (Promedio)	100, 25 (sd)	Significancia estadística (p<2e-16)
LR <1960	1.66E-04	3.44E-04	a	1.66E-04	2.46E-04	a	1.66E-04	1.92E-04	a
LR 1960-1980	1.79E-04	3.69E-04	b	1.79E-04	2.63E-04	b	1.79E-04	2.06E-04	b
LR >2000	2.04E-04	4.14E-04	c	2.04E-04	2.96E-04	c	2.04E-04	2.32E-04	c
Tamaño de ventana (kb), tamaño de ventana (kb)	20,5 (Promedio)	20,5 (sd)	Significancia estadística (p<2e-16)	50,10 (Promedio)	50, 10 (sd)	Significancia estadística(p<2e-16)	100, 25 (Promedio)	100, 25 (sd)	Significancia estadística (p<2e-16)
Zpr 1980	2.01E-04	4.23E-04	a	2.01E-04	3.00E-04	a	2.01E-04	2.33E-04	a
Zpr >2000	2.55E-04	5.19E-04	b	2.55E-04	3.70E-04	b	2.55E-04	2.88E-04	b
Zmx 1980	2.10E-04	4.45E-04	a	2.10E-04	3.17E-04	a	2.10E-04	2.46E-04	a
Zmx >2000	2.55E-04	5.19E-04	b	2.55E-04	3.70E-04	b	2.55E-04	2.88E-04	b
MV1	1.13E-04	2.43E-04	a	1.13E-04	1.74E-04	a	1.13E-04	1.36E-04	a
MV2	5.54E-05	1.28E-04	b	5.54E-05	9.07E-05	b	5.54E-05	7.07E-05	b

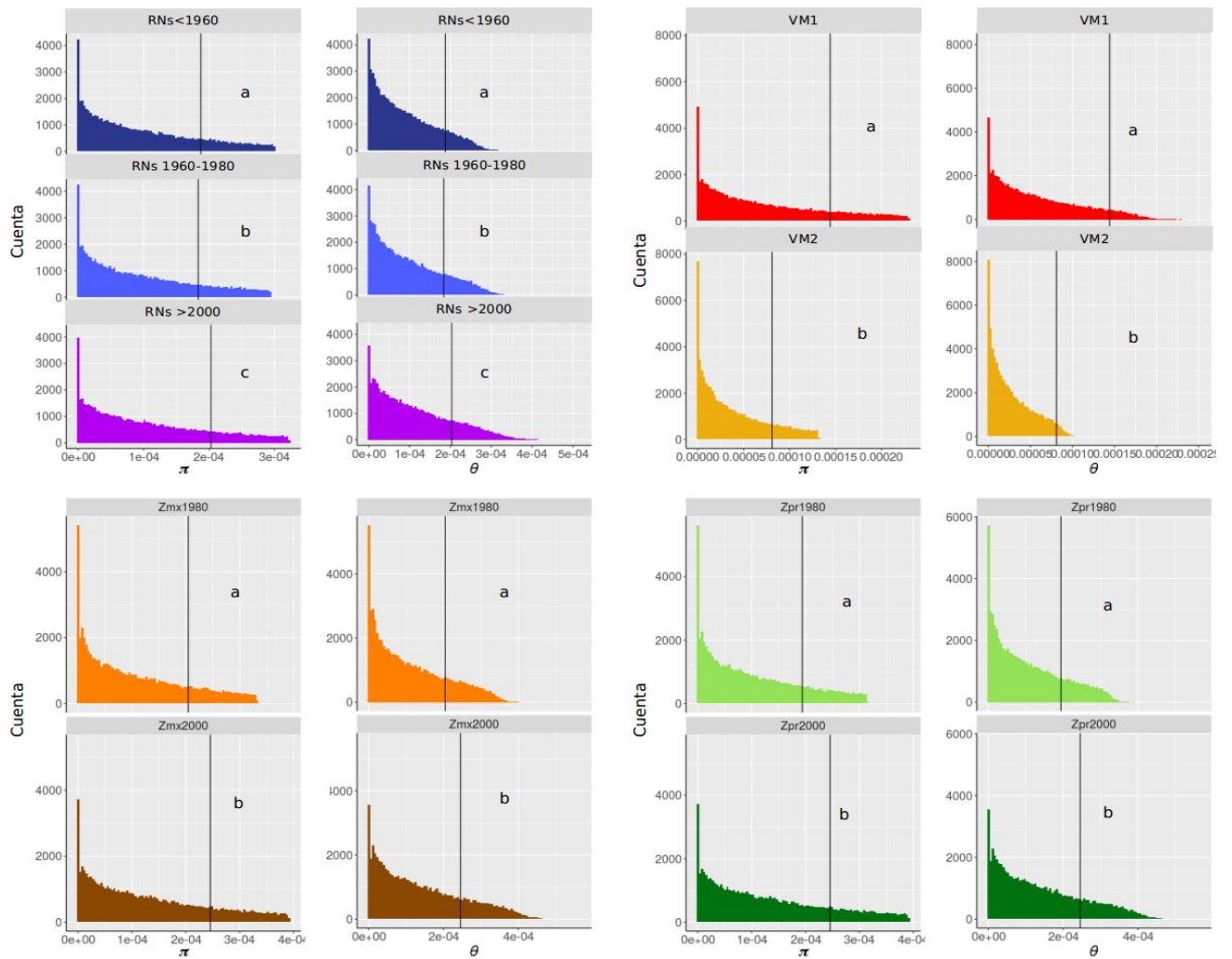


Figura 12. Estimados de diversidad genética  $\pi$  y  $\theta_w$  en razas, variedades modernas y parientes silvestres de maíz agrupados de acuerdo con su periodo de colecta. A) Razas (RNs), B) Variedades modernas, grupos genéticos: VM1 y VM2, C) *Zea mays* ssp. *mexicana*, y D) *Zea mays* ssp. *parviglumis*, las letras minúsculas indican diferencias significativas ( $p < 0.0001$ ) entre los genotipos de un mismo grupo comparados de acuerdo con el periodo de colecta. Las líneas verticales en cada caso muestran el promedio aritmético.

Tabla 5. Valores promedio de la D de Tajima calculada en 20, 50 y 100 kb y con pasos de 5, 10 y 25 kb, respectivamente; para razas (LR), mexicana (Zmx), *parviglumis* (Zpr) y variedades modernas (MVs) de maíz (Rojas-Barrera et al. 2019).

Grupo	Promedio	sd	IQR (Q3-Q1)
LR <1960	0.105 <sup>a</sup>	0.505	0.249
LR 1960-1980	- 0.032 <sup>b</sup>	0.453	0.125
LR >2000	- 0.077 <sup>c</sup>	0.546	0.236
MV1	0.277 <sup>d</sup>	0.576	0.572
MV2	0.468 <sup>e</sup>	0.724	0.895
Zmx 1980	-0.077 <sup>f</sup>	0.299	0.167
Zmx >2000	- 0.093 <sup>g</sup>	0.44	0.26
Zpr 1980	- 0.089 <sup>h</sup>	0.285	0.182
Zpr >2000	- 0.093 <sup>g</sup>	0.44	0.26

IQR = Rango intercuantílico



## 4. Discusión

### 4.1 La introducción de variedades mejoradas ha modificado la estructura poblacional de las razas de maíz

La correlación entre la ancestría y la altitud concuerda con trabajos previos que muestran la relevancia de la altitud en la estructuración poblacional del maíz (van Heerwaarden et al. 2007; van Heerwaarden et al. 2010; Arteaga et al. 2016), y es consistente con la adopción de las VMs en regiones de altitudes medias y bajas (Bjarnason 1994; Perales, Brush & Qualset 2003; Bellon & Brush 1994). A pesar de los esfuerzos para desarrollar VMs adaptadas a mayor altitud (Wellhausen 1961; Vázquez Carrillo et al. 2012; Tadeo Robledo et al. 2012), su presencia todavía es escasa en los valles altos (Donnet et al. 2012; Perales, Brush & Qualset 2003). Durante nuestras colectas no encontramos VMs en simpatria con la raza Chalqueño. Se ha reportado que las razas de valles altos carecen de tolerancia al calor y a la endogamia (Bjarnason 1994), características que han sido asociadas a una fuerte adaptación local y que probablemente han limitado la introducción de VMs en zonas altas. Así, a pesar de que los agricultores de baja escala han adoptado paquetes tecnológicos, la adopción de las VMs ha sido limitada (Perales, Brush & Qualset 2003 a,b).

Las VMs colectadas en simpatria con RNs formaron dos grupos genéticos: VM1 (rojo) y VM2 (amarillo) (Fig. 4B). Al ordenar las accesiones de las más antiguas a las más recientes, el incremento en la proporción de la ancestría asociada a VM1, fue notable y sugiere que se debe al flujo genético de las VMs hacia las RN (Fig. 4B y 5A-C). Este cambio en la estructura poblacional a través del periodo evaluado, parece estar asociado al contexto histórico, pues coincide con la adopción gradual de las VMs, que se inició en la década de 1940 con las primeras VMs generadas a partir de RNs sobresalientes en la producción de grano (Wellhausen 1961). El proceso de introducción de VMs se aceleró en la década de 1990, después de las modificaciones a la Ley Federal de Variedades Vegetales (Espinosa-Calderón et al. 2014); y los cambios en la estructura poblacional coinciden con los datos estadísticos del reporte "Uso de tecnología y servicios en la superficie agrícola nacional", que

indican un incremento sostenido en la adopción de VMs desde 2012 hasta 2016 (SIAP 2016).

Los resultados del PCA (Fig. 10) y los árboles de máxima verosimilitud (Fig. 11) agruparon a las VM1 y VM2, con los genotipos derivados de los acervos genéticos tropicales producidos por los programas de mejoramiento de México, Camerún y Nigeria. Ambos árboles ubicaron a los CMLs como el grupo basal del cual se derivan VM1 y VM2 (Fig. 11A-B). Aunque el análisis de ancestría que sólo incluyó a los genotipos tropicales (Fig. 11C) agrupó a dos accesiones del grupo VM1 (MsH22015 y ZmH42015) con los genotipos derivados del programa de mejoramiento de Camerún. En la década de 1980, el Instituto Internacional de Agricultura Tropical (IITA, por sus siglas en inglés), en conjunto con el Instituto de Investigaciones Agronómicas de Camerún (IRA, por sus siglas en inglés), inició un programa para desarrollar híbridos tropicales para zonas de altitud media (Menkir et al. 2004), algunas de estas líneas fueron incorporadas a la colección del Banco Nacional de Germoplasma de EUA. Esto implica que probablemente han sido empleadas para generar VMs con amplia distribución. Por lo anterior, no podemos descartar la posibilidad de que las VMs colectadas en simpatria con las razas Zamorano y Mushito en el estado de Michoacán, estén basadas en germoplasma exótico.

Al igual que con VM1, las CMLs resultaron basales al grupo genético VM2. No obstante, en todos los análisis de agrupamiento realizados, VM2 siempre forma un grupo distinto y homogéneo. La baja diversidad genética estimada para este grupo (Tablas 3 y 4) puede deberse a un cuello de botella producido durante la generación de líneas endogámicas que son la base de VM2, lo que explicaría su baja diversidad genética (Tabla 3 y 4) y los valores de  $F_{ST}$  observados (Fig. 9). Dos individuos colectados como razas en el 2015 se agruparon con VM2, lo que sugiere que se trata de semillas acriolladas derivadas de alguna VM. El acriollamiento de VMs, que implica la adaptación al manejo y ambientes locales son una práctica común en México (Perales, Brush & Qualset 2003; Bellon & Berthaud 2004).

Por otra parte, la distribución inusual de los teocintles que se agrupan con *mexicana*, pero que fueron colectados en San Lorenzo, Jalisco,



una localidad donde anteriormente ha sido colectado *parviglumis*, puede atribuirse a la movilidad de semillas propiciada por actividades humanas. Anteriormente se ha sugerido que teocintles han llegado a localidades en el estado de Puebla en las camionetas que transportan y distribuyen estiércol (Serratos 1997). Un caso más reciente fue el descubrimiento de un teocintle en la región de Aragón, España (Trtikova et al. 2017).

El cambio en la distribución geográfica del teocintle es un evento poco estudiado, que requiere que se recabe evidencia de manera sistemática para identificar las causas que promueven su movilidad y los nuevos puntos de distribución.

#### 4. 2 Flujo genético e introgresión de variedades modernas hacia razas nativas y parientes silvestres de maíz

El flujo genético de ambos grupos genéticos de VMs hacia las RN>2000 probablemente está asociada a se creciente adopción y a la introducción de germoplasma extranjero después de la década de 1990. La ley de semillas fue modificada en 1991 y 2007, lo que permitió la participación del sector privado en la investigación y comercio de VMs (Luna Mena et al. 2012; Espinosa-Calderón et al. 2014). Aún cuando las primeras VMs fueron liberadas en 1940 y su producción se incrementó en 1960 con el inicio de la revolución verde, la ancestría de las RNs 1960-1980 no presentan cambios en la estructura poblacional (Fig. 5B), es probable que esto se deba a que las primeras VMs se produjeron a partir de razas con alta producción de grano como *Tuxpeño* y *Celaya*, y el flujo genético, si lo hubo, pudo haber transferido alelos que ya estaban presentes en las RNs (Bellon & Brush 1994; Wellhausen 1961). En 1970, el sector privado comenzó a desarrollar sus propias VMs, aunque estas no fueron liberadas hasta 1980. Durante este periodo las compañías privadas no tenían acceso a las semillas producidas por el INIA (Luna Mena et al. 2012); lo que hace probable que las VMs de este periodo estuvieran basadas en el acervo genético de las líneas de maíz desarrolladas por el CIMMYT (CMLs) o en germoplasma foráneo.

Más allá del contexto sociopolítico, prácticas como el acriollamiento de la semillas de VMs han promovido el flujo genético -- deliberado o

accidental -- entre VMs y RNs (Bellon & Berthaud 2004; Bellon & Brush 1994; Perales, Brush & Qualset 2003). El flujo genético de las VMs hacia RNs ha sido documentado previamente con evidencia fenotípica (Hernández Xolocotzi & Ortega Paczka 1987; Perales, Brush & Qualset 2003); sin embargo, hasta donde tenemos conocimiento, este es el primer trabajo que provee evidencia genética del flujo de las VMs hacia las RNs y PSs en el centro de origen del maíz.

La ausencia de flujo genético en la única RN distribuida en valles altos, Chalqueño, puede atribuirse a la baja presencia de VMs en estas regiones (Donnet et al. 2012; Perales, Brush & Qualset 2003)(Figura 8, Anexo 2 ). También, se ha sugerido que para maíz la adaptación local en maíces de valles altos juega un papel crucial, debido a que el número de ambientes por unidad de área es mayor en tierras altas que en regiones de altitud media (Bjarnason 1994). Estos ambientes heterogéneos requieren el desarrollo de cultivares adaptados a condiciones específicas, una estrategia que difiere de las que se usan para generar VMs. Las VMs tropicales y subtropicales están adaptadas a las condiciones ambientalmente homogéneas dominantes en regiones de gran extensión geográfica, lo que facilita su adopción y uso a gran escala. En este contexto, es probable que en caso de haber flujo genético de VMs hacia las RNs locales, los alelos generalistas provenientes de las VMs serán eliminados por selección negativa en favor de los alelos asociados a la adaptación local (CIMMYT 1988; Mercer, Martínez-Vásquez, & Perales 2008).

En contraste con nuestros resultados y las observaciones anteriores, el trabajo realizado por Joost van Heerwaarden 2007 con microsatélites, observó que la introgresión ocurre en mayor proporción de VTs de tierras bajas hacia VTs de tierras altas. El autor concluyó que este resultado era esperado debido a la relativa facilidad con la que crecen las VTs de tierras bajas en tierras altas, en comparación con la condición inversa, VTs de tierras altas creciendo en tierras bajas. La discordancia con nuestras observaciones, podría deberse a que las VTs en comparación con las VMs, están adaptadas a otras condiciones como el manejo tradicional y baja proporción de insumos externos (Bellon et al. 2018), aunque para probar esta propuesta resultaría conveniente expandir el muestreo de maíces de

regiones de mayor altitud, pues nuestro trabajo sólo incluyó a la raza Chalqueño.

Sin embargo, la intensidad de la adaptación local representa una disyuntiva que puede limitar la capacidad de respuesta de las RNs de valles altos al cambio climático y hace más vulnerable a los agricultores de estas regiones (Bellon, Hodson, & Hellin 2011). Adicionalmente, la escasez de VMs en regiones altas también ha sido atribuida al fracaso de los programas institucionales, que no han logrado desarrollar estrategias que promuevan la adaptación de las VMs en tierras altas. Un estudio del realizado por Perales et al. en 2003 proyectó que en el mediano plazo (10-20 años), las razas de valles altos podrían ser desplazadas por las VMs. Sin embargo, 16 años después de la publicación de este trabajo las VMs aún son raras en regiones de valles altos, y la estructura genética de la raza Chalqueño no ha cambiado en el periodo evaluado (Fig. 4). Aunque, de acuerdo con lo observado en este trabajo, es de esperarse que cuando las VMs incrementen su frecuencia en valles altos, la estructura genética seguirá la misma tendencia que la observada en las RNs de altitudes medias y bajas.

El objetivo de incluir RNs simpátricas con VMs fue incrementar la probabilidad de detectar flujo genético. Sin embargo, nuestros resultados indican que la simpatría al momento del muestreo no es necesaria para la detección de flujo genético. Por ejemplo, detectamos flujo genético de VMs hacía las razas Vandeño y Pepitilla en ausencia de VMs en regiones vecinas al momento del muestreo, lo que implica que el flujo genético ocurrió en generaciones previas o es consecuencias de la simpatría con VMs acriolladas, como los individuos colectados como Vandeño que se agruparon con VM2 (Fig. 4B).

*Zea mays* tiene tres *loci* asociados a un sistema de incompatibilidad polen-pistilo que es heredable y no recíproco (ver sección 1.6 de la Introducción). Los pistilos de RNs y PSs que tienen el alelo Ga1-a pueden ser fertilizadas por polen con genotipo Ga-1s o Ga1-m, pero no son compatibles con el alelo recesivo ga1; y un patrón similar se ha reportado para el alelo Tcb-s (Kermicle 2006; Padilla García et al. 2012). Basados en esto, consideramos que la incompatibilidad gametofítica podría ser la responsable de la ausencia de flujo de las VMs hacia Chalqueño. No

obstante, el alelo dominante Ga1-m se ha detectado en Chalqueño y en una muestra representativa de VMs (de La Cruz Larios 2007; Padilla García et al. 2012). Por tanto, un impedimento en la fertilización cruzada no parece ser la responsable de la ausencia de flujo genético de las VMs hacia Chalqueño.

Este estudio provee evidencia genómica del flujo genético de VMs hacia *mexicana*. Nuestros resultados coinciden con evidencia experimental previa que consiste en cruzas de teocintles de la Planicie Central Mexicana y de la región de Chalco como progenitor femenino y la VM P36D1 como progenitor masculino; estas cruzas producen entre 5-10 % de semillas viables (Baltazar et al. 2005). Aún cuando la tasa de flujo es baja, esta cantidad es suficiente para producir la señal de flujo detectada. Adicionalmente, se ha inferido genéticamente que en poblaciones naturales que el flujo genético entre *mexicana* y maíz es bidireccional (Hufford et al. 2013), aunque ocurre con una tasa asimétrica que favorece el flujo de PSs hacía maíz (Baltazar et al. 2005).

La dirección del flujo de VMs hacia Zpr1980 indica que el flujo genético fue mayor en el pasado, aunque esto no descarta el flujo hacia las poblaciones más recientes Zpr2000. A diferencia de la población alopátrica de *mexicana*, las poblaciones alopátricas de *parviglumis*: Amatlán y Malinalco, comparten ancestría con *mexicana* (Fig. 2A Zpr-Allo), por lo que decidimos descartar su uso como control para evaluar el efecto de la simpatría con VMs.

Recientemente se ha sugerido que los PSs introgresados con alelos provenientes de cultivares representan bancos de germoplasma *in situ* (Ellstrand 2018), debido a que estas poblaciones silvestres deben tener una mayor facilidad para cruzarse con sus parientes domesticados, y los alelos introgresados tienen la cualidad de permanecer en las poblaciones bajo condiciones agrestes (Ellstrand 2018). Esta propuesta reevalúa a los parientes simpátricos introgresados con alelos domesticados y provee una hipótesis que puede ponerse a prueba en el futuro. Sin embargo, desde una perspectiva contraria, el flujo genético con VMs puede llevar a la pérdida de alelos silvestres asociados a adaptación local (Todesco et al. 2016).

### 4.3 La divergencia entre razas y parientes silvestres con VMs disminuye en el tiempo

El índice de divergencia ( $F_{ST}$ ) se ha usado ampliamente para evaluar la estructura poblacional (Meirmans & Hedrick 2011), la disminución de la divergencia entre RNs-VMs y PS-VMs concuerda con el cambio en la estructura poblacional en el periodo evaluado (Fig. 4 y 5).

Es interesante de apuntar que la  $F_{ST}$  entre *parviglumis* y VMs también disminuye a través del tiempo, a pesar de la ausencia de flujo para las poblaciones del 2000 (Fig. 8 C-F). Estos datos indican que el intercambio genético no es un proceso asociado a la disminución del valor de  $F_{ST}$  a lo largo del genoma, por lo tanto, es necesario identificar las regiones sujetas a introgresión en los genomas de las RNs y los PSs, y compararlas con aquellas regiones que no han sido afectadas por el flujo.

En un estudio realizado con *Heliconius* donde  $F_{ST}$  fue calculada en ventanas de 5kb a lo largo del genoma, el tamaño reducido de las ventanas se asoció con un incremento en la diferenciación en *H. melpomene* y *H. pachinus*, y con valores altos del estadístico  $D$  de Patterson (Kronforst et al. 2013). Lo anterior ilustra la manera en que  $F_{ST}$  y  $D$ , que evalúan la divergencia genética y el intercambio genético (asociado a la homogeneización de las poblaciones), respectivamente, pueden producir patrones discordantes cuando se estiman para regiones pequeñas (Martin, Davey, & Jiggins 2015), como los resultados obtenidos en este trabajo para  $F_{ST}$  y el estadístico  $D$ . Aunque, otro factor a considerar es el hecho de que el estadístico  $F_{ST}$  considera los SNPs distribuidos a lo largo de todo el genoma, mientras que el valor de la  $D$  de Patterson se basa sólo en unos cuantos *loci*.

### 4.4 Las CMLs son basales a los grupos VM1 y VM2

A pesar de que las VM1 y VM2 forman grupos definidos en el PCA que incluye las RNs (Fig. 10A), ambos grupos de VMs caen dentro de las distribuciones de los acervos tropicales generados por los programas de mejoramiento de México, Nigeria y Camerún (Fig. 10 B-D). El acervo genético de México incluido en el trabajo de Romay et al. 2013, se compone principalmente de líneas endogámicas de maíz desarrolladas por el CIMMYT (CMLs). Más del 80% de las VMs producidas por el sector

privado contienen germoplasma del CIMMYT (Aquino Mercado, Peña Bautista, & Ortiz-Monasterio 2008; Morris 2000). Lo anterior coincide con la observación de que las CMLs son los grupos basales en los árboles de máxima verosimilitud (Fig. 11A-B).

Sin embargo, las VMs producidas por compañías de semillas transnacionales son comercializadas en diferentes países, y cuando éstas usan material proveniente del CIMMYT suelen incorporar poblaciones y complejos con una base genética más amplia, que puede incluir germoplasma de diferentes regiones del mundo (Morris 2000), lo anterior podría explicar porque el grupo VM1 comparte ancestría con VMs de Camerún. No obstante, otra interpretación para este resultado es que las CMLs desarrolladas por el CIMMYT forman parte de la base genética de las VMs generadas por el programa de mejoramiento de Camerún, y es por lo que MV1 se agrupa con esos linajes.

Aunque las CMLs también son basales al grupo VM2, los análisis de ancestría realizados con accesiones de los programas tropicales no agrupan a MV2 con ningún genotipo, pero la ancestría de MV2 está presente en baja proporción en el acervo del programa de mejoramiento de México. Es probable que MV2 provenga de un acervo genético de introducción reciente, derivado de variedades foráneas o de un conjunto de RNs no incluidas en este trabajo. Otra posibilidad es que VM2 haya sido generado a partir de una segunda ronda de mejoramiento, lo que implicaría un cuello de botella que redujo su diversidad nucleotídica y produjo divergencia a nivel de todo el genoma por deriva genética (Excoffier & Ray 2008), esta explicación es la más plausible pues concuerda con los patrones observados con los análisis de agrupamiento (Fig.9) y los valores de la D de Tajima (Tabla 5).

#### 4.5 La diversidad nucleotídica incrementa en las razas y parientes silvestres muestreados después del 2000

El decremento del valor de  $\pi$  en el periodo 1960-1980 probablemente está asociado a la adopción y acriollamiento de VMs que se liberaron en la década de 1960 (Wellhausen 1961; Trueba Carranza 2012). Un estudio realizado en el sureste de Chiapas demostró que las VMs acriolladas eran plantadas en mayor proporción que las RNs locales, y que las semillas de

las VMs eran recicladas hasta 10 años después de su introducción (Bellon & Risopoulou 2001; Bellon & Brush 1994). Por otro lado, aun cuando la participación del sector privado en el comercio inició en la década de 1980, la adopción de las VMs no fue instantánea. En la década de los 90s, 45 compañías controlaban el 55% del mercado de semillas mejoradas, esta tendencia se consolidó en la siguiente década, y en 2009 dos compañías producían el 95% de las semillas mejoradas en México (Aboites, Martínez, & Torres 1999; Luna Mena et al. 2012). Por tanto, la adopción gradual de las VMs (SIAP 2016) y la reconfiguración de la producción de semillas mejoradas después de la década de 1990 puede explicar el incremento significativo de  $\pi$  en las RNs > 2000.

Anteriormente, se había reportado para maíz y arroz que el intercambio genético con VMs es seguido de un incremento de  $\pi$  (Bitocchi et al. 2015; Bellucci et al. 2013; Steele et al. 2009), particularmente si las VMs son producidas a partir de germoplasma foráneo. Sin embargo, a este incremento en la diversidad le sigue una caída en la diversidad, resultado del desplazamiento parcial o total por parte de las VMs (van de Wouw et al. 2010; Brush 1995). Respecto a los grupos VM1 y VM2, observamos una baja diversidad nucleotídica (Tabla 2) y una alta divergencia probablemente producidas por deriva genética (Excoffier & Ray 2008).

Aunque nuestra aproximación no fue diseñada para evaluar erosión genética, los resultados obtenidos son relevantes para la controversia suscitada sobre la posible erosión genética de maíz en su centro de origen (Dyer et al. 2014; Brush et al. 2015; Dyer et al. 2015). La erosión genética puede ser evaluada a nivel de cultivos, variedades o alelos (van de Wouw et al. 2010). En este trabajo logramos recolectar las siete RNs registradas en el base de datos de maíces nativos, que fueron recolectadas antes de 1960 y  $\pi$  resultó mayor para las RNs más recientes. Por tanto, no detectamos pérdida de variación genética, lo que concuerda con reportes previos (Bellon & Brush 1994; Louette, Charrier, & Berthaud 1997; Perales, Brush, & Qualset 2003), y tampoco hallamos evidencia de erosión a nivel de diversidad nucleotídica.

No obstante, resta explorar la pérdida o cambio en la frecuencia de alelos asociados a la adaptación a ambientes locales y a usos especiales como el valor estético, ritual y otras preferencias culturales (Perales,



Brush,&Qualset 2003; Bellon&Brush 1994; García-Salazar&Ramírez-Jaspeado 2014; Zarazúa-Escobar, Almaguer-Vargas,&Ocampo-Ledesma 2011).

Nuestros hallazgos apoyan el desplazamiento gradual de los sistemas bajo manejo tradicional por paquetes tecnológicos empleados en mejoramiento que incluyen a las VMs, los fertilizantes, la maquinaria y los pesticidas (Byerlee & de Polanco 1986). Esta transición de la organización comunal hacia un entorno que considera a los cultivos sólo como valor de cambio, coloca a los agricultores pequeños en una situación desventajosa a nivel social, económico y ecológico (Zarazua-Escobar, Almaguer-Vargas, & Ocampo-Ledesma 2011), y también puede ser precedente de un cuello de botella de modernización (van de Wouw et al. 2010), que debe ser monitoreado dados los servicios evolutivos y nutrimentales que presta los agricultores pequeños (Bellon et al. 2018; Faith et al. 2010).

Nuestros criterios de inclusión y clasificación de VT en RNs están basado en la definición de Anderson y Cutler (1942), quienes describieron a una raza como un conjunto de individuos relacionados con el suficiente número de características en común para ser reconocidas como grupo. Aun así, los resultados obtenidos apuntan a que la definición propuesta por Casañas et al.(2017) puede adaptarse mejor al caso del maíz; esta define a las RNs como variedades de plantas que han evolucionado y continúan evolucionando a través de técnicas de mejoramiento convencional o moderno, en ambientes tradicionales o tecnificados, dentro de un área ecogeográfica influenciada por la cultura local.

## 5. Conclusiones

En gran parte del territorio del centro y sur de México, las VTs nativas y sus PSs son simpátricas con VMs. Por tanto, conservar a las VTs y a sus PS en un estado estático no es posible ni deseable (Perales R, Brush, & Qualset 2003 a,b). Sin embargo, es posible preservar los mecanismos que permiten la evolución de los acervos genéticos domesticados y silvestres de maíz.

Nuestros resultados sugieren que la adopción de MVs ha tenido impactos significativos en la composición genética de las VTs y sus PSs. En este estudio demostramos la existencia de flujo genético de las MVs hacia VTs y mexicana, y que las frecuencias alélicas han cambiado en el periodo evaluado. Nuestras observaciones proveen evidencia de corto plazo del proceso evolutivo continuo de las VTs y de sus PSs en su centro de origen y a la vez evidencian la transición de un sistema agrícola tradicional hacia un sistema orientado a la agricultura comercial, lo que sugiere un cuello de botella de modernización (van de Wouw et al. 2010; Brush 1995).

Es importante cubrir las necesidades del pequeño agricultor y valorar los servicios ecológicos y evolutivos que este presta (Bellon et al. 2018). Enfrentar al pequeño agricultor al mercado de la semilla híbrida lo pone en desventaja y devalúa su labor. No se trata de limitar la introducción de VMs, sino de impulsar estrategias como el mejoramiento evolutivo participativo de razas locales (Murphy et al. 2005), fortalecer y fomentar las redes de bancos de semillas locales, y la generación de mercados que reconozcan el valor agregado de las razas: atributos estéticos, mayor calidad nutrimental, y valor culinario (Bellon et al. 2018; Brush 1995); de tal modo que se prevenga o amortigüe la erosión genética en vistas del cuello genético asociado a la modernización por el que probablemente atravesará el campo mexicano.

Finalmente, nuestros hallazgos pueden ser empleados como referencia para diseñar estrategias orientadas al monitoreo de la diversidad genética y políticas agrícolas encaminadas a cumplir las metas de biodiversidad (Aichi Biodiversity Targets 2010.) para conservar los reservorios genéticos en los centros de diversidad.

## 6. Perspectivas

El estudio del papel del flujo genético en la evolución de especies y poblaciones vegetales ha cobrado relevancia en las últimas décadas debido a la frecuencia con la que el intercambio e introgresión genéticos suelen suceder (Ellstrand 2014). Perspectivas recientes sugieren que la siguiente etapa en el estudio del flujo genético, es la estimación del número de eventos requeridos para observar los patrones de introgresión detectados al momento del muestreo (Rieseberg 2019). A partir de los resultados y la discusión presentados en este trabajo proponemos las siguientes perspectivas con el objetivo de evaluar la robustez de nuestras observaciones y modelar escenarios de flujo genético que nos permitan predecir cambios asociados al flujo genético de VMs hacia RNs.

1. Identificar las regiones introgresadas en el genoma de VTs empleando el estadístico  $f_d$  propuesto por (Martin, Davey, & Jiggins 2015) y evaluar si el índice de divergencia  $F_{ST}$  difiere en estas regiones respecto al resto del genoma
2. Determinar si los alelos asociados al mejoramiento de VMs de maíz (Hufford et al. 2012), están representados en las zonas introgresadas; y si se encuentran selección positiva y/o negativa.
3. Modelar escenarios de flujo genético, con herramientas como *fastsimcoal* (Excoffier & Foll 2011), para estimar la tasa de migración hacia RNs y PSs.
4. Evaluar parámetros indicativos de erosión genética en VTs y PSs, como corridas de homocigosis en el genoma y cambios en el tamaño poblacional ( $N_e$ ).
5. Ampliar el muestreo de razas de zonas altas, con el fin de evaluar si las VTs de valles altos (> 1800) están más estructuradas en comparación con las VTs de altitudes medias y bajas. En caso de confirmarse el patrón anterior, evaluaremos si la estructura está dada por la adaptación local asociada con la heterogeneidad de ambientes en valles altos.
6. La historia demográfica de las poblaciones de teocintles de Amatlán

de Quetzalcóatl, Morelos y Malinalco, Edo de México, permanece como una incógnita. Su distribución a la mitad del gradiente altitudinal de *mexicana* y *parviglumis* (Sánchez González et al. 2018), y el grado de hibridación o ancestría mestiza (Figura 4), sugiere que estas poblaciones de PSs podrían ser ancestrales a las poblaciones de *mexicana* (Fukunaga et al. 2005). Esta hipótesis puede ser explorada ampliando nuestro muestreo o incluyendo poblaciones de teocintles que se distribuyen en regiones altitud media; para evaluar las relaciones filogenéticas y escenarios de flujo genético entre las poblaciones de Malinalco y Amatlán, con poblaciones de *mexicana* distribuidas en la Meseta Central de México y en las regiones altas de Edo. de México.

## 7. Referencias

- Aboites, G., Martínez, F., & Torres, G. (1999). El negocio de la producción de semillas mejoradas y su rol en el proceso de privatización de la agricultura mexicana. *Espiral Estudios Sobre Estado Y Sociedad*, 6(16), 151-185.
- Acevedo, F., Huerta, E., Burgeff, C., Koleff, P., & Sarukhán, J. (2011). Is transgenic maize what Mexico really needs? *Nature Biotechnology*, 29(1), 23-24.
- Agapito-Tenfen, S., Lopez, F. R., Mallah, N., Abou-Slemayne, G., Trtikova, M., Nodari, R. O., & Wickson, F. (2017a). Transgene flow in Mexican maize revisited: Socio-biological analysis across two contrasting farmer communities and seed management systems. *Ecology and Evolution*, 7(22), 9461-9472.
- Agostini, M. G., Roesler, I., Bonetto, C., Ronco, A. E., & David, B. (2020). Pesticides in the real world: The consequences of GMO-based intensive agriculture on native amphibians. *Biological Conservation*, 241, 108355.
- Ahmed, A. U., Hoddinott, J. F., Abedin, N., & Hossain, N. Z. (2019). *Economic and Health Impacts of Genetically Modified Eggplant*. International Food Policy Research Institute. <http://www.indiaenvironmentportal.org.in/files/file/Economic%20and%20health%20impacts%20of%20genetically%20modified%20eggplant.pdf>
- Aichi Biodiversity Targets 2010. Retrieved September 26, 2018, from <https://www.cbd.int/sp/targets/>
- Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, 19(9), 1655-1664.
- Anderson, J. A., Ellsworth, P. C., Faria, J. C., Head, G. P., Owen, M. D. K., Pilcher, C. D., Shelton, A. M., & Meissle, M. (2019). Genetically Engineered Crops: Importance of Diversified Integrated Pest Management for Agricultural Sustainability. *Frontiers in Bioengineering and Biotechnology*, 7, 24.
- Aquino Mercado, P. R., Peña Bautista, R. J., & Ortiz-Monasterio, I. (2008). *México y el CIMMYT*. CIMMYT.
- Arteaga, M. C., Moreno-Letelier, A., Mastretta-Yanes, A., Vázquez-Lobo, A., Breña-Ochoa, A., Moreno-Estrada, A., Eguiarte, L. E., & Piñero, D. (2016). Genomic variation in recently collected maize landraces from Mexico. *Genomics Data*, 7, 38-45.
- Baltazar, B. M., Castro Espinoza, L., Espinoza Banda, A., de la Fuente Martínez, J. M., Garzón Tiznado, J. A., González García, J., Gutiérrez, M. A., Guzmán Rodríguez, J. L., Heredia Díaz, O., Horak, M. J., Madueño Martínez, J. I., Schapaugh, A. W., Stojšin, D., Uribe Montes, H. R., & Zavala García, F. (2015). Pollen-Mediated Gene Flow in Maize: Implications for Isolation Requirements and Coexistence in Mexico, the Center of Origin of Maize. *PLoS One*, 10(7), e0131549.
- Baltazar, B. M., Sánchez-González, J. de J., de la Cruz Larios, L., & Schoper, J. B. (2005). Pollination between maize and teosinte: an important determinant of gene flow in Mexico. *TAG. Theoretical and Applied Genetics. Theoretische Und Angewandte Genetik*, 110(3), 519-526.
- Beckie, H. J., & Hall, L. M. (2014). Genetically-modified herbicide-resistant (GMHR) crops a two-edged sword? An Americas perspective on development and effect on weed management. *Crop Protection*, 66, 40-45.
- Bellon, M. R., & Berthaud, J. (2004). Transgenic maize and the evolution of landrace diversity in Mexico. The importance of farmers' behavior. *Plant Physiology*, 134(3), 883-888.
- Bellon, M. R., & Brush, S. B. (1994). Keepers of maize in Chiapas, Mexico. *Economic Botany*, 48(2), 196-209.
- Bellon, M. R., Hodson, D., & Hellin, J. (2011). Assessing the vulnerability of traditional maize seed systems in Mexico to climate change. *Proceedings of the National Academy of Sciences of the United States of America*, 108(33), 13432-13437.

- Bellon, M. R., Mastretta-Yanes, A., Ponce-Mendoza, A., Ortiz-Santamaría, D., Oliveros-Galindo, O., Perales, H., Acevedo, F., & Sarukhán, J. (2018). Evolutionary and food supply implications of ongoing maize domestication by Mexican campesinos. *Proceedings. Biological Sciences / The Royal Society*, 285(1885). <https://doi.org/10.1098/rspb.2018.1049>
- Bellon, M. R., & Risopoulos, J. (2001). Small-scale farmers expand the benefits of improved maize germplasm: A case study from Chiapas, Mexico. *World Development*, 29(5), 799–811.
- Bellon, M. R., & Taylor, J. E. (1993). “Folk” Soil Taxonomy and the Partial Adoption of New Seed Varieties. *Economic Development and Cultural Change*, 41(4), 763–786.
- Bellucci, E., Bitocchi, E., Rau, D., Nanni, L., Ferradini, N., Giardini, A., Rodriguez, M., Attene, G., & Papa, R. (2013). Population structure of barley landrace populations and gene-flow with modern varieties. *PLoS One*, 8(12), e83891.
- Benbrook, C. M. (2012). Impacts of genetically engineered crops on pesticide use in the U.S. -- the first sixteen years. In *Environmental Sciences Europe* (Vol. 24, Issue 1). <https://doi.org/10.1186/2190-4715-24-24>
- Bitocchi, E., Bellucci, E., Rau, D., Albertini, E., Rodriguez, M., Veronesi, F., Attene, G., & Nanni, L. (2015). European flint landraces grown in situ reveal adaptive introgression from modern maize. *PLoS One*, 10(4), e0121381.
- Bitocchi, E., Nanni, L., Rossi, M., Rau, D., Bellucci, E., Giardini, A., Buonamici, A., Vendramin, G. G., & Papa, R. (2009). Introgression from modern hybrid varieties into landrace populations of maize (*Zea mays* ssp. *mays* L.) in central Italy. *Molecular Ecology*, 18(4), 603–621.
- Bjarnason, M. (Ed.). (1994). *The Subtropical, Midaltitude, and Highland Maize Subprogram*. CIMMYT.
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114–2120.
- Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., & Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics*, 23(19), 2633–2635.
- Brown, C. H., Clement, C. R., Epps, P., Luedeling, E., & Wichmann, S. (: January 19, 2013). The Paleobiolinguistics of Domesticated Chili Pepper (*Capsicum* spp.). *Ethnobiology Letters*, 4, 1–11.
- Brugière, N., Rothstein, S. J., & Cui, Y. (2000). Molecular mechanisms of self-recognition in Brassica self-incompatibility. *Trends in Plant Science*, 5(10), 432–438.
- Brush, S. B. (1992). Reconsidering the Green Revolution: Diversity and Stability in Cradle Areas of Crop Domestication. *Human Ecology*, 2, 145–167.
- Brush, S. B. (1995). In Situ Conservation of Landraces in Centers of Crop Diversity. *Crop Science*, 35, 346–354.
- Brush, S. B., Taylor, J. E., & Bellon, M. R. (1992). Technology adoption and biological diversity in Andean potato agriculture. *Journal of Development Economics*, 39(2), 365–387.
- Brush, S., Kesseli, R., Ortega, R., Cisneros, P., Zimmerer, K., & Quiros, C. (1995). Potato Diversity in the Andean Center of Crop Domestication. *Conservation Biology*, 9(5), 1189–1198.
- Brush, S. B., Bellon, M. R., Hijmans, R. J., Ramirez, Q. O., Perales, H. R., & van Etten, J. (2015). Assessing maize genetic erosion. *Proceedings of the National Academy of Sciences of the United States of America*, 112(1), E1–E1.
- Bulant, C., Gallais, A., Matthys-Rochon, E., & Prioul, J. L. (2000). Xenia Effects in Maize with Normal Endosperm. *Crop Science*, 40, 182–189.
- Byerlee, D., & de Polanco, E. H. (1986). Farmers’ Stepwise Adoption of Technological Packages: Evidence from the Mexican Altiplano. *American Journal of Agricultural Economics*, 68(3), 519–527.
- Casañas, F., Simó, J., Casals, J., & Prohens, J. (2017). Toward an evolved concept of landrace. *Frontiers in Plant Science*, 8, 145.
- CIMMYT. (1976). *CIMMYT review 1976*. CIMMYT.



- <http://repository.cimmyt.org/xmlui/handle/10883/3599>  
CIMMYT. (1988). *Maize production regions in developing countries*. CIMMYT.  
<http://repository.cimmyt.org/xmlui/handle/10883/705>
- CIP. International Potato Center. International Potato Center. Retrieved July 29, 2019, from <https://cipotato.org/>
- Colin McEwan, J (2016). *SNP filtering through HWE?* ResearchGate. Retrieved January 21, 2020, from [https://www.researchgate.net/post/SNP\\_filtering\\_through\\_HWE](https://www.researchgate.net/post/SNP_filtering_through_HWE)
- CONABIO. (2011a). *Proyecto global de maíces nativos*. Comisión Nacional Para El Conocimiento Y Uso de La Biodiversidad (CONABIO).
- CONABIO. (2011b). *Recopilación, generación, actualización y análisis de información acerca de la diversidad genética de maíces y sus parientes silvestres en México*. Comisión Nacional para el Conocimiento y Uso de la Biodiversidad.  
[http://www.biodiversidad.gob.mx/genes/pdf/proyecto/InformedeGestion\\_V1.pdf](http://www.biodiversidad.gob.mx/genes/pdf/proyecto/InformedeGestion_V1.pdf)
- CONABIO. *Centros de plantas cultivadas*. Biodiversidad Mexicana. Retrieved October 8, 2019, from <https://www.biodiversidad.gob.mx/genes/centrosOrigen/centrosPlantas1a.htm>
- Crow, J. F., & Dove, W. F. (1998). 90 years ago: the beginning of hybrid maize. *Genetics*, 148(3), 923–928.
- De La Cruz Larios, L. (2007). *Sistemas de incompatibilidad genética en maíz y teocintle (Zea spp.) en México* (J. de J. Sánchez González (Ed.)) [Ph.D., Universidad de Guadalajara].  
[http://repositorio.cucba.udg.mx:8080/xmlui/bitstream/handle/123456789/5149/Gonzalez\\_Ledesma\\_J\\_Santos.pdf?sequence=1](http://repositorio.cucba.udg.mx:8080/xmlui/bitstream/handle/123456789/5149/Gonzalez_Ledesma_J_Santos.pdf?sequence=1)
- Delgado-Salinas, A., Caballero, J., & Casas, A. (2004). Crop domestication in Mesoamerica. In R. M. Goodman (Ed.), *Encyclopedia of Plant and Crop Science* (1st ed.). CRC Press Taylor and Francis Group.
- de Nettancourt, D. (2001). *Incompatibility and Incongruity in Wild and Cultivated Plants*. Springer, Berlin, Heidelberg.
- Demont, M., & Stein, A. J. (2013). Global value of GM rice: a review of expected agronomic and consumer benefits. *New Biotechnology*, 30(5), 426–436.
- Domínguez-Valenzuela, J. A., Gherekhloo, J., Fernández-Moreno, P. T., Cruz-Hipolito, H. E., Alcántara-de la Cruz, R., Sánchez-González, E., & De Prado, R. (2017). First confirmation and characterization of target and non-target site resistance to glyphosate in Palmer amaranth (*Amaranthus palmeri*) from Mexico. *Plant Physiology and Biochemistry: PPB / Societe Francaise de Physiologie Vegetale*, 115, 212–218.
- Donnet, M. L., López, D., Arista, J., Carrión, F., Hernández, V., & González, A. (2012). *El potencial de mercado de semillas mejoradas de maíz en México*. CIMMYT.
- Duncan, B., Leyva-Guerrero, E., Werk, T., Stojšin, D., Baltazar, B. M., García-Lara, S., Zavala-López, M., de la Fuente-Martínez, J. M., & Meng, C. (2019). Assessment of potential impacts associated with gene flow from transgenic hybrids to Mexican maize landraces. *Transgenic Research*, 28(5–6), 509–523.
- Durand, E. Y., Patterson, N., Reich, D., & Slatkin, M. (2011). Testing for ancient admixture between closely related populations. *Molecular Biology and Evolution*, 28(8), 2239–2252.
- Dyer, G. A., López-Feldman, A., Yúnez-Naude, A., & Edward Taylor, J. (2014). Genetic erosion in maize's center of origin. *Proceedings of the National Academy of Sciences of the United States of America*, 111(39), 14094–14099.
- Dyer, G. A., López-Feldman, A., Yúnez-Naude, A., Edward Taylor, J., & Ross-Ibarra, J. (2015). Reply to Brush et al.: Wake-up call for crop conservation science. *Proceedings of the National Academy of Sciences of the United States of America*, 112(1), E2–E2.
- Ellstrand, N. C. (1992). Gene flow among seed plant populations. *New Forests*, 6(1–4), 241–256.
- Ellstrand, N. C. (2014). Is gene flow the most important evolutionary force in



- plants? *American Journal of Botany*, 101(5), 737–753.
- Elstrand, N. C. (2018). Does introgression of crop alleles into wild and weedy living populations create cryptic in situ germplasm banks? [Review of *Does introgression of crop alleles into wild and weedy living populations create cryptic in situ germplasm banks?*]. *Molecular Ecology*, 27(1), 38–40.
- Elstrand, N. C., & Elam, D. R. (1993). Population Genetic Consequences of Small Population Size: Implications for Plant Conservation. In *Annual Review of Ecology and Systematics* (Vol. 24, Issue 1, pp. 217–242). <https://doi.org/10.1146/annurev.es.24.110193.001245>
- Elstrand, N. C., Meirmans, P., Rong, J., Bartsch, D., Ghosh, A., de Jong, T. J., Haccou, P., Lu, B.-R., Snow, A. A., Neal Stewart, C., Jr., Strasburg, J. L., van Tienderen, P. H., Vrieling, K., & Hooftman, D. (2013). Introgression of Crop Alleles into Wild or Weedy Populations. *Annual Review of Ecology, Evolution, and Systematics*, 44(1), 325–345.
- Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One*, 6(5), e19379.
- Escudero, M., Eaton, D. A. R., Hahn, M., & Hipp, A. L. (2014). Genotyping-by-sequencing as a tool to infer phylogeny and ancestral hybridization: a case study in *Carex* (Cyperaceae). *Molecular Phylogenetics and Evolution*, 79, 359–367.
- Espinosa-Calderón, A., Turrent-Fernández, A., Tadeo-Robledo, M., Vicente-Tello, S., Gómez-Montiel, N., Valdivia-Bernal, R., Sierra Macías, M., & Zamudio-González, B. (2014). Ley de Semillas y Ley Federal de Variedades Vegetales y transgénicos de maíz en México. *Rev Mex De Cienc Agric*, 5(2), 293–308.
- Evans, M. M. S., & Kermicle, J. L. (2001). Teosinte crossing barrier1, a locus governing hybridization of teosinte with maize. *Theoretical and Applied Genetics*, 103(2-3), 259–265.
- Evenson, R. E., & Gollin, D. (2003). Assessing the impact of the green revolution, 1960 to 2000. *Science*, 300(5620), 758–762.
- Excoffier, L., Dupanloup, I., Huerta-Sánchez, E., Sousa, V. C., & Foll, M. (2013). Robust demographic inference from genomic and SNP data. *PLoS Genetics*, 9(10), e1003905.
- Excoffier, L., & Foll, M. (2011). fastsimcoal: a continuous-time coalescent simulator of genomic diversity under arbitrarily complex evolutionary scenarios. *Bioinformatics*, 27(9), 1332–1334.
- Excoffier, L., & Ray, N. (2008). Surfing during population expansions promotes genetic revolutions and structuration. *Trends in Ecology & Evolution*, 23(7), 347–351.
- Faith, D. P., Magallón, S., Hendry, A. P., Conti, E., Yahara, T., & Donoghue, M. J. (2010). Ecosystem services: an evolutionary perspective on the links between biodiversity and human well-being. *Current Opinion in Environmental Sustainability*, 2(1-2), 66–74.
- Farraia, M., Cavaleiro Rufo, J., Castro Mendes, F., Delgado, L., Laerte Boechat, J., & Moreira, A. (2019). Metabolic interactions in asthma. *European Annals of Allergy and Clinical Immunology*. <https://doi.org/10.23822/EurAnnACI.1764-1489.101>
- FIDA. (2014). *Agricultura familiar en América Latina: Un nuevo análisis comparativo*. Fondo Internacional de Desarrollo Agrícola (FIDA). [https://www.ifad.org/documents/38714170/39135645/Family+farming+in+Latin+America+-+A+new+comparative+analysis\\_s.pdf/9330a6c4-c897-4e1c-9c05-1144ebec0457](https://www.ifad.org/documents/38714170/39135645/Family+farming+in+Latin+America+-+A+new+comparative+analysis_s.pdf/9330a6c4-c897-4e1c-9c05-1144ebec0457)
- Franks, S. J., Kane, N. C., O'Hara, N. B., Tittes, S., & Rest, J. S. (2016). Rapid genome-wide evolution in Brassica rapa populations following drought revealed by sequencing of ancestral and descendant gene pools. *Molecular Ecology*, 25(15), 3622–3631.
- Freytag, G. F., & Debouck, D. G. (2002). *Taxonomy, distribution, and ecology of the genus phaseolus (Leguminosae-papilionoideae) in North America, Mexico and Central America*. Botanical Research Institute of Texas (BRIT).

- Fukunaga, K., Hill, J., Vigouroux, Y., Matsuoka, Y., Sanchez G, J., Liu, K., Buckler, E. S., & Doebley, J. (2005). Genetic diversity and population structure of teosinte. *Genetics*, 169(4), 2241–2254.
- García-Salazar, J. A., & Ramírez-Jaspeado, R. (2014). El mercado de la semilla mejorada de maíz (*Zea mays* L.) en México: análisis del saldo comercial por entidad federativa. *Revista Fitotecnia Mexicana*, 37(1), 69–77.
- Gassmann, A. J., Petzold-Maxwell, J. L., Clifton, E. H., Dunbar, M. W., Hoffmann, A. M., Ingber, D. A., & Keweshan, R. S. (2014). Field-evolved resistance by western corn rootworm to multiple *Bacillus thuringiensis* toxins in transgenic maize. *Proceedings of the National Academy of Sciences of the United States of America*, 111(14), 5141–5146.
- Gepts, P. (2008). Tropical Environments, Biodiversity, and the Origin of Crops. In P. H. Moore & R. Ming (Eds.), *Genomics of Tropical Crop Plants* (Vol. 1, pp. 1–20). Springer, New York, NY. [https://doi.org/10.1007/978-0-387-71219-2\\_1](https://doi.org/10.1007/978-0-387-71219-2_1)
- Green, R. E., Krause, J., Briggs, A. W., Maricic, T., Stenzel, U., Kircher, M., Patterson, N., Li, H., Zhai, W., Fritz, M. H.-Y., Hansen, N. F., Durand, E. Y., Malaspinas, A.-S., Jensen, J. D., Marques-Bonet, T., Alkan, C., Prüfer, K., Meyer, M., Burbano, H. A., ... Pääbo, S. (2010). A draft sequence of the Neandertal genome. *Science*, 328(5979), 710–722.
- Guerra-García, A., Suárez-Atilano, M., Mastretta-Yanes, A., Delgado-Salinas, A., & Piñero, D. (2017). Domestication Genomics of the Open-Pollinated Scarlet Runner Bean (L.). *Frontiers in Plant Science*, 8, 1891.
- Harlan, J. R. (1971). Agricultural origins: centers and noncenters. *Science*, 174(4008), 468–474.
- Harlan, J. R. (1992). *Crops and Man*. Crop Science Society of Amer.
- Harlan, J. R. (1998). *The Living Fields: Our Agricultural Heritage*. Cambridge University Press.
- Harris, D. R. (1990). Vavilov's concept of centres of origin of cultivated plants: its genesis and its influence on the study of agricultural origins. *Biological Journal of the Linnean Society. Linnean Society of London*, 39(1), 7–16.
- Hawkes, J. G. (1983). *The Diversity of Crop Plants*. Harvard University Press.
- Heng, L., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25(16), 2078–2079.
- Hernández-Terán, A., Wegier, A., Benítez, M., Lira, R., Sosa Fuentes, T. G., & Escalante, A. E. (2019). performance in cotton plants with different genetic backgrounds: the case of in Mexico, and its implications for germplasm conservation. *PeerJ*, 7, e7017.
- Hernández Xolocotzi, E., & Ortega Paczka, R. (1987). *Variación en maíz y cambios socioeconomicos en Chiapas, México, 1946-1971*. <http://www.sidalc.net/cgi-bin/wxis.exe/?IsisScript=CIMMYT.xis&method=post&formato=2&cantidad=1&expresion=mfn=002126>
- Herten, K., Hestand, M. S., Vermeesch, J. R., & Van Houdt, J. K. J. (2015). GBSX: a toolkit for experimental design and demultiplexing genotyping by sequencing experiments. *BMC Bioinformatics*, 16, 73.
- Hufford, M. B., Lubinsky, P., Pyhäjärvi, T., Devengenzo, M. T., Ellstrand, N. C., & Ross-Ibarra, J. (2013). The genomic signature of crop-wild introgression in maize. *PLoS Genetics*, 9(5), e1003477.
- Hufford, M. B., Xu, X., van Heerwaarden, J., Pyhäjärvi, T., Chia, J.-M., Cartwright, R. A., Elshire, R. J., Glaubitz, J. C., Guill, K. E., Kaeppler, S. M., Lai, J., Morrell, P. L., Shannon, L. M., Song, C., Springer, N. M., Swanson-Wagner, R. A., Tiffin, P., Wang, J., Zhang, G., ... Ross-Ibarra, J. (2012). Comparative population genomics of maize domestication and improvement. *Nature Genetics*, 44(7), 808–811.
- Jenkins, P. A., Song, Y. S., & Brem, R. B. (2012). *Genealogy-Based Methods for Inference of Historical Recombination and Gene Flow and Their Application in Saccharomyces cerevisiae* (J. J. Welch (Ed.)) [Data set]. <https://doi.org/10.1371/journal.pone.0046947>

- Jones, D. M. (1977). *The green revolution in America: Success or failure?* 6, 55–63.
- Jónsson, H., Schubert, M., Seguin-Orlando, A., Ginolhac, A., Petersen, L., Fumagalli, M., Albrechtsen, A., Petersen, B., Korneliusson, T. S., Vilstrup, J. T., Lear, T., Myka, J. L., Lundquist, J., Miller, D. C., Alfarhan, A. H., Alquraishi, S. A., Al-Rasheid, K. A. S., Stagegaard, J., Strauss, G., ... Orlando, L. (2014). Speciation with gene flow in equids despite extensive chromosomal plasticity. In *Proceedings of the National Academy of Sciences* (Vol. 111, Issue 52, pp. 18655–18660). <https://doi.org/10.1073/pnas.1412627111>
- Kato Y, T. A. (1997). Review of Introgression Between Maize and Teosinte. In J. A. Serratos, M. A. Willcox, & F. Castillo (Eds.), *Gene Flow Among Maize Landraces, Improved Maize Varieties, and Teosinte: Implications for Transgenic Maize* (pp. 44–53). CIMMYT.  
[https://books.google.com/books/about/Gene\\_Flow\\_Among\\_Maize\\_Landraces\\_Improved.html?hl=&id=qGVIMII0daoC](https://books.google.com/books/about/Gene_Flow_Among_Maize_Landraces_Improved.html?hl=&id=qGVIMII0daoC)
- Kermicle, J. L. (2006). A selfish gene governing pollen-pistil compatibility confers reproductive isolation between maize relatives. *Genetics*, 172(1), 499–506.
- Kermicle, J. L., & Evans, M. M. S. (2010). The *Zea mays* sexual compatibility gene *ga2*: naturally occurring alleles, their distribution, and role in reproductive isolation. *The Journal of Heredity*, 101(6), 737–749.
- Kermicle, J. L., Allen, J. P., & Others. (1990). Cross-incompatibility between maize and teosinte. *Maydica*, 35(4), 399–408.
- Khoury, C. K., Achicanoy, H. A., Bjorkman, A. D., Navarro-Racines, C., Guarino, L., Flores-Palacios, X., Engels, J. M. M., Wiersema, J. H., Dempewolf, H., Sotelo, S., Ramírez-Villegas, J., Castañeda-Álvarez, N. P., Fowler, C., Jarvis, A., Rieseberg, L. H., & Struik, P. C. (2016). Origins of food crops connect countries worldwide. *Proceedings of the Royal Society B: Biological Sciences*, 283(1832), 20160792.
- Kirchhoff, P. (1943) Mesoamérica. *Acta Americana. Revista de la Sociedad de Antropología y Geografía* 1:92–107
- Korneliusson, T. S., Albrechtsen, A., & Nielsen, R. (2014). ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics*, 15, 356.
- Kraft, K. H., Brown, C. H., Nabhan, G. P., Luedeling, E., Luna Ruiz, J. de J., Coppens d'Eeckenbrugge, G., Hijmans, R. J., & Gepts, P. (2014). Multiple lines of evidence for the origin of domesticated chili pepper, *Capsicum annuum*, in Mexico. *Proceedings of the National Academy of Sciences of the United States of America*, 111(17), 6165–6170.
- Kronforst, M. R., Hansen, M. E. B., Crawford, N. G., Gallant, J. R., Zhang, W., Kulathinal, R. J., Kapan, D. D., & Mullen, S. P. (2013). Hybridization reveals the evolving genomic architecture of speciation. *Cell Reports*, 5(3), 666–677.
- Kulathinal, R. J., Stevison, L. S., & Noor, M. A. F. (2009). The genomics of speciation in *Drosophila*: diversity, divergence, and introgression estimated using low-coverage genome sequencing. *PLoS Genetics*, 5(7), e1000550.
- Kumar, V., Lammers, F., Bidon, T., Pfenninger, M., Kolter, L., Nilsson, M. A., & Janke, A. (2017). The evolutionary history of bears is characterized by gene flow across species. *Scientific Reports*, 7, 46487.
- Kwak, M., Kami, J. A., & Gepts, P. (2009). The Putative Mesoamerican Domestication Center of Is Located in the Lerma-Santiago Basin of Mexico. *Crop Science*, 49(2), 554.
- Larson, G., Karlsson, E. K., Perri, A., Webster, M. T., Ho, S. Y. W., Peters, J., Stahl, P. W., Piper, P. J., Lingaas, F., Fredholm, M., Comstock, K. E., Modiano, J. F., Schelling, C., Agoulnik, A. I., Leegwater, P. A., Dobney, K., Vigne, J.-D., Vilà, C., Andersson, L., & Lindblad-Toh, K. (2012). Rethinking dog domestication by integrating genetics, archeology, and biogeography. *Proceedings of the National Academy of Sciences of the United States of America*, 109(23), 8878–8883.
- Louette, D., Charrier, A., & Berthaud, J. (1997). In Situ conservation of maize in Mexico: Genetic diversity and Maize seed management in a traditional community. *Economic Botany*, 51(1), 20–38.
- Luna Mena, B. M., Hinojosa Rodríguez, M. A., Ayala Garay, Ó. J., Castillo González, F., & Mejía Contreras, J. A. (2012a). Perspectivas de desarrollo de la industria

- semillera de maíz en México. *Revista Fitotecnia Mexicana*, 35(1), 1–7.
- Martin, M. D., Vieira, F. G., Ho, S. Y. W., Wales, N., Schubert, M., Seguin-Orlando, A., Ristaino, J. B., & Gilbert, M. T. P. (2016). Genomic Characterization of a South American Phytophthora Hybrid Mandates Reassessment of the Geographic Origins of Phytophthora infestans. *Molecular Biology and Evolution*, 33(2), 478–491.
- Martin, S. H., Davey, J. W., & Jiggins, C. D. (2015). Evaluating the use of ABBA-BABA statistics to locate introgressed loci. *Molecular Biology and Evolution*, 32(1), 244–257.
- Matsuoka, Y., Vigouroux, Y., Goodman, M. M., Sanchez G, J., Buckler, E., & Doebley, J. (2002). A single domestication for maize shown by multilocus microsatellite genotyping. *Proceedings of the National Academy of Sciences of the United States of America*, 99(9), 6080–6084.
- Maune, J. F., Camadro, E. L., & Erazzú, L. E. (2018). Cross-incompatibility and self-incompatibility: unrelated phenomena in wild and cultivated potatoes? *Botany*, 96(1), 33–45.
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., & DePristo, M. A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, 20(9), 1297–1303.
- Meirmans, P. G., & Hedrick, P. W. (2011). Assessing population structure: FST and related measures. *Molecular Ecology Resources*, 11(1), 5–18.
- Menkir, A., Melake-Berhan, A., The, C., Ingelbrecht, I., & Adepoju, A. (2004). Grouping of tropical mid-altitude maize inbred lines on the basis of yield data and molecular markers. *TAG. Theoretical and Applied Genetics. Theoretische Und Angewandte Genetik*, 108(8), 1582–1590.
- Mercer, K., Martínez-Vásquez, Á., & Perales, H. R. (2008). Asymmetrical local adaptation of maize landraces along an altitudinal gradient. *Evolutionary Applications*, 1(3), 489–500.
- Miramontes-Piña, C. U. (2006). Situación actual y perspectivas del maíz en México 1996-2010. *SIAP. Distrito Federal, México*.
- Monte, M. N., Burusco, M. F. R., Carboni, M. F., Castellote, M. A., Sucar, S., Norero, N. S., Colman, S. L., Massa, G. A., Colavita, M. L., & Feingold, S. E. (2018). Genetic Diversity in Argentine Andean Potatoes by Means of Functional Markers. *American Journal of Potato Research*, 95(3), 286–300.
- Morris, M. L. (2000). *Impactos del mejoramiento de maíz en América Latina, 1966-1997*. CIMMYT.
- Morris, M. L., & Bellon, M. R. (2004). Participatory plant breeding research: Opportunities and challenges for the international crop improvement system. In *Euphytica* (Vol. 136, Issue 1, pp. 21–35). <https://doi.org/10.1023/b:euph.0000019509.37769.b1>
- Murphy, K., Lammer, D., Lyon, S., Carter, B., & Jones, S. S. (2005). Breeding for organic and low-input farming systems: An evolutionary-participatory breeding method for inbred cereal grains. *Renewable Agriculture and Food Systems*, 20(1), 48–55.
- Muthuselvi, R., & Praneetha, S. (2019). Molecular basis of self-incompatibility in vegetable crops. *International Journal of Chemical Studies*, 7(4), 506–510.
- Myhr, A. I., & Traavik, T. (2002). 10.1023/A:1013814108502. *Journal of Agricultural and Environmental Ethics*, 15(1), 73–86.
- Neigel, J. E. (1997). A Comparison of Alternative Strategies for Estimating Gene Flow from Genetic Markers. *Annual Review of Ecology and Systematics*, 28(1), 105–128.
- Nei, M. (1973). Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences of the United States of America*, 70(12), 3321–3323.
- Nelson, O. E. (1953). A genic substitute for isolation in hybrid corn seed production. *Economic Botany*, 7(4), 382–384.
- Newbigin, E., Anderson, M. A., & Clarke, A. E. (1993). Gametophytic Self-



- Incompatibility Systems. *The Plant Cell*, 1315–1324.
- Norsworthy, J. K., Schwartz, L. M., & Barber, L. T. (2016). The Incidence and Ramifications of Glyphosate Resistance in Cotton. *Outlooks on Pest Management*, 27(1), 31–35.
- Ortega Paczka, R. A. (1973). *Variación en maíz y cambios socio-económicos en Chiapas, Méx.; 1946-1971* (E. Hernández Xolocotzi (Ed.)) [Master of Science, College of Postgraduates]. <http://www.sidalc.net/cgi-bin/wxis.exe/?IsisScript=UACHBC.xis&method=post&formato=2&cantidad=1&expresion=mfn=072870>
- Ortiz-García, S., Ezcurra, E., Schoel, B., Acevedo, F., Soberón, J., & Snow, A. A. (2005). Absence of detectable transgenes in local landraces of maize in Oaxaca, Mexico (2003-2004). *Proceedings of the National Academy of Sciences of the United States of America*, 102(35), 12338–12343.
- Owens, G. L., Baute, G. J., & Rieseberg, L. H. (2016). Revisiting a classic case of introgression: hybridization and gene flow in Californian sunflowers. *Molecular Ecology*, 25(11), 2630–2643.
- Padilla García, J. M. (2011). *Relaciones de incompatibilidad en razas e híbridos comerciales de maíz* (J. de J. Sánchez González (Ed.)) [Ph.D.]. Universidad de Guadalajara.
- Padilla-García, J. M., DeLaCruz-Larios, L., Avendaño-López, A. N., & Sánchez-Martínez, J. (2016). Sistemas de incompatibilidad genética de híbridos comerciales de maíz. *Revista de Sistemas Experimentales*, 3(6), 51–59.
- Padilla García, J. M., Sánchez González, J. de J., de la Cruz Larios, L., Ruiz Corral, J. A., Ron Parra, J., & Morales Rivera, M. M. (2012). Incompatibilidad gametofítica en las razas mexicanas de maíz. *Rev Mex De Cienc Agric*, 3(3), 525–537.
- Pandolfo, C. E., Presotto, A., Carbonell, F. T., Ureta, S., Poverene, M., & Cantamutto, M. (2018). Transgene escape and persistence in an agroecosystem: the case of glyphosate-resistant Brassica rapa L. in central Argentina. *Environmental Science and Pollution Research*, 25(7), 6251–6264.
- Papa, R., & Gepts, P. (2003). Asymmetry of gene flow and differential geographical structure of molecular diversity in wild and domesticated common bean (*Phaseolus vulgaris* L.) from Mesoamerica. *TAG. Theoretical and Applied Genetics. Theoretische Und Angewandte Genetik*, 106(2), 239–250.
- Perales, H. R., Brush, S. B., & Qualset, C. O. (2003). Landraces of Maize in Central Mexico: An Altitudinal Transect. *Economic Botany*, 57(1), 7–20.
- Perales R, H., Brush, S. B., & Qualset, C. O. (2003). Dynamic management of maize landraces in Central Mexico. *Economic Botany*, 57(1), 21.
- Piñeyro-Nelson, A., Van Heerwaarden, J., Perales, H. R., Serratos-Hernández, J. A., Rangel, A., Hufford, M. B., Gepts, P., Garay-Arroyo, A., Rivera-Bustamante, R., & Alvarez-Buylla, E. R. (2009b). Transgenes in Mexican maize: molecular evidence and methodological considerations for GMO detection in landrace populations. *Molecular Ecology*, 18(4), 750–761.
- Piperno, D. R., Ranere, A. J., Holst, I., Iriarte, J., & Dickau, R. (2009). Starch grain and phytolith evidence for early ninth millennium B.P. maize from the Central Balsas River Valley, Mexico. *Proceedings of the National Academy of Sciences of the United States of America*, 106(13), 5019–5024.
- Potrykus, I. (2001). Golden rice and beyond. *Plant Physiology*, 125(3), 1157–1161.
- Pressoir, G., & Berthaud, J. (2004). Patterns of population structure in maize landraces from the Central Valleys of Oaxaca in Mexico. *Heredity*, 92(2), 88–94.
- Price, M. N., Dehal, P. S., & Arkin, A. P. (2009). FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Molecular Biology and Evolution*, 26(7), 1641–1650.
- Proyecto global de maíces nativos*. (n.d.). Retrieved August 6, 2019, from <https://www.biodiversidad.gob.mx/genes/proyectoMaices.html>
- Quist, D., & Chapela, I. H. (2001). Transgenic DNA introgressed into traditional maize landraces in Oaxaca, Mexico. *Nature*, 414(6863), 541–543.
- Ranere, A. J., Piperno, D. R., Holst, I., Dickau, R., & Iriarte, J. (2009). The cultural

- and chronological context of early Holocene maize and squash domestication in the Central Balsas River Valley, Mexico. *Proceedings of the National Academy of Sciences of the United States of America*, 106(13), 5014–5018.
- Rendón-Aguilar, B., Bernal-Ramírez, L. A., Bravo-Avilez, D., & Rocha-Munive, M. G. (2019). Temporal dynamics of detected transgenes in maize landraces in their center of origin. *Revista Mexicana de Biodiversidad*, 90(0).  
<https://doi.org/10.22201/ib.20078706e.2019.90.2653>
- Rhymer, J. M., & Simberloff, D. (1996). Extinction by hybridization and introgression. *Annual Review of Ecology and Systematics*, 27(1), 83–109.
- Rieseberg, L. H. (2019). Mapping footprints of past genetic exchange [Review of *Mapping footprints of past genetic exchange*]. *Science*, 366(6465), 570–571.
- Rieseberg, L. H., Kim, S.-C., Randell, R. A., Whitney, K. D., Gross, B. L., Lexer, C., & Clay, K. (2007). Hybridization and the colonization of novel habitats by annual sunflowers. *Genetica*, 129(2), 149–165.
- Rocha-Munive, M. G., Soberón, M., Castañeda, S., Niaves, E., Scheinvar, E., Eguiarte, L. E., Mota-Sánchez, D., Rosales-Robles, E., Nava-Camberos, U., Martínez-Carrillo, J. L., Blanco, C. A., Bravo, A., & Souza, V. (2018). Evaluation of the Impact of Genetically Modified Cotton After 20 Years of Cultivation in Mexico. *Frontiers in Bioengineering and Biotechnology*, 6, 82.
- Romay, M. C., Millard, M. J., Glaubitz, J. C., Peiffer, J. A., Swarts, K. L., Casstevens, T. M., Elshire, R. J., Acharya, C. B., Mitchell, S. E., Flint-Garcia, S. A., McMullen, M. D., Holland, J. B., Buckler, E. S., & Gardner, C. A. (2013). Comprehensive genotyping of the USA national maize inbred seed bank. *Genome Biology*, 14(6), R55.
- Ryffel, G. U. (2014). Transgene flow: facts, speculations and possible countermeasures. *GM Crops & Food*, 5(4), 249–258.
- Sánchez González, J. de J., Ruiz Corral, J. A., García, G. M., Ojeda, G. R., Larios, L. D. la C., Holland, J. B., Medrano, R. M., & García Romero, G. E. (2018). Ecogeography of teosinte. *PLoS One*, 13(2), e0192676.
- Schaffer, H. E., & Johnson, F. M. (1974). Isozyme allelic frequencies related to selection and gene-flow hypotheses. *Genetics*, 77(1), 163–168.
- Sedlazeck, F. J., Rescheneder, P., & von Haeseler, A. (2013). NextGenMap: fast and accurate read mapping in highly polymorphic genomes. *Bioinformatics*, 29(21), 2790–2791.
- Serratos, J. A. (1997). *Gene Flow Among Maize Landraces, Improved Maize Varieties, and Teosinte: Implications for Transgenic Maize* (J. A. Serratos, M. A. Willcox, & F. Castillo (Eds.)). CIMMYT.
- SIAP. (2016). *Uso de tecnología y de servicios en el campo Cuadros tabulares 2016*. Servicio de Información Agroalimentaria y Pesquera .  
<https://www.gob.mx/siap/documentos/tecnificacion>
- Slatkin, M. (1985). Gene Flow in Natural Populations. In *Annual Review of Ecology and Systematics* (Vol. 16, Issue 1, pp. 393–430).  
<https://doi.org/10.1146/annurev.es.16.110185.002141>
- Slatkin, M. (1987). Gene flow and the geographic structure of natural populations. *Science*, 236(4803), 787–792.
- Smith, B. D. (1997). The Initial Domestication of Cucurbita pepoin the Americas 10,000 Years Ago. *Science*, 276(5314), 932–934.
- Sonnenfeld, D. A. (1992). Mexico's "Green Revolution," 1940-1980: Towards an Environmental History. *Environmental History Review: EHR: A Publication of the American Society for Environmental History*, 16(4), 28–52.
- Spooner, D. M., McLean, K., Ramsay, G., Waugh, R., & Bryan, G. J. (2005). A single domestication for potato based on multilocus amplified fragment length polymorphism genotyping. *Proceedings of the National Academy of Sciences of the United States of America*, 102(41), 14694–14699.
- Steele, K. A., Gyawali, S., Joshi, K. D., Shrestha, P., Sthapit, B. R., & Witcombe, J. R. (2009). Has the introduction of modern rice varieties changed rice genetic diversity in a high-altitude region of Nepal? *Field Crops Research*, 113(1), 24–30.
- Stewart, C. N., Jr, Halfhill, M. D., & Warwick, S. I. (2003). Transgene introgression



- from genetically modified crops to their wild relatives. *Nature Reviews. Genetics*, 4(10), 806–817.
- Tadeo Robledo, M., Espinosa Calderón, A., Arteaga Escamilla, I., Trejo Pastor, V., Sierra Macías, M., Valdivia Bernal, R., & Zamudio González, B. (2012). Productividad de variedades precoces de maíz de grano amarillo para Valles Altos. *Revista Mexicana de Ciencias Agrícolas*, 3(7), 1417–1423.
- Takahata, N., & Nei, M. (1984). F and g statistics in the finite island model. *Genetics*, 107(3), 501–504.
- The Heliconius Genome Consortium. (2012). Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. In *Nature* (Vol. 487, Issue 7405, pp. 94–98). <https://doi.org/10.1038/nature11041>
- Thiele, G. (1999). Informal potato seed systems in the Andes: Why are they important and what should we do with them? *World Development*, 27(1), 83–99.
- Tixier-Boichard, M., Bed'hom, B., & Rognon, X. (2011). Chicken domestication: from archeology to genomics. *Comptes Rendus Biologies*, 334(3), 197–204.
- Todesco, M., Pascual, M. A., Owens, G. L., Ostevik, K. L., Moyers, B. T., Hübner, S., Heredia, S. M., Hahn, M. A., Caseys, C., Bock, D. G., & Rieseberg, L. H. (2016). Hybridization and extinction. *Evolutionary Applications*, 9(7), 892–908.
- Trtikova, M., Lohn, A., Binimelis, R., Chapela, I., Oehen, B., Zemp, N., Widmer, A., & Hilbeck, A. (2017). Teosinte in Europe - Searching for the Origin of a Novel Weed. *Scientific Reports*, 7(1), 1560.
- Trueba Carranza, A. J. (2012). *Semillas mexicanas mejoradas de maíz: su potencial productivo*. Colegio de Postgraduados.
- van de Wouw, M., Kik, C., van Hintum, T., van Treuren, R., & Visser, B. (2010). Genetic erosion in crops: concept, research results and challenges. *Plant Genetic Resources: Characterization and Utilization*, 8(01), 1–15.
- van Heerwaarden, J. (2007). *Population genetics of traditionally managed maize : farming practice as a determinant of genetic structure and identity of maize landraces in Mexico*. Chapter 3: Determinantsof regional genetic structure in Mexican maize landraces (M. J. Kropff (Ed.)) [Ph.D.]. Wageningen University, the Netherlands.
- van Heerwaarden, J., Doebley, J., Briggs, W. H., Glaubitz, J. C., Goodman, M. M., de Jesus Sanchez Gonzalez, J., & Ross-Ibarra, J. (2010). Genetic signals of origin, spread, and introgression in a large sample of maize landraces. *Proceedings of the National Academy of Sciences*, 108(3), 1088–1092.
- Vargas-Parada, L. (2014). GM maize splits Mexico. *Nature*, 511(7507), 16–17.
- Vázquez Carrillo, M. G., Santiago Ramos, D., Salinas Moreno, Y., Rojas Martínez, I., Arellano Vázquez, J. L., Velázquez Cardelas, G. A., & Espinosa Calderón, A. (2012). Interacción genotipo-ambiente del rendimiento y calidad de grano y tortilla de híbridos de maíz en Valles Altos de Tlaxcala, México. *Revista Fitotecnia Mexicana*, 35(3), 229–237.
- Wang, M., Chen, Z., Zhang, H., Chen, H., & Gao, X. (2018). Transcriptome Analysis Provides Insight into the Molecular Mechanisms Underlying - Mediated Cross-Incompatibility in Maize. *International Journal of Molecular Sciences*, 19(6). <https://doi.org/10.3390/ijms19061757>
- Warwick, S. I., -J. Simard, M., Légère, A., Beckie, H. J., Braun, L., Zhu, B., Mason, P., Séguin-Swartz, G., & Stewart, C. N. (2003). Hybridization between transgenic Brassica napus L. and its wild relatives: Brassica rapa L., Raphanus raphanistrum L., Sinapis arvensis L., and Erucastrum gallicum (Willd.) O.E. Schulz. *TAG. Theoretical and Applied Genetics. Theoretische Und Angewandte Genetik*, 107(3), 528–539.
- Warwick, S. I., Légère, A., Simard, M. -J, & James, T. (2007). Do escaped transgenes persist in nature? The case of an herbicide resistance transgene in a weedy Brassica rapa population. *Molecular Ecology*, 17(5). <https://doi.org/10.1111/j.1365-294X.2007.03567.x>
- Wegier, A., Piñeyro-Nelson, A., Alarcón, J., Gálvez-Mariscal, A., Alvarez-Buylla, E. R., & Piñero, D. (2011). Recent long-distance transgene flow into wild populations conforms to historical patterns of gene flow in cotton (*Gossypium hirsutum*) at its centre of origin. *Molecular Ecology*, 20(19), 4182–4194.

- Wei, F., Coe, E., Nelson, W., Bharti, A. K., Engler, F., Butler, E., Kim, H., Goicoechea, J. L., Chen, M., Lee, S., Fuks, G., Sanchez-Villeda, H., Schroeder, S., Fang, Z., McMullen, M., Davis, G., Bowers, J. E., Paterson, A. H., Schaeffer, M., ... Wing, R. A. (2007). Physical and genetic structure of the maize genome reflects its complex evolutionary history. *PLoS Genetics*, 3(7), e123.
- Wellhausen, E. J. (1961). El mejoramiento del maíz en México: Avances actuales y proyección hacia el futuro. *Revista de La Sociedad Mexicana de Historia Natural*, 2, 435-462.
- Wichmann, M. C., Alexander, M. J., Soons, M. B., Galsworthy, S., Dunne, L., Gould, R., Fairfax, C., Niggemann, M., Hails, R. S., & Bullock, J. M. (2009). Human-mediated dispersal of seeds over long distances. *Proceedings. Biological Sciences / The Royal Society*, 276(1656), 523-532.
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer.
- Wolfgruber, T. K., Sharma, A., Schneider, K. L., Albert, P. S., Koo, D.-H., Shi, J., Gao, Z., Han, F., Lee, H., Xu, R., Allison, J., Birchler, J. A., Jiang, J., Dawe, R. K., & Presting, G. G. (2009). Maize centromere structure and evolution: sequence analysis of centromeres 2 and 5 reveals dynamic Loci shaped primarily by retrotransposons. *PLoS Genetics*, 5(11), e1000743.
- Wright, S. (1931). Evolution in Mendelian Populations. *Genetics*, 16(2), 97-159.
- Wright, S. (1943). Isolation by distance. *Genetics*, 28, 139-156.
- Wright, S. (1946). Isolation by distance under diverse systems of mating. *Genetics*, 31, 39-59.
- Yamasaki, M., Tenaillon, M. I., Bi, I. V., Schroeder, S. G., Sanchez-Villeda, H., Doebley, J. F., Gaut, B. S., & McMullen, M. D. (2005). A large-scale screen for artificial selection in maize identifies candidate agronomic loci for domestication and crop improvement. *The Plant Cell*, 17(11), 2859-2872.
- Yan, J., Shah, T., Warburton, M. L., Buckler, E. S., McMullen, M. D., & Crouch, J. (2009). Genetic characterization and linkage disequilibrium estimation of a global maize collection using SNP markers. *PloS One*, 4(12), e4451.
- Zapiola, M. L., Campbell, C. K., Butler, M. D., & Mallory-Smith, C. A. (2008). Escape and establishment of transgenic glyphosate-resistant creeping bentgrass *Agrostis stolonifera* in Oregon, USA: a 4-year study. *The Journal of Applied Ecology*, 45(2), 486-494.
- Zarazúa-Escobar, J. A., Almaguer-Vargas, G., & Ocampo-Ledesma, J. G. (2011). El programa de apoyos directos al campo (PROCAMPO) y su impacto sobre la gestión del conocimiento productivo y comercial de la agricultura del Estado de México. *Agricultura, Sociedad Y Desarrollo*, 8(1), 89-105.
- Zheng, X., Levine, D., Shen, J., Gogarten, S. M., Laurie, C., & Weir, B. S. (2012). A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics*, 28(24), 3326-3328.
- Zheng, Y., & Janke, A. (2018). Gene flow analysis method, the D-statistic, is robust in a wide parameter space. *BMC Bioinformatics*, 19(1), 10.
- Zizumbo-Villarreal, D., & Colunga-García Marín, P. (2010). Origin of agriculture and plant domestication in West Mesoamerica. *Genetic Resources and Crop Evolution*, 57(6), 813-825.

# Anexo 1

Datos de pasaporte de las muestras y accesiones incluidas en este trabajo.

Accesión-ID	Localidad	Estado	Accesión (ID Banco de germoplasma)	Longitud	Latitud	Altitud (msnm)	N
AmTp2015(1)	Amatlán de Quetzalcóatl	Morelos	Zea mays ssp. parviglumis allopatric 2015	-99.034083	18.982028	1678	5
BoH12015	San Cristobal Honduras	Oaxaca	Modern Variety	-97.02734	16.32723	1280	5
BoNa1943	San Miguel Etlá	Oaxaca	Bolita 1943 (CIMMYT 2267)	-96.78517	17.207304	1684	6
BoNa1970	Ejutla & Agua de Sol	Oaxaca	Bolita 1970 (CIMMYT 6019)	-96.73333333	16.46666666 7	1471	5
BoNa2003	San Cristobal Honduras	Oaxaca	Bolita 2003 (CUCBA, JSG- RMM-LCL-484)	-97.02194444	16.3236111	1320	7
BoNa2015a	San Cristobal Honduras	Oaxaca	Bolita 2015	-97.02734	16.32723	1280	6
BoNa2015b	San Cristobal Honduras	Oaxaca	Bolita 2015	-97.03871	16.32532	1129	5
BoTN2015(3)	San Cristobal Honduras	Oaxaca	Putative maize- teosinte hybrid 2015*	-97.02734	16.32723	1280	4
BoTp2003	San Cristobal Honduras	Oaxaca	Zea mays ssp. Parviglumis (CUCBA JSG- RMM-LCL-483)	-97.0325	16.3238889	1272	5
BoTp2015	San Cristobal Honduras	Oaxaca	Zea mays ssp. parviglumis	-97.02734	16.32723	1280	7
ChNa1958	Chalco	Estado de México	Chalqueño 1958 (CIMMYT 2235)	-98.8833	19.25	2247	9
ChNa1972	Ozumba de Alzate	Estado de México	Chalqueño 1972 (INIFAP 4013- MEX-612)	-98.79361	19.03917	2300	3
ChNa2003	San Pedro Nexapa	Estado de México	Chalqueño 2003 (CUCBA JSG- RMM-LCL-478)	-98.72694444	19.0827777 8	2705	5
ChNa2015a	San Pedro Nexapa	Estado de México	Chalqueño 2015	-98.7273	19.08338	2702	3
ChNa2015b	San Pedro Nexapa	Estado de México	Chalqueño 2015	-98.727	19.08214	2703	3
ChNa2015c	San Pedro Nexapa	Estado de México	Chalqueño 2015	-98.73696	19.07522	2644	2
ChTm1984	Amecameca	Estado de México	Zea mays ssp. mexicana (INIFAP 10046)	-98.76666667	19.133333 33	2390	7

Accesión-ID	Localidad	Estado	Accesión (ID Banco de germoplasma)	Longitud	Latitud	Altitud (msnm)	N
ChTm2003	San Pedro Nexapa	Estado de México	Zea mays ssp. mexicana (CUCBA JSG- RMM-LCL-477)	-98.72694444	19.0827777 8	2705	5
ChTm2014	San Pedro Nexapa	Estado de México	Zea mays ssp. mexicana	-98.824167	19.0814833	2522	4
ChTm2015	San Pedro Nexapa	Estado de México	Zea mays ssp. mexicana	-98.7273	19.08338	2522	6
MITp2015(2)	Malinalco	Estado de México	Zea mays ssp. parviglumis allopatric 2015	-99.5005	18.952722	1131	5
MsH12014	San Agustín del maíz	Michoacán	Modern Variety	-101.15015	19.9068	1825	5
MsH22015	San Agustín del maíz	Michoacán	Modern Variety	-101.17	19.895	1845	6
MsNa1952	Huandacare o	Michoacán	Mushito 1952 (INIFAP 4314- MICH-186)	-101.2666667	20	1885	3
MsNa1970	Lomas de Irapeo	Michoacán	Mushito 1970 (CIMMYT 5567)	-103.497012	19.716303	1517	5
MsNa1979	Huandacare o	Michoacán	Mushito 1979 (INIFAP 9217- MICH-497)	-101.2666667	20	1800	4
MsNa2004	Batuecas	Michoacán	Mushito 2004 (JSG-LCL-563)	-101.4702778	20.0361111 1	2180	6
MsNa2014	San Agustín del maíz	Michoacán	Mushito 2014	-101.15015	19.9068	1825	5
MsNa2015	San Agustín del maíz	Michoacán	Mushito 2015	-101.1692	19.8958	1840	6
MsTm2004	San Agustín del maíz	Michoacán	Zea mays ssp. mexicana (JSG- RMM-450-04)	-101.4702778	20.0361111 1	2180	5
MsTm2014	San Agustín del maíz	Michoacán	Zea mays ssp. mexicana	-101.15015	19.9068	1825	1
MsTm2015	San Agustín del maíz	Michoacán	Zea mays ssp. mexicana	-101.1692	19.8958	1840	6
PpNa1943	Huitzucó	Guerrero	Pepitilla 1943 (CIMMYT 219)	-99.330802	18.290454	982	5
PpNa1960	Huitzucó	Guerrero	Pepitilla 1960 (CIMMYT 54)	-99.330802	18.290454	982	6
PpNa1973	Ixcateopán	Guerrero	Pepitilla 1973 (INIFAP 2359- GRO-280)	-99.73139	18.40833	2201	3
PpNa2003a	Huitzucó	Guerrero	Pepitilla 2003 (JSG-RMM-LCL- 508)	-99.19083333	18.248888 89	1205	8

Accesión-ID	Localidad	Estado	Accesión (ID Banco de germoplasma)	Longitud	Latitud	Altitud (msnm)	N
PpNa2003b	Ixcateopán	Guerrero	Pepitilla 2003 (JSG-RMM-LCL- 496) Zea mays ssp. parviglumis	-99.78444444	18.5047222 2	1890	6
PpTp1978	Teloloapán	Guerrero	(INIFAP 10002-C- 10-78) Zea mays ssp. parviglumis	-99.9	18.4	1450	1
PpTp1978	Ixcateopán	Guerrero	(INIFAP 10000-C- 8-78) Zea mays ssp. parviglumis	-99.8	18.5	1920	1
PpTp1984a	Teloloapán	Guerrero	(INIFAP 10062- JSG-LOS-117) Zea mays ssp. parviglumis	-99.9	18.4	1465	4
PpTp1984b	Ixcateopán	Guerrero	(INIFAP 10058- JSG-LOS-111) Zea mays ssp. parviglumis	-99.78333333	18.5	1795	1
PpTp2003a	Huitzucó	Guerrero	(JSG-RMM-LCL- 504) Zea mays ssp. parviglumis	-99.19083333	18.2488888 9	1205	3
PpTp2003b	Ixcateopán	Guerrero	(JSG-RMM-LCL- 495) Zea mays ssp. parviglumis	-99.78444444	18.5047222 2	1890	6
PpTp2010	Huitzucó	Guerrero	Zea mays ssp. parviglumis	-99.218	18.23769	1101	3
PpTp2011	Teloloapán	Guerrero	Zea mays ssp. parviglumis	-99.90842	18.411	1439	5
SJH12015	San Jeronimo Ayotlán	Jalisco	Zea mays ssp. mexicana allopatric 2015	-102.356356	20.418405	1577	5
SJTm2002	San Jeronimo Ayotlán	Jalisco	Zea mays ssp. mexicana allopatric 2002	-102.3333333	20.4166666 7	1597	6
SJTm2015	San Jeronimo Ayotlán	Jalisco	Zea mays ssp. mexicana allopatric 2015	-102.356356	20.418405	1577	5
TbH12015	San Lorenzo	Jalisco	Modern Variety	-103.9885	19.9614	1071	6
TbNa1946a (El Grullo)	El Grullo	Jalisco	Tabloncillo 1946 (CIMMYT 313)	-104.2	19.8	870	8
TbNa1946b (Cd.Guzmán)	Ciudad Guzmán	Jalisco	Tabloncillo 1946 (CIMMYT 598)	-103.497012	19.716303	1517	5
TbNa1960	Ciudad Guzmán	Jalisco	Tabloncillo 1960 (CIMMYT 347)	-103.497012	19.716303	1517	6
TbNa2003	San Lorenzo	Jalisco	Tabloncillo 2003 (RMM-15)	-103.9952778	19.9433333 3	982	5
TbNa2014	San Lorenzo	Jalisco	Tabloncillo 2014	-103.99635	19.94355	982	3
TbNa2015	San Lorenzo	Jalisco	Tabloncillo 2015	-103.9885	19.9614	1071	5

Accesión-ID	Localidad	Estado	Accesión (ID Banco de germoplasma)	Longitud	Latitud	Altitud (msnm)	N
TbTp1983	El palmar	Jalisco	Zea mays ssp. parviglumis (INIFAP 10036- JSG-LOS-73)	-104.0666667	19.95	980	4
TbTp2003	San Lorenzo	Jalisco	Zea mays ssp. parviglumis (RMM-13)	-103.9952778	19.943333 33	982	5
TbTp2014	San Lorenzo	Jalisco	Zea mays ssp. parviglumis	-103.99635	19.94355	982	3
TbTp2015	San Lorenzo	Jalisco	Zea mays ssp. parviglumis	-103.9885	19.9614	1071	5
VaNa1946	Huetamo	Michoacán	Vandeño 1946 (CIMMYT 5332)	18.615695	100.90177 1	298	6
VaNa2003	Huetamo	Michoacán	Vandeño 2003 (JSG-RMM-LCL- 521)	-100.9161111	18.820555 56	672	5
VaNa2015	Huetamo	Michoacán	Vandeño 2015 (Probably recyflad VM seeds)	-100.8998	18.8339	721	7
VaTp2003	Huetamo	Michoacán	Zea mays ssp. parviglumis ( JSG-RMM-457- 2002)	-100.9161111	18.820555 56	672	3
ZmH12015	Churintzio	Michoacán	Modern Variety	-102.098333	20.198117	1779	5
ZmH22015	Churintzio	Michoacán	Modern Variety	-102.1011	20.192	1763	5
ZmH32015	Churintzio	Michoacán	Modern Variety	-102.098333	20.198117	1779	5
ZmH42015	Churintzio	Michoacán	Modern Variety	-102.09175	20.185483	1779	5
ZmNa1944	Penjamillo	Michoacán	Zamorano 1944 (INIFAP 4191- MICH-66)	-101.95	20.1	1645	7
ZmNa1979	Chilchota	Michoacán	Zamorano 1979 (INIFAP 9203- MICH-483 )	-102.1166667	19.85	1720	3
ZmNa2004	Churintzio	Michoacán	Zamorano 2004 (JSG-LCL-561)	-102.060833	20.078333	1910	5
ZmNa2014	Churintzio	Michoacán	Zamorano 2014	-102.06897	20.15188	1834	5
ZmNa2015	Churintzio	Michoacán	Zamorano 2015	-102.07424	20.1503	1848	5
ZmTm1983	Churintzio	Michoacán	Zea mays ssp. mexicana (INIFAP 10026- JSG-LOS-48)	-102.0833333	20.15	1800	5
ZmTm2004	Churintzio	Michoacán	Zea mays ssp. mexicana (JSG- RMM-438-04)	-102.060833	20.078333	1910	6



Accesión-ID	Localidad	Estado	Accesión (ID Banco de germoplasma)	Longitud	Latitud	Altitud (msnm)	N
ZmTm2014	Churintzio	Michoacán	<i>Zea mays</i> ssp. mexicana	-102.06897	20.15188	1834	3
ZmTm2015	Churintzio	Michoacán	<i>Zea mays</i> ssp. mexicana	-102.075	20.1666	1820	5
ZmTm2015	Churintzio	Michoacán	<i>Zea mays</i> ssp. mexicana	-102.0742333	20.1503	1820	6
JaZd1978	Cuautitlán	Jalisco	<i>Zea diploperennis</i> (CIMMYT10003)	-104.2166667	19.53333333 3	2250	7
JaZdJSG2003	Cuautitlán	Jalisco	<i>Zea diploperennis</i> (JSG-RMM-LCL-551)	-104.2780556	19.5905555 6	1870	5

(1) De acuerdo con nuestros resultados la población de AmTp2015 (Amatlán, Morelos) se agrupa con *Zea mays* ssp. *parviglumis*. Sin embargo, no existe una clasificación definitiva para esta población.

(2) La población de Malinalco tiene una alta proporción de *Zea mays* ssp. *parviglumis*. Sin embargo no hay una clasificación definitiva para esta población.

(3) Estos individuos fueron clasificados como raza Bolita, pero no se incluyeron en los modelos de flujo genético evaluados.

## Anexo 2

Modelos empleados para evaluar el flujo genético de los dos grupos de variedades modernas 1 y 2 (VM1 y VM2, respectivamente) hacia razas (RNs) y parientes silvestres (PSs) simpátricos colectados en 2014-2015.

Altitud (msnm)	H1	H2	H3	H4	D	Z	p Valor	Significancia estadística	Límite de significancia después de la corrección de Bonferroni
	RN<1960	RN 1960-1980	VM1	Zd	-0.008287	-4.480013	0.000007	*	0.001612903
	RN<1960	RN>2000	VM1	Zd	0.024032	15.195294	0	*	0.001612903
	RN1960-1980	RN>2000	VM1	Zd	0.0318	19.175892	0	*	0.001612903
	Zpr1980	Zpr>2000	VM1	Zd	-0.020839	-6.241593	0	*	0.004166667
	Zmx1980	Zmx>2000	VM1	Zd	0.024102	8.035662	0	*	0.001666667
	Zmx1980	ZmxAlo	VM1	Zd	-0.041687	-10.456314	0	*	0.001666667
	Zmx>2000	ZmxAlo	VM1	Zd	-0.066214	-21.855613	0	*	0.001666667
	RN<1960	RN 1960-1980	VM2	Zd	-0.010764	-5.361846	0	*	0.001612903
	RN<1960	RN>2000	VM2	Zd	0.017933	10.524891	0	*	0.001612903
	RN 1960-1980	RN>2000	VM2	Zd	0.028463	16.407875	0	*	0.001612903
	Zpr1980	Zpr>2000	VM2	Zd	-0.021781	-6.22325	0	*	0.004166667
	Zmx1980	Zmx>2000	VM2	Zd	0.021183	6.789144	0	*	0.001666667
	Zmx1980	ZmxAlo	VM2	Zd	-0.038636	-9.110458	0	*	0.001666667
	Zmx>2000	ZmxAlo	VM2	Zd	-0.058835	-18.97634	0	*	0.001666667
	Zmx1980	Zmx>2000	RN>2000	Zd	0.02806	10.083215	0	*	0.016666667
	Zmx>2000	ZmxAlo	RN>2000	Zd	-0.06216	-16.003632	0	*	0.016666667
Unidad de análisis RN<1960 - RN 2015									
2702	ChNa1958	ChNa2015	VM1	Zd	-0.001342	-0.401428	0.688105	Not	9.35E-05
1848	ZmNa1944	ZmNa2015	VM1	Zd	0.055965	16.033591	0	*	0.000833333
1840	MsNa1952	MsNa2015	VM1	Zd	0.028572	7.512065	0	*	0.000833333
1890	PpNa1943	PpNa2003b	VM1	Zd	-0.02209	-5.517421	0	*	0.00047619
1205	PpNa1943	PpNa2003a	VM1	Zd	0.020145	5.664354	0	*	0.00047619
1280	BoNa1943	BoNa2015	VM1	Zd	0.076672	21.729999	0	*	0.000833333
1071	TbNa1946a	TbNa2015	VM1	Zd	0.119777	31.798555	0	*	0.00047619
1071	TbNa1946b	TbNa2015	VM1	Zd	0.158882	36.933912	0	*	0.00047619
721	VaNa1943	VaNa2015	VM1	Zd	0.183309	41.273831	0	*	0.000248756
2702	ChNa1958	ChNa2015	VM2	Zd	-0.002774	-0.800584	0.423373	Not	9.35E-05
1848	ZmNa1944	ZmNa2015	VM2	Zd	0.060873	14.596478	0	*	0.000833333
1840	MsNa1952	MsNa2015	VM2	Zd	0.018897	4.413926	0.00001	*	0.000833333
1890	PpNa1943	PpNa2003b	VM2	Zd	-0.024808	-6.09141	0	*	0.00047619
1205	PpNa1943	PpNa2003a	VM2	Zd	0.015528	4.053485	0.00005	*	0.00047619
1280	BoNa1943	BoNa2015	VM2	Zd	0.074645	17.197648	0	*	0.000833333
1071	TbNa1946a	TbNa2015	VM2	Zd	0.090895	20.757376	0	*	0.00047619

Altitud (msnm) RN 2015 or PS 2015	H1	H2	H3	H4	D	Z	p Valor	Significancia estadística	Límite de significancia después de la corrección de Bonferroni
1071	TbNa1946b	TbNa2015	VM2	Zd	0.130421	26.757641	0	*	0.00047619
721	VaNa1943	VaNa2015	VM2	Zd	0.34904	73.461675	0	*	0.000248756
Unidad de análisis RN 1980 - RN 2015									
2702	ChNa1972	ChNa2015	VM1	Zd	-0.024769	-4.903818	0.000001	*	9.35E-05
1848	ZmNa1979	ZmNa2015	VM1	Zd	0.053922	13.868437	0	*	0.000833333
1840	MsNa1979	MsNa2015	VM1	Zd	0.026584	7.00883	0	*	0.000833333
Unidad de análisis RN 1980 - RN 2015									
1890	PpNa1960	PpNa2003b	VM1	Zd	-0.003875	-1.085932	0.277509	Not	0.00047619
1205	PpNa1960	PpNa2003a	VM1	Zd	0.038952	11.803186	0	*	0.00047619
1890	PpNa1973	PpNa2003b	VM1	Zd	0.00774	1.843039	0.065323	Not	0.00047619
1205	PpNa1973	PpNa2003a	VM1	Zd	0.050744	12.133471	0	*	0.00047619
1280	BoNa1970	BoNa2015	VM1	Zd	0.070321	21.376411	0	*	0.000833333
1071	TbNa1960	TbNa2015	VM1	Zd	0.154168	34.890898	0	*	0.00047619
2702	ChNa1972	ChNa2015	VM2	Zd	-0.020586	-3.849977	0.000118	Not	9.35E-05
1848	ZmNa1979	ZmNa2015	VM2	Zd	0.051586	11.671773	0	*	0.000833333
1840	MsNa1979	MsNa2015	VM2	Zd	0.026224	5.982135	0	*	0.000833333
1890	PpNa1960	PpNa2003b	VM2	Zd	-0.008121	-2.199891	0.027815	Not	0.00047619
1205	PpNa1960	PpNa2003a	VM2	Zd	0.032431	9.368245	0	*	0.00047619
1890	PpNa1973	PpNa2003b	VM2	Zd	0.009303	2.128797	0.033271	Not	0.00047619
1205	PpNa1973	PpNa2003a	VM2	Zd	0.048464	11.437568	0	*	0.00047619
1280	BoNa1970	BoNa2015	VM2	Zd	0.068274	17.833988	0	*	0.000833333
1071	TbNa1960	TbNa2015	VM2	Zd	0.126984	27.946086	0	*	0.00047619
Unidad de análisis RN 2003 - RN 2015									
2702	ChNa2003	ChNa2015	VM1	Zd	-0.004847	-1.596648	0.110344	Not	9.35E-05
1848	ZmNa2004	ZmNa2015	VM1	Zd	0.062717	18.413997	0	*	0.000833333
1840	MsNa2004	MsNa2015	VM1	Zd	0.01561	4.72616	0.000002	*	0.000833333
1280	BoNa2003	BoNa2015	VM1	Zd	0.041215	13.339116	0	*	0.000833333
1071	TbNa2003	TbNa2015	VM1	Zd	0.088824	23.614443	0	*	0.00047619
721	VaNa2003	VaNa2015	VM1	Zd	0.042156	10.427473	0	*	0.001666667
2702	ChNa2003	ChNa2015	VM2	Zd	0.001127	0.366161	0.714245	Not	9.35E-05
1848	ZmNa2004	ZmNa2015	VM2	Zd	0.067556	17.260138	0	*	0.000833333
1840	MsNa2004	MsNa2015	VM2	Zd	0.007102	1.926898	0.053992	Not	0.000833333
1280	BoNa2003	BoNa2015	VM2	Zd	0.036196	10.60275	0	*	0.000833333
1071	TbNa2003	TbNa2015	VM2	Zd	0.062262	15.284395	0	*	0.00047619
721	VaNa2003	VaNa2015	VM2	Zd	0.046699	10.186425	0	*	0.001666667

Altitud (msnm) RN 2015 or PS 2015	H1	H2	H3	H4	D	Z	p Valor	Significancia estadística	Límite de significancia después de la corrección de Bonferroni
Unidad de análisis PS 1980 - PS 2015									
2522	ChTm1984	ChTm2015	VM1	Zd	0.014591	4.016982	0.000059	*	0.001666667
1820	ZmTm1983	ZmTm2015	VM1	Zd	0.014093	2.604482	0.009201	Not	0.001666667
1890	PpTp1984a	PpTp2003b	VM1	Zd	0.013724	3.020606	0.002523	Not	0.000110375
1205	PpTp1984a	PpTp2003a	VM1	Zd	-0.0799	-12.18395	0	*	0.000833333
1101	PpTp1984a	PpTp2010	VM1	Zd	-0.070469	-11.306946	0	*	0.000833333
982	TbTp1983	TbTp2003	VM1	Zd	-0.029353	-4.723257	0.000002	*	0.001666667
2522	ChTm1984	ChTm2015	VM2	Zd	0.013037	3.391936	0.000694	*	0.001666667
Unidad de análisis PS 1980 - PS 2015									
1820	ZmTm1983	ZmTm2015	VM2	Zd	0.012711	2.252715	0.024277	Not	0.001666667
1890	PpTp1984a	PpTp2003b	VM2	Zd	0.016318	3.557798	0.000374	Not	0.000110375
1205	PpTp1984a	PpTp2003a	VM2	Zd	-0.073667	-11.433777	0	*	0.000833333
1101	PpTp1984a	PpTp2010	VM2	Zd	-0.060633	-9.418095	0	*	0.000833333
982	TbTp1983	TbTp2003	VM2	Zd	-0.034559	-5.25627	0	*	0.001666667
Unidad de análisis PS (2003/2004) - PS 2015									
2522	ChTm2003	ChTm2015	VM1	Zd	0.012822	2.877701	0.004006	Not	0.001666667
1820	ZmTm2004	ZmTm2015	VM1	Zd	-0.03474	-6.835837	0	*	0.001666667
1840	MsTm2004	MsTm2015	VM1	Zd	-0.040576	-6.823226	0	*	0.004166667
1280	BoTp2003	BoTp2015	VM1	Zd	-0.011295	-1.911071	0.055995	Not	0.001666667
1101	PpTp2003a	PpTp2010	VM1	Zd	0.010167	2.004586	0.045007	Not	0.000833333
1101	PpTp2003b	PpTp2010	VM1	Zd	-0.085578	-13.602248	0	*	0.000833333
2522	ChTm2003	ChTm2015	VM2	Zd	0.012532	2.708388	0.006761	Not	0.001666667
1820	ZmTm2004	ZmTm2015	VM2	Zd	-0.030976	-5.757694	0	*	0.001666667
1840	MsTm2004	MsTm2015	VM2	Zd	-0.044013	-7.225905	0	*	0.004166667
1280	BoTp2003	BoTp2015	VM2	Zd	-0.005832	-0.907886	0.363938	Not	0.001666667
1101	PpTp2003a	PpTp2010	VM2	Zd	0.016495	3.161088	0.001572	Not	0.000833333
1101	PpTp2003b	PpTp2010	VM2	Zd	-0.076343	-11.988195	0	*	0.000833333

## Anexo 3

Heterocigosis observada (Ho) y estimada (Hs) para las accesiones incluidas en este trabajo. DS = Desviación estándar.

<b>Población / Accesión</b>	<b>Promedio Ho</b>	<b>DS Ho</b>	<b>Varianza Ho</b>	<b>Promedio Hs</b>	<b>DS Hs</b>	<b>Varianza Hs</b>	<b>Límite inferior Hs</b>	<b>Límite superior Hs</b>
JaZd1980	0.060	0.143	0.020	0.091	0.178	0.032	0.323	0.332
JaZdJSG	0.046	0.141	0.020	0.072	0.177	0.031	0.331	0.342
AmTp2015	0.062	0.156	0.024	0.081	0.172	0.030	0.233	0.243
MITp2015	0.084	0.181	0.033	0.100	0.185	0.034	0.147	0.155
BoTN2015	0.074	0.178	0.032	0.082	0.170	0.029	0.085	0.094
BoTp2003	0.048	0.134	0.018	0.082	0.182	0.033	0.408	0.416
BoTp2015	0.051	0.120	0.014	0.095	0.177	0.031	0.464	0.470
PpTp1978	0.087	0.279	0.078	0.075	0.222	0.049	-11.469	-10.587
PpTp1984a	0.058	0.149	0.022	0.099	0.202	0.041	0.404	0.414
PpTp1984b	0.082	0.274	0.075	NA	NA	NA	NA	NA
PpTp2003a	0.062	0.175	0.031	0.086	0.201	0.040	0.248	0.260
PpTp2003b	0.078	0.161	0.026	0.107	0.187	0.035	0.265	0.273
PpTp2010	0.061	0.169	0.029	0.093	0.211	0.044	0.319	0.329
PpTp2011	0.065	0.145	0.021	0.104	0.191	0.036	0.375	0.383
TbTp1983	0.039	0.132	0.017	0.066	0.176	0.031	0.389	0.401
TbTp2003	0.040	0.129	0.017	0.063	0.164	0.027	0.343	0.354
TbTp2014	0.056	0.159	0.025	0.101	0.224	0.050	0.418	0.428

<b>Población / Accesión</b>	<b>Promedio Ho</b>	<b>DS Ho</b>	<b>Varianza Ho</b>	<b>Promedio Hs</b>	<b>DS Hs</b>	<b>Varianza Hs</b>	<b>Límite inferior Hs</b>	<b>Límite superior Hs</b>
TbTp2015	0.064	0.141	0.020	0.121	0.201	0.041	0.471	0.478
VaTp2003	0.062	0.166	0.028	0.098	0.214	0.046	0.346	0.356
ChTm1984	0.068	0.143	0.021	0.104	0.182	0.033	0.340	0.347
ChTm2003	0.051	0.133	0.018	0.091	0.191	0.037	0.435	0.444
ChTm2014	0.068	0.146	0.021	0.133	0.217	0.047	0.485	0.492
ChTm2015	0.070	0.149	0.022	0.104	0.187	0.035	0.322	0.329
MsTm2004	0.052	0.134	0.018	0.090	0.187	0.035	0.416	0.425
MsTm2014	0.087	0.282	0.080	NA	NA	NA	NA	NA
MsTm2015	0.069	0.155	0.024	0.103	0.192	0.037	0.326	0.334
ZmTm1983	0.075	0.163	0.026	0.116	0.207	0.043	0.348	0.356
ZmTm2004	0.075	0.156	0.024	0.108	0.188	0.035	0.301	0.308
ZmTm2014	0.074	0.186	0.035	0.104	0.217	0.047	0.263	0.273
ZmTm2015	0.075	0.131	0.017	0.116	0.173	0.030	0.352	0.357
SJTm2002	0.080	0.168	0.028	0.101	0.183	0.034	0.204	0.211
SJTm2015	0.083	0.184	0.034	0.095	0.184	0.034	0.123	0.131
BoNa1943	0.063	0.159	0.025	0.075	0.165	0.027	0.148	0.157
BoNa1970	0.080	0.177	0.031	0.095	0.183	0.034	0.151	0.159
BoNa2003	0.073	0.162	0.026	0.087	0.168	0.028	0.161	0.169
BoNa2015	0.075	0.146	0.021	0.090	0.157	0.025	0.160	0.165



<b>Población / Accesión</b>	<b>Promedio Ho</b>	<b>DS Ho</b>	<b>Varianza Ho</b>	<b>Promedio Hs</b>	<b>DS Hs</b>	<b>Varianza Hs</b>	<b>Límite inferior Hs</b>	<b>Límite superior Hs</b>
ChNa1958	0.065	0.143	0.020	0.086	0.166	0.027	0.241	0.248
ChNa1972	0.022	0.120	0.014	0.059	0.208	0.043	0.419	0.445
ChNa2003	0.094	0.194	0.037	0.098	0.178	0.032	0.036	0.044
ChNa2015	0.081	0.155	0.024	0.099	0.168	0.028	0.179	0.185
MsNa1952	0.057	0.166	0.028	0.086	0.207	0.043	0.288	0.301
MsNa1970	0.066	0.151	0.023	0.094	0.182	0.033	0.295	0.303
MsNa1979	0.060	0.148	0.022	0.095	0.194	0.038	0.369	0.378
MsNa2004	0.079	0.159	0.025	0.099	0.173	0.030	0.192	0.199
MsNa2014	0.068	0.152	0.023	0.095	0.180	0.032	0.274	0.282
MsNa2015	0.078	0.165	0.027	0.093	0.172	0.030	0.153	0.161
PpNa1943	0.067	0.168	0.028	0.078	0.171	0.029	0.143	0.153
PpNa1960	0.079	0.163	0.027	0.094	0.172	0.030	0.163	0.171
PpNa1973	0.074	0.204	0.041	0.092	0.222	0.049	0.152	0.165
PpNa2003a	0.070	0.143	0.021	0.091	0.164	0.027	0.230	0.236
PpNa2003b	0.076	0.152	0.023	0.111	0.184	0.034	0.313	0.321
TbNa1946a	0.062	0.146	0.021	0.080	0.162	0.026	0.215	0.223
TbNa1946b	0.053	0.138	0.019	0.084	0.179	0.032	0.356	0.364
TbNa1960	0.078	0.168	0.028	0.094	0.176	0.031	0.165	0.173
TbNa2003	0.071	0.160	0.025	0.090	0.175	0.031	0.207	0.215

<b>Población / Accesión</b>	<b>Promedio Ho</b>	<b>DS Ho</b>	<b>Varianza Ho</b>	<b>Promedio Hs</b>	<b>DS Hs</b>	<b>Varianza Hs</b>	<b>Límite inferior Hs</b>	<b>Límite superior Hs</b>
TbNa2014	0.064	0.173	0.030	0.086	0.197	0.039	0.234	0.245
TbNa2015	0.056	0.139	0.019	0.086	0.177	0.031	0.339	0.347
VaNa1943	0.060	0.145	0.021	0.093	0.178	0.032	0.348	0.356
VaNa2003	0.069	0.161	0.026	0.084	0.171	0.029	0.176	0.185
VaNa2015	0.047	0.125	0.016	0.082	0.175	0.031	0.420	0.428
ZmNa1944	0.060	0.144	0.021	0.080	0.164	0.027	0.252	0.261
ZmNa1979	0.068	0.180	0.032	0.096	0.208	0.043	0.226	0.238
ZmNa2004	0.076	0.160	0.026	0.099	0.180	0.032	0.234	0.241
ZmNa2014	0.078	0.166	0.028	0.097	0.179	0.032	0.188	0.195
ZmNa2015	0.073	0.158	0.025	0.095	0.178	0.032	0.231	0.238
BoH12015	0.053	0.139	0.019	0.083	0.180	0.032	0.362	0.371
Msh12014	0.049	0.150	0.023	0.061	0.163	0.027	0.192	0.204
Msh22015	0.064	0.150	0.023	0.097	0.181	0.033	0.332	0.340
SJH12015	0.049	0.158	0.025	0.052	0.148	0.022	0.044	0.056
TbH12015	0.059	0.158	0.025	0.070	0.160	0.026	0.152	0.161
ZmH12015	0.042	0.140	0.020	0.055	0.154	0.024	0.217	0.230
ZmH22015	0.035	0.137	0.019	0.044	0.144	0.021	0.201	0.215
ZmH32015	0.044	0.148	0.022	0.050	0.149	0.022	0.111	0.123
ZmH42015	0.051	0.146	0.021	0.068	0.167	0.028	0.242	0.252

## Anexo 4

Artículo requisito basado en este trabajo de investigación.



# Contemporary evolution of maize landraces and their wild relatives influenced by gene flow with modern maize varieties

Idalia C. Rojas-Barrera<sup>a,b,1</sup>, Ana Wegier<sup>c</sup>, José de Jesús Sánchez González<sup>d</sup>, Gregory L. Owens<sup>e,f</sup>, Loren H. Rieseberg<sup>e</sup>, and Daniel Piñero<sup>b</sup>

<sup>a</sup>Posgrado en Ciencias Bioquímicas, Universidad Nacional Autónoma de México (UNAM), C.P. 04510 Ciudad de México, México; <sup>b</sup>Departamento de Ecología Evolutiva, Instituto de Ecología, UNAM, 04510 Ciudad de México, México; <sup>c</sup>Laboratorio de Genética de la Conservación, Jardín Botánico, Instituto de Biología, UNAM, 04510 Ciudad de México, México; <sup>d</sup>Centro Universitario de Ciencias Biológicas y Agropecuarias, Universidad de Guadalajara, Nextipac Zapopan, 44600 Jalisco, México; <sup>e</sup>Department of Botany and Biodiversity Research Centre, University of British Columbia, Vancouver, BC V6T 1Z4, Canada; and <sup>f</sup>Department of Integrative Biology, University of California, Berkeley, CA 94720

Edited by Edward S. Buckler, US Department of Agriculture-Agricultural Research Service/Cornell University, Ithaca, NY, and approved September 6, 2019 (received for review October 12, 2018)

Mexico is recognized as the center of origin and domestication of maize. Introduction of modern maize varieties (MVs) into Mexico raised concerns regarding the possible effects of gene flow from MVs into maize landraces (LRs) and their wild relatives (WRs), teosintes. However, after more than 60 y from the release of the first MVs, the impact of the sympatry with LRs and their WRs has not been explored with genetic data. In this work, we assessed changes in the genomes of 7 maize LRs and 2 WR subspecies from collections spanning over 70 y. We compared the genotypes obtained by genotyping by sequencing (GBS) for LRs and WRs before and after the adoption of MVs, and observed introgression from sympatric MVs into LRs and into the WR *Zea mays ssp. mexicana* sampled after the year 2000. We also found a decrease in the paired divergence index ( $F_{ST}$ ) between MV-LR and MV-WR over the same time frame. Moreover, we determined that LR genetic diversity increased after 2000, probably as a result of gene flow from MVs introduced in the 1990s. Our findings allowed us to identify ongoing changes in the domesticated and wild maize genetic pools, and concur with previous works that have evaluated short-term gene flow from MVs into LRs in other crops. Our approach represents a useful tool for tracking evolutionary change in wild and domesticated genetic resources, as well as for developing strategies for their conservation.

modern varieties | gene flow | landraces | crop wild relatives | population genomics

The Mesoamerican centers of crop origins are characterized by crop landraces (LRs) living in sympatry with their wild relatives (WRs). Both are important genetic resources; thus, conservation of their unique genetics and minimizing their genetic erosion are priorities for the Convention on Biological Diversity and to ensure food security (1–3).

Maize is a genetically diverse crop (4–6), probably because of the broad range of environments in which it grows, as well as both ancient and contemporary introgression with sympatric WRs (5, 7–9). Contemporary gene flow with modern varieties (MVs) has been assessed for genetically modified (GM) cultivars only (10–12), without considering that in recent history, non-GM MVs have been cultivated in sympatry with both LRs and WRs. This raises the possibility that MV alleles may be spreading into LR or WR gene pools, potentially replacing indigenous alleles. Nevertheless, the impact of sympatry of Mexican LRs and WRs with MVs has not yet been explored using a wide range of temporal sampling to show the complete time line of effects.

There is evidence that adaptive introgression from the WR *Zea mays ssp. mexicana* (hereafter, *mexicana*) into maize LRs promoted the latter's adaptation to the highlands of Mexico (9). Likewise, adaptive introgression by artificial selection has been

documented for maize LRs in Italy, where changes in population structure, as well as new alleles acquired from MVs, were identified in local LRs (13). The use of MVs also has provoked concerns about the displacement of LRs by MVs and the loss of alleles present in LRs adapted to local and heterogeneous environments (1, 14–16). This occurred in fewer than 5 y in the midwestern United States, where open-pollinated maize cultivars were replaced by maize MVs (17). In contrast, it has been reported for both potato and rice that after MV introductions, genetic diversity initially increases (18–20), although this increase is followed by a decline in diversity (21, 22) that stops after the partial substitution of LRs by MVs (15).

In Mexico, the first maize MVs were open-pollinated varieties under mass selection released in 1946. These were produced by the Mexican Department of Agriculture in agreement with The Rockefeller Foundation. A few years later, in 1950, the first hybrid maize lines were released, which represented crosses between open-pollinated varieties that had been selfed for 1 generation prior to hybrid production. At this stage, the efforts of improvement programs focused on the central Mexican Plateau and the central northwest and southwest of Mexico (23). The next

## Significance

Crop diversity may be essential for adaptation to diverse future climates, and its conservation depends on human practices and preferences. Besides these, we show here that regulations promoting the adoption of modern cultivars can promote rapid changes in the genetic pools of indigenous landraces (LRs) and crop wild relatives (WRs). We compared a wide range of temporal samples of maize LRs and WRs. Modern varieties (MVs), LRs, and their WRs have been in sympatry for over 60 y. We provide genomic evidence of ongoing evolution of *Zea mays* L. due to introgression from MVs. These findings should foster monitoring strategies and policies that use and safeguard the genetic diversity of maize and its WRs at their center of origin.

Author contributions: I.C.R.-B., A.W., and D.P. designed research; I.C.R.-B. performed research; J.d.J.S.-G., G.L.O., L.H.R., and D.P. contributed new reagents/analytic tools; I.C.R.-B. analyzed data; and I.C.R.-B., L.H.R., and D.P. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Published under the PNAS license.

Data deposition: The data reported in this paper have been deposited in the Open Science Framework (OSF), <https://osf.io/pqvt4/>, and GitHub, [https://github.com/yetzehev/Ongoing\\_Evol\\_Landraces](https://github.com/yetzehev/Ongoing_Evol_Landraces).

<sup>1</sup>To whom correspondence may be addressed. Email: [icrojasb@ieecologia.unam.mx](mailto:icrojasb@ieecologia.unam.mx).

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1817664116/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1817664116/-DCSupplemental).

decade, the 1960s, has been recognized as the beginning of the green revolution that led to a worldwide production increase of different cereal grains. This revolution relied on multiple strategies, including mechanization, advances in plant genetics and breeding, synthesis and application of fertilizers and pesticides, and adoption of high-yielding modern crop varieties (24, 25). Mexico is historically linked to the green revolution because of the successful wheat breeding program headed by Norman Borlaug. In the 1960s, 2 important institutions were created: the National Research Institute in Agriculture (INIA, its acronym in Spanish) and the International Maize and Wheat Improvement Center (CIMMYT, its acronym in Spanish). However, the acceptance of maize MVs was limited and occurred gradually. By 1976, maize MVs covered less than 15% of land dedicated to maize production (26). In the southeast of Mexico, MVs were adopted and replanted as open-pollinated varieties; this process, known as creolization, led to a decrease in land dedicated to LRs (14, 16, 20, 22, 27, 28). During this period, the trade of maize MVs was controlled by the government, whereas it is currently controlled by private companies (29, 30).

The first Mexican Seed Act of 1961 established the authority of the National Seed Producer (PRONASE, its acronym in Spanish) over the trade of genetically improved seeds produced by INIA. However, the private sector found this to be restrictive and pressed for authorization to perform seed research. Permission was granted in 1976, and its first MVs were released in 1983 (30, 31). In 1991, the seed act was modified to allow unrestricted private sector participation in seed research, production, and trade, and the Mexican Seed Act of 1996 gave the private seed companies access to INIA MVs. The next decade was crucial; the seed act was modified again in 2007 and PRONASE disappeared, and by 2010, private companies had almost total control of the MV seed markets (30, 32).

Current estimates of the proportion of land used for maize production that is devoted to MVs range from 42.5 (29) to 58% (33). However, reports on the penetrance of MVs into the maize market (34) imply that the latter proportion could be overestimated (35).

Maize MVs are mainly produced and consumed in the northwest and central west of Mexico, in contrast to the rest of the country, where LRs are predominant (29). The partial adoption of MVs can be explained by the added value of maize LRs, such as sensory attributes, higher quality for specialty foods (36–38), and adaptation to marginal environments and poor soil conditions (22, 37). Consequently, LRs are the base of traditional Mexican cuisine (39, 40), while MVs are used for tortillas, for the corn flour agroindustrial market, or to feed cattle (29, 40).

The partial adoption of MVs has led to frequent sympatry with local LRs and their WRs, which are grown in neighboring fields or in the same locality. Sympatry and creolization appear to have facilitated gene flow from MVs into LRs, which is promoted by farmers in some cases. This genetic exchange has only been documented through phenotypic evidence; so far, there is no genetic evidence of its extent or consequences (20, 39, 41).

Here, we tested for gene flow from maize MVs into local LRs, and their WRs, teosintes: *Zea mays* ssp. *parviglumis* (hereafter, *parviglumis*) and *mexicana*. These wild subspecies currently are sympatric with MVs mainly in the central Mexican Plateau and the Balsas River Basin, located in the central-southwestern part of Mexico, the region recognized as the center of domestication of maize (42–44). The teosintes *mexicana* and *parviglumis* exhibit different altitudinal distributions: *parviglumis* inhabits subhumid tropical environments between 143 and 1,960 meters above sea level (masl), whereas *mexicana* is found in subtropical to temperate conditions in highlands from 1,500 to 2,990 masl (45). We hypothesized that gene flow has continuously occurred from MVs into LRs and WRs in the Mexican Plateau and the Balsas River Basin regions.

To assess levels of introgression from maize MVs into native LRs and their WRs, we compared the genotypes of individuals currently in sympatry with MVs, including 7 LRs and the 2 teosinte subspecies mentioned above. The LRs included were classified according to Anderson and Cutler (46), who defined a LR as a group of related individuals with enough characteristics in common to permit their recognition as a group. The 7 LRs were sorted by sampling period (before 1960, from 1960 to 1980, and after 2000); *mexicana* and *parviglumis* individuals were sampled during the 1980s and after 2000. We also included MVs found in sympatry with LRs and WRs and, finally, allopatric populations of WRs to assess the role of sympatry. Genotypes obtained via genotyping by sequencing (GBS) (47) were compared according to sampling time period. Comparisons of LRs and WRs sampled at different time periods allowed us to confirm gene flow from maize MVs into LRs and into their WRs, detect rapid changes in the population structure of maize LRs, and document higher nucleotide diversity in the most recent LRs and WRs.

## Materials and Methods

**Samples and Accessions.** We sampled LRs currently cultivated in sympatry with their WRs and commercial MVs, and then compared them with the oldest samples and accessions sorted into different periods in order to evaluate genetic exchange between these 3 groups.

To accomplish this aim, we examined the Native Maize Project record from the National Commission for the Knowledge and Use of Biodiversity. The purpose of this project was to update the distribution information for maize LRs and their WRs and to determine diversity centers in Mexico (48). We looked for LRs sampled in Chalco, the central Mexican Plateau, the Balsas River Basin, and the Oaxacan regions where WRs are distributed (49). We set 3 temporal periods: (1) The first was used as a negative control and included samples collected before 1960, prior to the adoption of MVs; (2) the second period was set from 1960 to 1980; and (3) the third period included samples collected after 2000. Since we were interested in potential LR-WR introgression, we selected both LR and WR samples from regions where their distributions overlap and are currently in close geographic proximity (*SI Appendix, Fig. S1*).

Seven maize LRs (Chalqueño, Zamorano, Mushito, Pepitilla, Tabloncillo, Bolita, and Vandeño) met our criteria for inclusion. We considered each sympatric group as a single analysis unit (i.e., 1 LR, 1 WR, 1 or more MVs). To analyze global patterns per sampling period, we grouped the genotypes as shown in Table 1.

Once we selected the groups of LRs and WRs, we visited the localities where these LRs were already reported as sympatric with *mexicana* or *parviglumis*. Four of the 7 sympatric LRs-WRs were growing in proximity to 1 or more commercial MVs. We included the 3 localities where we did not find MVs to test if sympatry at the time sampled was required to detect gene flow. All of the MVs included are the first generation (F1) of commercial varieties produced through controlled mating, which were collected between 2014 and 2015. As a control to test for gene flow from sympatric MVs into WRs, allopatric populations of *mexicana* and *parviglumis* were included.

For LRs and MVs sampled in 2014 to 2015, we collected at least 5 complete ears from randomly selected plants, with the exception of Vandeño LR (Huetamo, Michoacán), for which farmers could only provide us with seeds. The identity of LRs sampled in 2014 was established by Juan Manuel Hernández Casillas from the National Research Institute in Forestry, Agriculture and Livestock (INIFAP, its acronym in Spanish), and the same localities were revisited in 2015. For *mexicana* and *parviglumis*, we sampled seeds from at least 5 plants at each locality. Seeds from each locality were pooled, and 5 randomly selected seeds were germinated and genotyped. LR and WR accessions collected before or during the 1980s were obtained from CIMMYT and INIFAP. Samples from 2002 to 2004 were obtained from the University Center of Biological and Agricultural Sciences of the Universidad de Guadalajara. The passport information for all samples and accessions included in this work is available in *SI Appendix, Table S1*.

**DNA Extraction and Genotyping.** Maize and WR seeds were germinated in a CONVRON environmental chamber at 25 °C with a period of 12-h/12-h darkness and light, respectively. We used fresh leaf tissue from each plant to extract DNA using the DNeasy Plant Mini Kit (QIAGEN, Inc.). The quality of the DNA was assessed on an 0.8% agarose gel, and the DNA samples were quantified by fluorometry using Qubit 3.0 and the Quantification Starter Kit (Thermo Fisher Scientific, Inc.).



**Table 1. Time periods of sample collections: Maize LRs, MVs, and WRs**

Population	Population code	N
LR sampled between 1940 and 1959	LR < 1960	49
LR samples between 1960 and 1980	LR 1960 to 1980	35
LR sampled between 2000 and 2015	LR > 2000	106
MV sampled between 2014 and 2015	MVs	47
<i>Z. mays</i> ssp. <i>mexicana</i> sampled in 1984	Zmx1980	12
<i>Z. mays</i> ssp. <i>mexicana</i> sampled between 2000 and 2015	Zmx2000	48
Allopatric <i>Z. mays</i> ssp. <i>mexicana</i> sampled between 2002 and 2015	Allopatric <i>mexicana</i>	11
<i>Z. mays</i> ssp. <i>parviglumis</i> sampled between 1978 and 1984	Zpr1980	11
<i>Z. mays</i> ssp. <i>parviglumis</i> sampled between 2000 and 2015	Zpr2000	44
Allopatric <i>Z. mays</i> ssp. <i>parviglumis</i> sampled in 2015	Allopatric <i>parviglumis</i>	10
<i>Z. diploperennis</i>	<i>Z. diploperennis</i>	12
Total		385

Zmx, *Z. mays* ssp. *mexicana*; Zpr, *Z. mays* ssp. *parviglumis*.

One DNA sample for each plant was submitted to the Biotechnology Resource Center of Cornell University for GBS on a 96-plex plate. Samples were digested with ApeKI enzyme, and libraries were developed according to the protocol standardized for maize by Elshire et al. (47). Four samples were submitted twice and used as controls to assess the reproducibility of sequencing.

Samples included 7 native maize LRs, collected from 1943 to 2015, and their WRs, collected from 1978 to 2015; all of the MVs were collected between 2014 and 2015. We included the species *Zea diploperennis* as an outgroup. Altogether, 385 samples were genotyped in this work.

#### Bioinformatics Workflow.

**Variants discovery.** Fastq files were demultiplexed with GBSx (50), and reads were trimmed with Trimmomatic 0.36 (51). Alignments were performed with Nextgenmap 0.5.3 (52) using the B73 genome (AGPv4) (<https://www.maizedb.org/assembly>) as a reference, and all alignments were converted to binary files with Samtools 1.5 (53). Variants were discovered for each sample using the HaplotypeCaller, and genotypes were merged with GenotypeGVCFs; both tools are from the Genome Analysis Toolkit (GATK 3.8.0) (54). We filtered single-nucleotide polymorphisms (SNPs) with a minor allele frequency (MAF) of 0.01, maximum missing data of 20% per SNP, and minimum mean depth of 2X; this step was done with vcftools 0.1.15 (55), and scripts are available at Open Science Framework (OSF) (<https://osf.io/pqvt4/>) (56) and GitHub ([https://github.com/yetzehev/Ongoing\\_Evol\\_Landraces](https://github.com/yetzehev/Ongoing_Evol_Landraces)). After filtering, we kept 316,294 SNPs distributed across the 10 chromosomes and contigs in the reference genome.

**Ancestry analysis.** Admixture 1.3.0 (57) was used to infer the ancestry assignment for WRs, LRs, and MVs. We ran models from 2 to 8 genetic groups (K) and selected the K with the lowest cross-validation (CV) error. The Admixture results were used as criteria to remove accessions TbTp2014 and TbTp2015, which were collected as *parviglumis* (SI Appendix, Table S1) but assigned to *mexicana* ancestry, as well as 2 individuals collected as Vandeño LR that clustered with MVs. We ran a second ancestry assignment for 463 genotypes, including the 2 clusters that grouped the MVs collected for this work and the tropical breeding pool genotypes from the 2013 dataset of Romay et al. (65). We ran models from 1 to 20 groups and again selected the K with the lowest CV error.

**Gene flow test (ABBA-BABA).** To test for gene flow from MVs into native LRs and WRs, we calculated Patterson's *D*-score, also known as the ABBA-BABA test. The test is based on a resolved phylogeny among 4 taxa ((H1,H2), H3), H4) and determines if the proportion of derived states is influenced by gene flow.

To compute this test, we used the BAM files from each individual and ran the analysis with the multipop ABBA-BABA module from the ANGSD package (Analysis of Next Generation Sequencing Data; <https://github.com/ANGSD/angsd>) (58); the parameters can be accessed at [https://github.com/yetzehev/Ongoing\\_Evol\\_Landraces](https://github.com/yetzehev/Ongoing_Evol_Landraces). The statistical threshold employed was  $P < 0.05$  with a Bonferroni correction.

In our model, we placed LR samples from different time periods in positions H1 and H2, and then asked if 1 time period shared more derived alleles with any of the MVs in position H3, with *Z. diploperennis* as the H4. If the tree has an excess of ABBA or BABA patterns, it indicates gene flow between H3 and H2 (ABBA) or between H3 and H1 (BABA) (59–61).

We also ran models for each analysis unit represented by LRs-WRs sympatric or allopatric with MVs. Then, we tested for gene flow with MVs that were sympatric at the sampling moment, and, again, we asked whether allele sharing between LRs and MVs differed between time periods. For those LRs

that were in sympatry with only one of the 2 genetic clusters identified for MVs with the ancestry analysis, or for those that were allopatric for MVs, we used the accession ZmH12015 (SI Appendix, Table S1) to survey gene flow from MV cluster 1 (MV1) and the accession SJH12015 (SI Appendix, Table S1) to survey gene flow from MV cluster 2 (MV2).

**Divergence index ( $F_{ST}$ ) across the genome.** To survey population divergence between the MV clusters, LRs and WRs sampled at different periods, we calculated the divergence index ( $F_{ST}$ ) with custom scripts using biallelic SNPs ([https://github.com/owensgl/reformat/blob/master/vcf2vertical\\_bi\\_basic.pl](https://github.com/owensgl/reformat/blob/master/vcf2vertical_bi_basic.pl) and [https://github.com/owensgl/pop\\_gen/blob/master/SNPtable2Fst.pl](https://github.com/owensgl/pop_gen/blob/master/SNPtable2Fst.pl)). We then averaged  $F_{ST}$  values in 5-kb windows considering linkage disequilibrium decay (62) and performed a nonparametric Kruskal–Wallis test, followed by a pairwise Mann–Whitney *U* test, to survey for changes in  $F_{ST}$  over time. In addition, we calculated  $F_{ST}$  values in 20-Mb windows per chromosome to plot the distributions with ggplot2 (63).

To assess temporal population divergence at the level of each analysis unit (*Samples and Accessions*), we calculated  $F_{ST}$  between the MVs and each of the 7 LRs and their sympatric WRs (SI Appendix, Fig. S3). For the LRs Chalqueño, Pepitilla, and Vandeño, which were not found in sympatry with MVs, we calculated  $F_{ST}$  using the accession ZmH12015 for MV1 and the accession SJH12015 for MV2.

**Clustering analysis.** To assess the germplasm source of MVs sampled in sympatry with LR-WR populations, we performed SNP calling for both LRs and MVs, using Tassel 5 (64) and the SNP database Tags On Physical Map v2.7 (<http://cbsusrv04.tc.cornell.edu/users/panzea/download.aspx?filegroupid=4>), which uses B73 version 2 as reference genome (AGPv2). Then, the genotypes generated for this work were merged and compared with a comprehensive public dataset of 2,578 genotypes from the US national maize inbred seed bank (65) available for AGPv2. Data were filtered with a MAF of 0.01 and maximum missing data of 20% per SNP, for a final total of 13,953 SNPs.

A first principal component analysis (PCA) was performed for the LRs and the MVs with our set of 319,294 SNPs described in *Variants discovery*, using the R package SNPRelate v1.12.0 (66). Subsequently, to explore the genetic relationships between the MVs sampled for this work and other breeding programs, 3 additional PCAs were run for MVs and the genotypes from the US national maize inbred seed bank, using the set of 13,953 SNPs described in the previous paragraph.

In addition, we constructed 2 maximum-likelihood (ML) trees, based on the generalized time-reversible (GTR) model, for 1,002 samples and the set of 13,953 SNPs using FastTree software (67) and plotted with FigTree. The first tree (SI Appendix, Fig. S6) included MV1 and MV2, as well as US inbred seed bank genotypes having a breeding program declaration (65). The second tree was run for 463 genotypes; this dataset included MV1; MV2; and the tropical breeding pools from Mexico, Nigeria, and Cameroon (SI Appendix, Fig. S6). Both trees were rooted with *parviglumis*.

**Nucleotide diversity.** We calculated the nucleotide diversity ( $\pi$ ) in 20-, 50-, and 100-kb windows (SI Appendix, Table S2); the Watterson estimator ( $\theta_W$ ); and Tajima's *D* (SI Appendix, Table S3) with ANGSD (SI Appendix, Table S3) (58) for the LRs and WRs sorted into different periods. The  $\pi$  and  $\theta_W$  distributions were compared to assess changes in genetic diversity over time; statistical significance was evaluated with a nonparametric Kruskal–Wallis test, followed by pairwise Mann–Whitney *U* tests (SI Appendix, Fig. S6). Significance was established at  $P < 0.05$  with Bonferroni correction. The frequency distribution was plotted with ggplot2 (63).



## Results

**Ancestry Analysis (Admixture).** Our ancestry analysis for LRs, MVs, and WRs found that  $K = 7$  exhibited the lowest CV error, so we used 7 clusters for further analyses (Fig. 1). The MVs were grouped into 2 clusters; MV1 (Fig. 1, *Center*, red) was more abundant and included 6 of the 9 MVs analyzed. MV1 shares ancestry with subtropical and lowland LRs and is present in all of the temporal periods. MV2 (Fig. 1, *Upper*, yellow) included 3 MVs sampled at 2 different localities. MV2 ancestry is missing in the first time period and is only detected at low levels (0.06) in samples of Bolita LR in the second time period. However, after 2000, MV2 ancestry is detected in a low proportion in 6 of the 7 LRs, except Chalqueño, and it became more abundant in lowland LRs. Furthermore, 2 individuals collected in 2015 as Vandeño LR have ancestry that is entirely MV2 (Fig. 1, *Upper*, yellow).

We also found that LR ancestry is significantly correlated with altitude (Fig. 1, *Lower Right*), and that all LRs, except Chalqueño, include some ancestry from an MV cluster. The ancestry group that included MV1 and middle and lowland LRs (Fig. 1, *Lower Right*, red) increases its abundance over time (*SI Appendix*, Fig. S2). Meanwhile, Chalqueño LR appears as a single homogeneous group, and its ancestry significantly decreases with altitude and time (Fig. 1 and *SI Appendix*, Fig. S2).

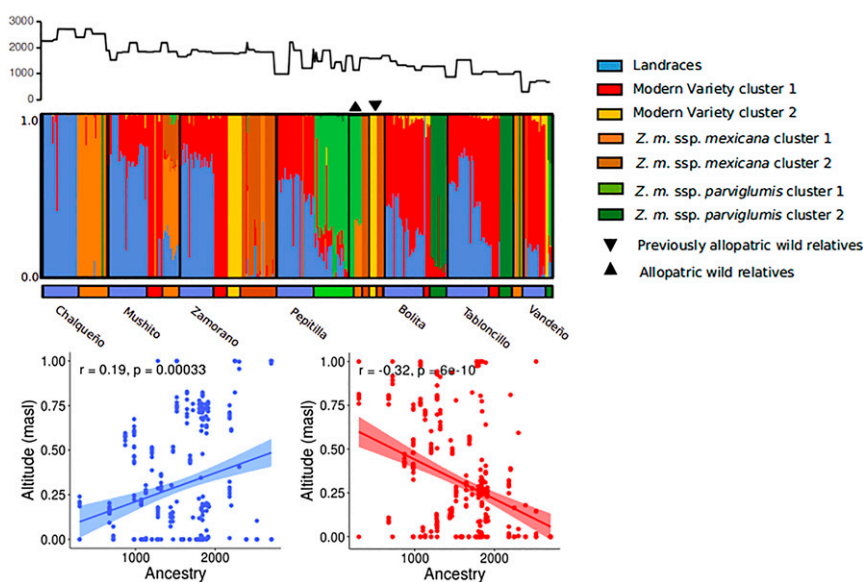
With respect to the WRs, we observed 2 genetic clusters for *mexicana* (Fig. 1, *Upper*, orange and brown) and 2 for *parviglumis* (Fig. 1, *Upper*, dark green and light green). As with the LRs, their ancestries are related to the altitude gradient; however, we did not find a difference in ancestry proportions of individuals collected prior to 1980 versus those sampled after 2000 (*SI Appendix*, Fig. S2). Surprisingly, we collected individuals assigned to *mexicana* ancestry in 2014 and 2015 from croplands near San Lorenzo, Jalisco. San Lorenzo is located in the western portion of central Mexico, a low-altitude region where *parviglumis* is typically found.

**Gene Flow Tests (ABBA-BABA).** The ancestry assignment of samples into different time periods suggests gene flow from MVs into LRs. Based on this, we ran several ABBA-BABA tests to evaluate whether the gene flow from MVs was associated with the change in population structure of LRs during the study period.

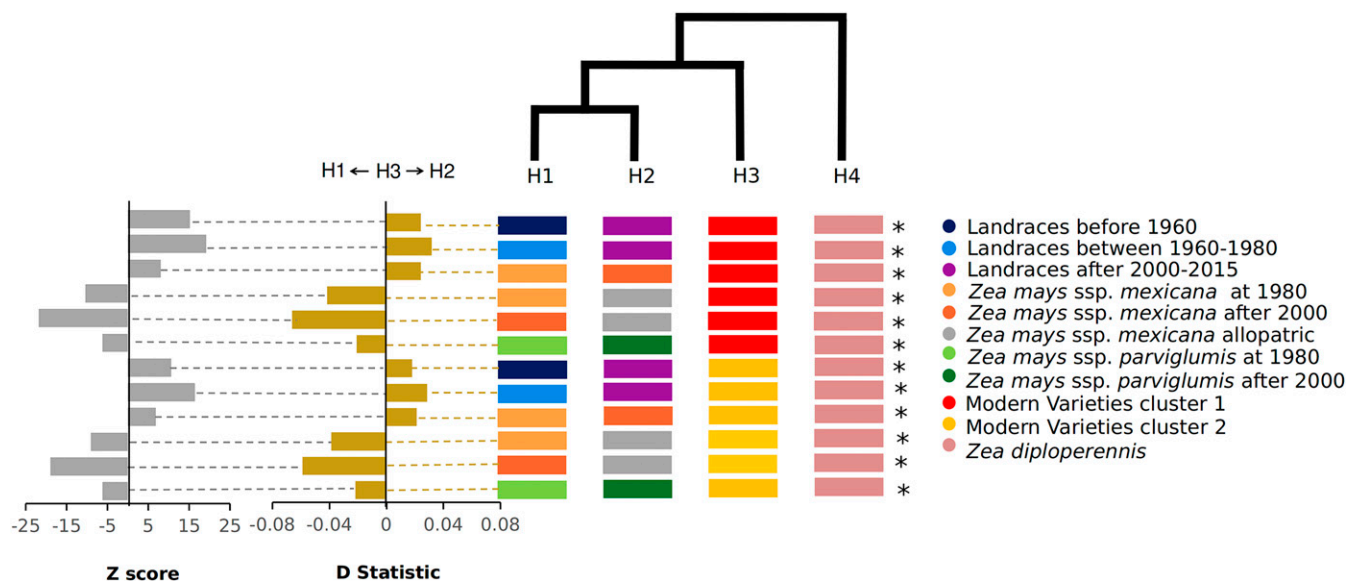
We tested 12 gene flow models to survey for gene flow from MV1 and MV2 into LR > 2000 and WR > 2000, and we set old LRs and old WRs as the sister groups (H1), respectively (Fig. 2). In effect, this test is asking if there are more MV alleles in more recent samples than in older collections. We found gene flow from MV1/MV2 into LR > 2000 ( $P < 0.0001$ ). This result was independent of the H1 group used, which was LR < 1960 or LR 1960 to 1980. We also explored the particular cases of each analysis unit. The results revealed gene flow from both MV clusters into 6 of the 7 LRs collected after 2000. The exception was Chalqueño, the highland LR sampled from 2,200 to 2,700 masl (*SI Appendix*, Table S4).

We observed gene flow from MVs into sympatric *mexicana* in both time periods, when allopatric *mexicana* was used as a negative control. This signal is stronger in the later sympatric population (>2000), as seen in *D*-score values with the allopatric comparison and when explicitly testing earlier and later *mexicana* populations ( $P < 0.05$ ). This was seen when either MV1 or MV2 was used as the H3 sample. To assess whether the gene flow signal from MVs to *mexicana* was associated with ancient and/or current exchange between *mexicana* and LRs, we set LR > 2000 in the H3 position as the donor for *mexicana* individuals collected in 1980 or after 2000. We observed gene flow from LR > 2000 into *mexicana* > 2000 (*SI Appendix*, Table S4). The *D*-score is higher when the donor group is LR > 2000 in comparison to MVs, which indicates that the gene flow signal for MVs-*mexicana* > 2000 is picking up gene flow between LRs-*mexicana*. Nevertheless, this signal is heterogeneous in time and space; when we assessed gene flow for each analysis unit, there was no gene flow from MVs into *mexicana* from the Central Plateau (ZmTm2015, *SI Appendix*, Table S4).

In addition, we detected gene flow from MV1 and MV2 into *parviglumis* sampled in 1980 ( $P < 0.05$ ); this result indicates that gene flow was higher in the past, but it does not rule out the possibility of gene flow for *parviglumis* sampled after 2000. We did not run a gene flow model for allopatric *parviglumis* populations because the classification of the Amatlán population remains unclear (9, 68), and the ancestry assignment for *parviglumis* sampled at Malinalco showed admixture with *mexicana* (Fig. 1).



**Fig. 1.** Ancestry analysis. (*Upper*) Altitudinal distributions of the samples. (*Center*) From left to right, samples are ordered from the oldest to the most recently collected, and from higher to lower altitude sampling unit. Each unit is separated by a black line, and these were named after the LRs found in sympatry with *mexicana* or *parviglumis* and MVs, ancestry for  $K = 7$  (lowest CV error). (*Lower*) Spearman rank correlation test between ancestry assignment and altitudinal distribution of samples.

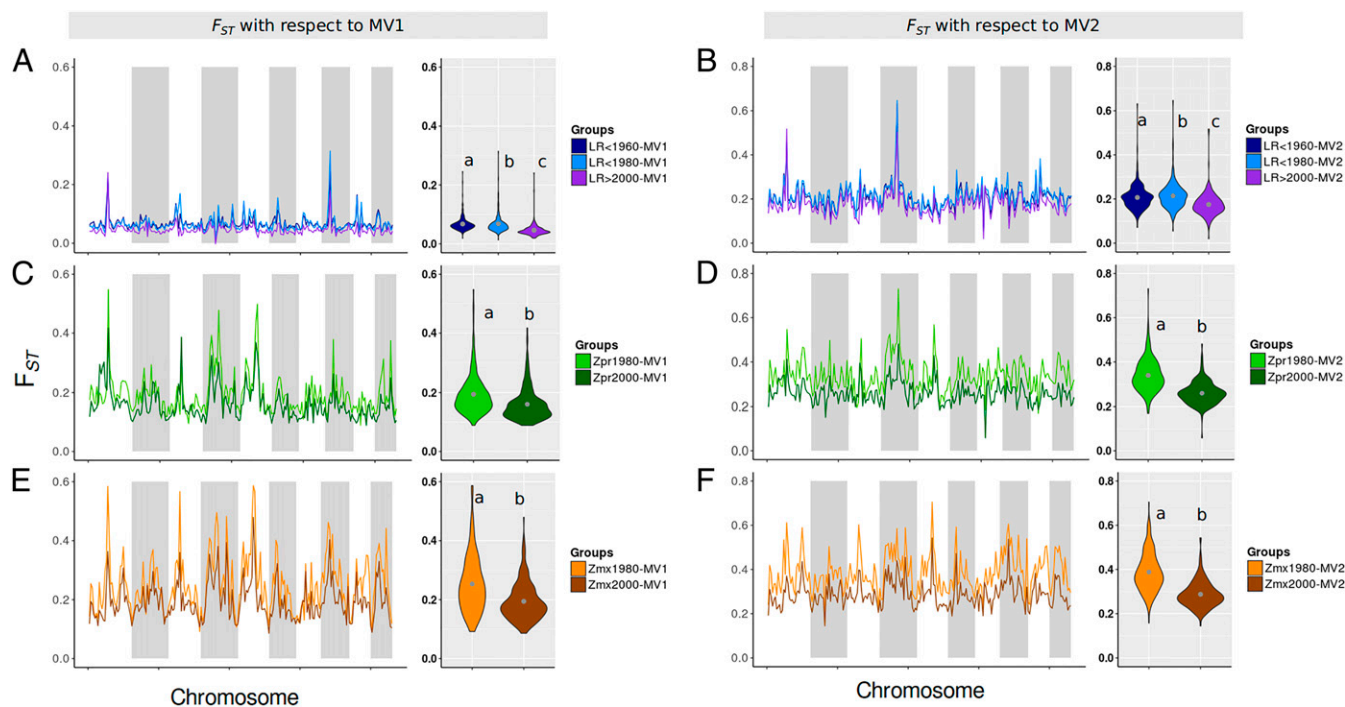


**Fig. 2.** Gene flow models (ABBA-BABA). (Right) Tree shows the phylogenetic relations assumed for the gene flow scenarios explored. Each model is represented by the color codes assigned to the different groups. \* $P < 0.05$  (statistical significance). (Left) Plots show the Z-score and Patterson's  $D$ -statistic: When  $D$  is positive, there is gene flow between H3 and H2; when  $D$  is negative, there is gene flow between H3 and H1.

**Differentiation Across the Genome ( $F_{ST}$ ).** We used the  $F_{ST}$  index to estimate the temporal genetic divergence between LR-MV and also between WR-MV. We observed that  $F_{ST}$  decreased significantly throughout the period studied, which means that allelic frequencies of LRs, *parviglumis* and *mexicana*, sampled after 2000 are more similar to MVs (Fig. 3).

In addition,  $F_{ST}$  was significantly lower between LR-MV1 than between LR-MV2 in the 3 periods evaluated (Fig. 3 *A* and *B*).

The same trend was observed when we compared  $F_{ST}$  between the 2: WR-MV1 and WR-MV2. Again, both WRs, *mexicana* and *parviglumis*, are significantly more similar genetically to MV1 than to MV2 (Fig. 3 *C-F*). When we calculated the  $F_{ST}$  for each of the 7 analysis units (SI Appendix, Fig. S3), we observed that for 6 of the 7 LRs and their sympatric WRs, there were lower differentiation estimates for MV1 compared with those found for MV2. This trend was independent of sympatry.



**Fig. 3.** Pairwise  $F_{ST}$  across the 10 chromosomes of maize and their WRs, *mexicana* and *parviglumis*, with respect to MVs.  $F_{ST}$  with respect to MV1 (Left) and MV2 (Right) is shown. The alternating shaded areas represent the 10 chromosomes of maize and their WRs; letters above the violin plots show statistical significance ( $P < 0.01$ ).  $F_{ST}$  of LRs was sampled at different periods, with respect to MV1 (A) and MV2 (B).  $F_{ST}$  of *Z. mays* ssp. *parviglumis* (Zpr) was sampled at different periods with respect to MV1 (C) and MV2 (D).  $F_{ST}$  of *Z. mays* ssp. *mexicana* (Zmx) was sampled at different periods with respect to MV1 (E) and MV2 (F).

**Genotypic Clustering.** When we ran the PCA, we observed that the distance between MV1 and MV2 was larger than the distance between MV1 and the LRs (Fig. 4A). To better understand this unexpected distribution, we used a dataset from the US seed bank, which included CIMMYT maize line (CML) genotypes from the Mexican breeding program and genotypes belonging to other breeding programs around the world (65). Both MV clusters fell within the distribution of the tropical breeding programs from Mexico, Nigeria, and Cameroon (Fig. 4B and C).

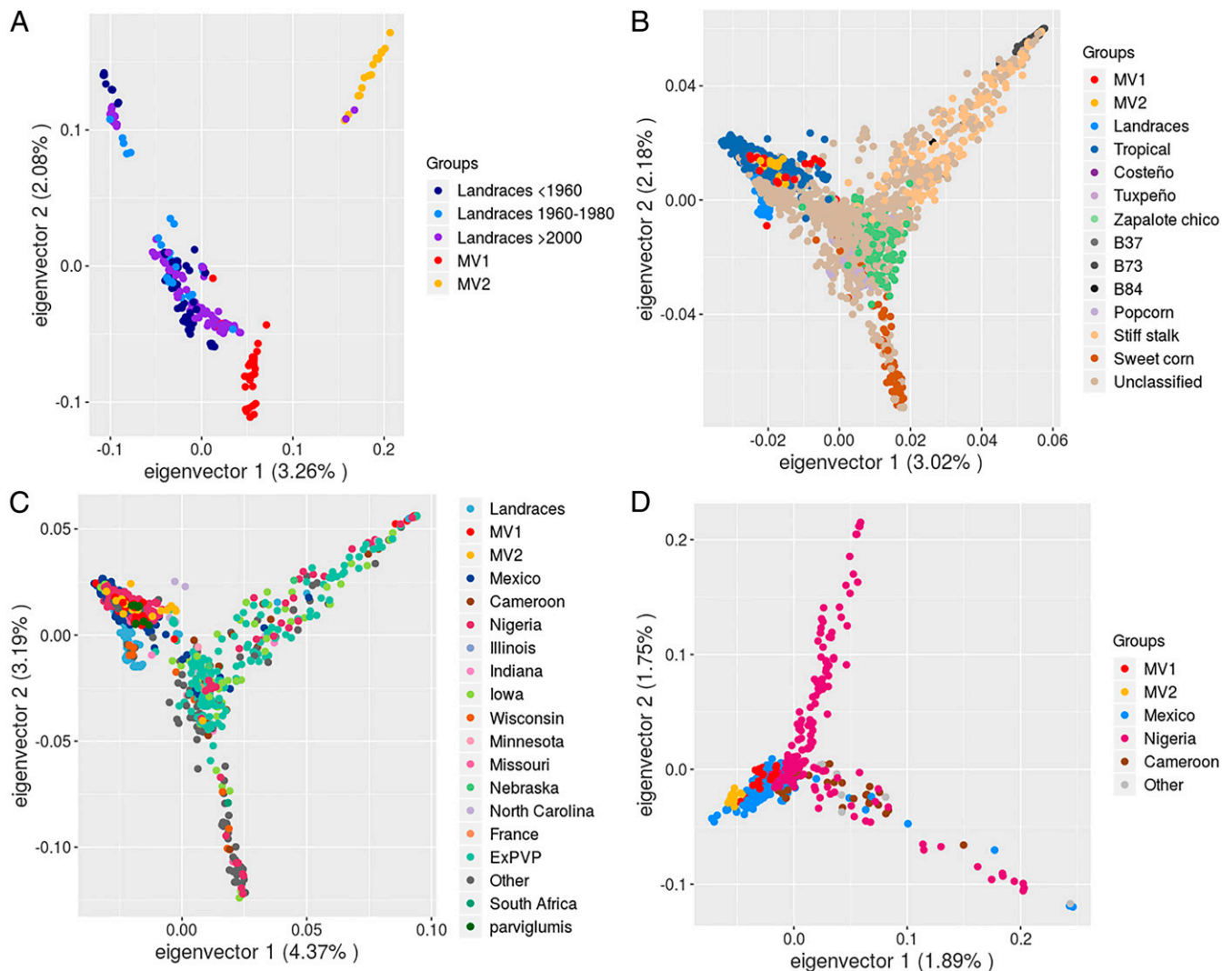
Ancestry analysis further showed that the tropical Mexican breeding program is the main germplasm source for MV1 and MV2 (SI Appendix, Fig. S4), although 2 MVs belonging to MV1 shared ancestry with genotypes from the Cameroon breeding program. These results agreed with the maximum likelihood (ML) tree for MV1, MV2, and the tropical lines, which showed that a group of CMLs is basal to both clusters (SI Appendix, Fig. S5).

**Nucleotide Diversity.** Finally, to test if changes in ancestry proportion and gene flow have had an impact on  $\pi$ , we compared LR < 1960, LR 1960 to 1980, and LR > 2000 (SI Appendix, Fig.

S5). We found that  $\pi$  for LR > 2000 ( $2.03 \times 10^{-4}$ ) is significantly higher than  $\pi$  for the oldest samples, sorted into LR < 1960 ( $1.87 \times 10^{-4}$ ) and LR 1960 to 1980 ( $1.84 \times 10^{-4}$ ). Interestingly,  $\pi$  for the intermediate period, LR 1960 to 1980, was significantly lower than for the other 2 periods. In addition,  $\pi$  for MV2 was significantly lower than  $\pi$  for MV1 (SI Appendix, Table S2). Because  $\pi$  is influenced by intermediate frequency alleles, we decided to estimate  $\theta_W$  to survey rare allele dynamics. In contrast to  $\pi$ ,  $\theta_W$  increased over the period evaluated (SI Appendix, Fig. S6). This trend matched Tajima's  $D$ , which is positive for LR < 1960 ( $0.105 \pm 0.505$ ), and negative for LR 1960 to 1980 ( $-0.032 \pm 0.453$ ) and LR > 2000 ( $-0.077 \pm 0.546$ ).

**Discussion**

**Ancestry Changes with Altitude and Sampling Time.** There was a significant correlation between the LR and WR ancestries with altitude (Fig. 1, Lower Right). This correlation concurs with previous works (4, 69), and it is consistent with the successful adoption of MVs in medium- and low-altitude regions (22, 36, 70). We observed 1 predominant ancestry group in middle and lowland LRs and



**Fig. 4.** MVs clustering with germplasm from different breeding programs. (A) Genotype distribution of MV1 and MV2 sampled in sympatry with LRs. LRs are colored by sampling period: earlier than 1960 (LR < 1960), dark blue; between 1960 and 1980 (LR 1960 to 1980); light blue; and later than 2000 (LR > 2000), purple. (B) Distribution for MV1, MV2, and the LRs collected for this work with a sample subset from the US national maize inbred seed bank (2,578 genotypes and 13,953 SNPs). (C) Clustering of genotypes colored by the breeding program (1,002 genotypes, 13,953 SNPs). ExPVP, expired plant variety protection. (D) Tropical breeding pools from Mexico, Nigeria, Cameroon, MV1, and MV2 (463 genotypes and 13,953 SNPs).



MV1. This ancestry increases its proportion in LRs over time (*SI Appendix, Fig. S2*), probably due to the continuous increase in the adoption of MVs. We also expected a broader geographic distribution for MV1 ancestry, because lowland alleles are more broadly adapted, as measured by grain yield, than highland alleles (71, 72).

The system to develop MVs changed drastically after 1990 when the Mexican Seed Act was modified (30, 32), which allowed the private sector to capture 95% of the seed market (30). The LR ancestry associated with MV2 was less prevalent than MV1 ancestry. According to the ancestry assignment with the US inbred seed bank genotypes (*SI Appendix, Fig. S4*), the MV2 genetic pool shares ancestry with the tropical breeding programs from Mexico. MV2 ancestry has increased over time, and if it is a recently introduced genetic pool, then it likely will follow the same trend as MV1. In fact, 2 individuals that were referred to as LRs by a farmer in 2015 clustered with MV2. It is likely that these individuals are the offspring of MVs that were replanted, which is a common practice known as creolization (39, 41).

Despite efforts to develop MVs genetic pools adapted to the highlands (23, 73, 74), their presence in the highlands is still scarce (34, 36). We did not find any MVs in sympatry with the Chalqueño highland LR. It has been reported that highland LRs lack tolerance to heat and inbreeding (70), features that are associated with strong local adaptation and probably have limited the introduction of MVs in the highlands. It can further explain why small farmers have adopted new technology, but they have not adopted MVs (36).

The pattern that we observed for WR ancestries (*Fig. 1, Upper*, orange and green clusters) matched the known altitudinal distribution of *mexicana* and *parviglumis* (45), except for individuals collected in San Lorenzo, Jalisco, located at 1,071 masl, a region where *parviglumis* is typically distributed. We collected teosintes at this locality in 2014 and 2015 in different croplands; these samples are clustered with subspecies *mexicana*. This unusual distribution can be attributed to the mobility of seeds fostered by human practices; it has been suggested that maize mimetic teosinte traveled from the Chalco region to Puebla and Toluca in trucks that transport manure (14). The nonintentional human transfer of teosintes also has been documented in Spain, where a nonclassified variety became a weed that damaged maize production in Western Europe and the north of Spain (75). The changes in the geographic distribution of teosintes require further evidence and a formal analysis of the conditions that promote their mobility.

**Gene Flow from MVs into LRs and WRs.** Gene flow from both MV clusters into LR > 2000 appears to be a consequence of the gradual adoption of MVs. Even though the first MVs were released in 1960, ancestry analysis does not show major changes during the 1960 to 1980 period (*SI Appendix, Fig. S2*), perhaps because the first MVs were based on high-yield LRs, such as Tuxpeño and Celaya; thus, gene flow from MV to LR would transfer alleles already present in LRs (22, 23). In 1970, the private sector began developing its own MVs, but these were not released until 1980. During this period, private companies did not have access to the seeds produced by INIA (30); therefore, MVs from this period are probably based on improved CIM-MYT lines or foreign germplasm. Beyond the sociopolitical causes, the resowing of locally grown MVs led to deliberate and/or accidental crosses between LRs and MVs (22, 39, 41). Gene flow from MVs into LRs has been previously documented with phenotypic evidence (14, 39). However, this work provides genomic evidence of the occurrence and extent of gene flow from MVs into LRs.

Specifically, we detected gene flow from MVs into 6 LRs distributed at medium and low altitudes, but not for Chalqueño, a highland LR sampled from 2,200 to 2,700 masl. The lack of gene flow into the only highland LR included can be explained

by the rarity of MVs in highland regions (34, 36). However, it is also possibly associated with the strength of local adaptation due to a higher number of maize environments per land unit in the highlands than in the midlands and lowlands (70). The tropical and subtropical MVs are adapted to more homogeneous environmental conditions, which has facilitated large-scale adoption of MVs. Under this hypothesis, MV introgressions into highland LRs are likely to be removed by selection against the generalist MV alleles versus the locally adapted LR alleles (72, 76).

However, the strength of local adaptation is a trade-off that may limit the capacity of highland LRs to respond to climate change and might make small farmers located in these regions more vulnerable (77). Conversely, the scarcity of MVs in the highlands also has been attributed to the failure of institutional breeding programs. In 2003, it was projected that in a fairly short time frame (10 to 20 y), highland LRs could be displaced by MVs (36). However, 12 y after the publication of this prediction, MVs are still rare in the highlands, and the highland genetic group had minimal changes, although we would expect that the highlands would follow the same trend for midland and lowland crop areas, if MVs are adopted at a higher frequency in the highlands.

*Z. mays* has a pollen/pistil incompatibility system that is heritable and nonreciprocal. Two main loci have been described: Gametophyte factor (*Ga1*) and Teosinte cross-barrier 1 (*Tcb1*). While we considered gametophytic incompatibility between the MVs and the Chalqueño highland LR as a possible explanation for the lack of gene flow, the codominant allele *Ga1-m* has been detected in Chalqueño and in a representative sample of MVs previously evaluated (78, 79). Thus, cross-fertilization incompatibility does not appear to account for the absence of gene flow from MVs into Chalqueño.

We selected LRs growing in sympatry with MVs when collected to increase the probability of detecting gene flow. However, our results indicate that sympatry at time of sampling was not needed to detect gene flow. For example, we observed gene flow from MVs into Vandeño and Pepitilla in the absence of MVs at sampled locations, which implies that gene flow occurred in previous generations or is a consequence of sympatry with creolized MVs, such as the 2 individuals that were collected as Vandeño but assigned genetically to MV2.

We found evidence of gene flow from MVs into *mexicana*. Our results are supported by experimental evidence: Crosses of teosintes from the Central Plateau and Chalco regions as the female parent with the MV P36D14 as male parent showed that 5 to 10% of the seeds produced were viable (80). Even when the proportion of viable seeds is low, such hybrids are sufficiently fertile to produce the signal of gene flow we detected. Furthermore, gene flow is bidirectional between *mexicana* and maize in natural populations (9), although it occurs at an asymmetrical rate that favors introgression from WRs into maize (80).

However, the gene flow test between LR > 2000-*mexicana* revealed that the gene flow signal between MVs-*mexicana* is picking up current gene flow between LRs-*mexicana*. Although this signal is heterogeneous in space, we detected gene flow from LRs into the Chalco region teosintes, but not for Central Plateau teosintes. In order to distinguish between these gene flow events, it is necessary to identify introgressed genomic regions to look for markers associated with domestication or improvement (81).

We found that the direction of gene flow signal from both MV groups goes toward *parviglumis* individuals sampled in 1980, a result suggesting that gene flow was stronger in the past. However, it does not preclude the possibility of gene flow into *parviglumis* collected after 2000, although to test this, we would require an allopatric population. Even beyond the objectives of the present paper, direct estimates of fitness and local adaptation will help to establish the consequences of introgression. Additionally, analyses of local adaptation by assessing the abundance of deleterious alleles and signatures of selection in candidate

genes would clarify if the ongoing gene flow is being counteracted by artificial and natural selection, and would help to establish the possible consequences for conservation.

**$F_{ST}$  between MVs with LRs and WRs Decays over Time.** We found that  $F_{ST}$  between MVs and LRs, and MVs with *mexicana*, decays over time (Fig. 3) and that this decay can be partially explained by modification of LR and WR genomes by recent or continuous gene flow from MVs (82). Also,  $F_{ST}$  between *parviglumis* and MVs decreased through time despite the fact that according to Patterson's  $D$ -statistics, gene flow was stronger for older *parviglumis* samples (Fig. 3 *C* and *D*). However, the effect size is small, consistent with previous observations that *parviglumis* rarely hybridizes with maize (83). Moreover, since the ABBA-BABA test is polarized with *Z. diploperennis*, it is based on a smaller number of markers than  $F_{ST}$ . Thus, the results from the ABBA-BABA test in this instance should be viewed with caution. Nonetheless, it would be of interest to identify introgressed regions in LR and WR, and to compare their divergence levels with genomic regions that have not been affected by gene flow, in order to determine whether other processes are affecting genome-wide  $F_{ST}$  values.

$F_{ST}$  between MV1 with LRs, *mexicana* and *parviglumis*, was significantly lower in all 3 periods than the divergence with MV2. This result is consistent with the early presence of MV1 and the abundance of the cultivated lines from MV1 versus MV2.

**MV1 and MV2 Appear to Be Derived from the Tropical Mexican Breeding Program.** More than 80% of MVs produced by the private sector in Mexico contain CIMMYT germplasm (84, 85), and we found that both MV1 and MV2 fell into the germplasm of the tropical breeding programs (Fig. 4). When we performed an ancestry analysis with the tropical breeding programs only (SI Appendix, Fig. S4), we observed that MV1 was more heterogeneous and appears to be mostly derived from the CIMMYT Mexican breeding program. However, a group of individuals from MV1 grouped with accessions from the Cameroon breeding program started in the 1980s. Germplasm for this program was based on inbred maize populations resistant to the streak virus, as well as other diverse sources (86). Thus some MVs collected for this work might be derived from foreign tropical breeding programs, although the Mexican breeding program appears to be the main source, as shown by the ML phylogenetic tree (SI Appendix, Fig. S5).

Finally, according to the results obtained with the PCA, ancestry assignment, and  $F_{ST}$  analysis, MV2 is a clearly differentiated group that likely derives from a few genotypes from the Mexican breeding program. Such a bottleneck and associated genetic drift may account for its reduced nucleotide diversity (SI Appendix, Table S2) and high divergence (87) (SI Appendix, Fig. S3).

**Nucleotide Diversity Changes with Time.** The nucleotide diversity for LRs decreased for the intermediate period, LR 1960 to 1980, but  $\pi$  is significantly higher for LR > 2000. The decrease of  $\pi$  in the 1960 to 1980 period was unexpected. However, it could be associated with the adoption and creolization of MVs released in 1960 that were based on a few outstanding Mexican LRs (23, 88). A study in the state of Chiapas in southeast Mexico showed that creolized MVs were planted in a higher proportion than local LRs and were resown up to 10 y after their introduction (22, 28).

On the other hand, although the participation of the private sector in trade started in 1980, the adoption of MVs was not instantaneous. Thus, it is improbable that MVs produced by the private sector were responsible for the significant decrease in  $\pi$  for LR 1960 to 1980. However, the gradual adoption of MVs over time (33) and the drastic changes in MV production after 1990 could explain the significant increase in  $\pi$  in LR > 2000. By the 1990s, 45 companies controlled 55% of the maize seed market.

There was further consolidation over the next decade; in 2009, 2 companies produced 95% of the maize MVs used (30, 31). We expected that  $\pi$  would increase after gene flow from MVs, as has been reported for potato and rice (13, 18, 89), particularly if the MVs were foreign. However, it has been reported that the trend for increased diversity is followed by a decline when LRs are displaced by MVs (1, 15).

In contrast to  $\pi$ , the estimate  $\theta_W$  tends to increase over the period evaluated (SI Appendix, Fig. S6). It matches with Tajima's  $D$  (SI Appendix, Table S3), which is positive for LR < 1960 and negative for LRs sorted into later periods. These parameters suggest a demographic expansion that is probably related to the sustained increase in demand for maize.

Although our approach was not designed to evaluate genetic erosion, our results are relevant to the controversy over genetic erosion at the center of origin of maize (90–92). Genetic erosion can be measured at crop, variety, and allele levels (15). We resampled the 7 LRs reported in the maize native record before 1960, and  $\pi$  was significantly higher for the most recent LRs (22, 39, 93). Also, note that the approach used in this work cannot detect genetic erosion at the level of specific alleles. However, the increase over time in the proportion of the predominant ancestry of MV1 in LRs, as well as the gene flow from MVs into LRs reported here, indicates that the gene pools present in LR < 1960 have been substantially modified.

Mexican agriculture is quite heterogeneous because farmers keep growing local LRs due to their sensory, aesthetic, ritualistic, and cultural values, in comparison to MVs (22, 29, 38, 39). Nevertheless, our results support a gradual replacement of traditional management systems with improved agricultural technologies, based on both MVs and their technology packages (94). This transition from community organization to crops seen only as exchange value can be socially, economically, and ecologically disadvantageous for small farmers (38), and also may imply a modernization bottleneck (15), which should be monitored in view of the nutritional and evolutionary services of small-scale farmers (35, 95).

Keeping LRs in a state of stasis is not possible and perhaps not desirable (39). However, it is possible to promote efforts to preserve wild and domesticated gene pools and allow for evolution. These efforts should include improving local LRs, encouraging seed savers networks, and developing markets for LRs that recognize their added value and the evolutionary services of small-scale farmers (1, 35).

In conclusion, our results show that the widespread adoption of MVs in Mexico has had significant impacts on the genetic composition of LRs and their WRs. We have demonstrated that there is substantial gene flow from MVs into LRs and into *mexicana*. As a result, genetic ancestry and allelic frequencies have changed across the time period evaluated. The changes detected provide short-term evidence of the ongoing evolution of maize LRs and their WRs at the center of origin of maize, and support a transition from traditional toward commercial agricultural systems, which suggests the early stages of a modernization bottleneck (1, 15). These findings can be used to design monitoring strategies and agricultural policies to reach biodiversity targets (3) for the conservation of gene reservoirs at the centers of diversity.

**ACKNOWLEDGMENTS.** We thank Azalea Guerra, Leopoldo Vázquez, Alejandro Martínez, Diana Rivera, Alicia Mastretta, Myriam Campos, Verónica Reyes, and Alfredo Villarruel for fieldwork assistance; Tania Garrido for technical laboratory assistance; Ernesto Campos Murillo for bioinformatic assistance in executing analyses in a cluster environment; Juan Manuel Hernández Casillas for LR identification; and Mauricio Bellon for reading and suggestions made to this work. We also thank the reviewers for their constructive comments and feedback on earlier versions of the manuscript. Computational analyses were performed in the computing cluster of the National Commission for the Knowledge and Use of Biodiversity (CONABIO, its acronym in Spanish), which was partially funded by the Secretaría de Medio Ambiente y Recursos Naturales through the grant "Contribución de

la Biodiversidad para el Cambio Climático” to CONABIO. We thank the CIMMYT, INIFAP, and University Center of Biological and Agricultural Sciences germplasm banks for providing old seeds included in this work, and especially the Mexican farmers who shared their seeds and knowledge. This work was supported by the Consejo Nacional de Ciencia Tecnología through PhD Scholarship 384904 (to I.C.R.-B.) and by the Secretaría de Medio

Ambiente y Recursos Naturales and CONABIO through the grant “Monitoreo de las razas criollas y de los linajes geográficos del maíz en México usando un enfoque genómico: una estrategia para su uso sustentable y su conservación” (to D.P.). This work is part of the PhD thesis of I.C.R.-B. from the Programa de Doctorado en Ciencias Bioquímicas, Universidad Nacional Autónoma de México.

1. S. B. Brush, In situ conservation of landraces in centers of crop diversity. *Crop Sci.* **35**, 346–354 (1995).
2. P. Gepts, Plant genetic resources conservation and utilization. *Crop Sci.* **46**, 2278–2292 (2006).
3. Convention on Biological Diversity, Aichi Biodiversity Targets (Convention on Biological Diversity, 2010). <https://www.cbd.int/sp/targets/>. Accessed 26 September 2018.
4. M. C. Arteaga *et al.*, Genomic variation in recently collected maize landraces from Mexico. *Genom. Data* **7**, 38–45 (2015).
5. J. F. Doebley, M. M. Goodman, C. W. Stuber, Isozymatic variation in Zea (Gramineae). *Syst. Bot.* **9**, 203–218 (1984).
6. Y. Vigouroux *et al.*, An analysis of genetic diversity across the maize genome using microsatellites. *Genetics* **169**, 1617–1630 (2005).
7. A. Eyre-Walker, R. L. Gaut, H. Hilton, D. L. Feldman, B. S. Gaut, Investigation of the bottleneck leading to the domestication of maize. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 4441–4446 (1998).
8. L. Wang *et al.*, The interplay of demography and selection during maize domestication and expansion. *Genome Biol.* **18**, 215 (2017).
9. M. B. Hufford *et al.*, The genomic signature of crop-wild introgression in maize. *PLoS Genet.* **9**, e1003477 (2013).
10. A. Piñeyro-Nelson *et al.*, Transgenes in Mexican maize: Molecular evidence and methodological considerations for GMO detection in landrace populations. *Mol. Ecol.* **18**, 750–761 (2009).
11. S. Ortiz-García *et al.*, Absence of detectable transgenes in local landraces of maize in Oaxaca, Mexico (2003–2004). *Proc. Natl. Acad. Sci. U.S.A.* **102**, 12338–12343 (2005).
12. D. Quist, I. H. Chapela, Transgenic DNA introgressed into traditional maize landraces in Oaxaca, Mexico. *Nature* **414**, 541–543 (2001).
13. E. Bitocchi *et al.*, European flint landraces grown in situ reveal adaptive introgression from modern maize. *PLoS One* **10**, e0121381 (2015).
14. J. A. Serratos, *Gene Flow Among Maize Landraces, Improved Maize Varieties, and Teosinte: Implications for Transgenic Maize*, J. A. Serratos, M. A. Willcox, F. Castillo, Eds. (CIMMYT, 1997).
15. M. van de Wouw, C. Kik, T. van Hintum, R. van Treuren, B. Visser, Genetic erosion in crops: Concept, research results and challenges. *Plant Genet. Resour.* **8**, 1–15 (2010).
16. J. van Heerwaarden, J. Hellin, R. F. Visser, F. A. van Eeuwijk, Estimating maize genetic erosion in modernized smallholder agriculture. *Theor. Appl. Genet.* **119**, 875–888 (2009).
17. J. F. Crow, 90 years ago: The beginning of hybrid maize. *Genetics* **148**, 923–928 (1998).
18. K. A. Steele *et al.*, Has the introduction of modern rice varieties changed rice genetic diversity in a high-altitude region of Nepal? *Field Crops Res.* **113**, 24–30 (2009).
19. E. Bitocchi *et al.*, Introgression from modern hybrid varieties into landrace populations of maize (*Zea mays* ssp. *mays* L.) in central Italy. *Mol. Ecol.* **18**, 603–621 (2009).
20. R. A. Ortega Paczka, Variación en maíz y cambios socioeconómicos en Chiapas, Méx. 1946–1971. M.Sc. thesis, Colegio de Postgraduados, Chapingo, Mexico (1973).
21. S. B. Brush, J. E. Taylor, M. R. Bellon, Technology adoption and biological diversity in Andean potato agriculture. *J. Dev. Econ.* **39**, 365–387 (1992).
22. M. R. Bellon, S. B. Brush, Keepers of maize in Chiapas, Mexico. *Econ. Bot.* **48**, 196–209 (1994).
23. E. J. Wellhausen, El mejoramiento del maíz en México: Avances actuales y proyección hacia el futuro. *Revista de la Sociedad Mexicana de Historia Natural* **21**, 435–462 (1961).
24. R. E. Evenson, D. Gollin, Assessing the impact of the green revolution, 1960 to 2000. *Science* **300**, 758–762 (2003).
25. G. S. Khush, Green revolution: The way forward. *Nat. Rev. Genet.* **2**, 815–822 (2001).
26. CIMMYT, CIMMYT review 1976 (CIMMYT, 1976). <https://repository.cimmyt.org/xmlui/handle/10883/3599>. Accessed 27 May 2018.
27. D. M. Jones, *The Green Revolution in America: Success or Failure?* (University of Texas Press, 1977), pp. 55–63.
28. M. R. Bellon, J. Risopoulou, Small-scale farmers expand the benefits of improved maize germplasm: A case study from Chiapas, Mexico. *World Dev.* **29**, 799–811 (2001).
29. J. A. García-Salazar, R. Ramírez-Jaspeado, El mercado de la semilla mejorada de maíz (*Zea mays* L.) en México: Análisis del saldo comercial por entidad federativa. *Rev. Fitotec. Mex.* **37**, 69–77 (2014).
30. B. M. Luna Mena, M. A. Hinojosa Rodríguez, Ó. J. Ayala Garay, F. Castillo González, J. A. Mejía Contreras, Perspectivas de desarrollo de la industria semillera de maíz en México. *Rev. Fitotec. Mex.* **35**, 1–7 (2012).
31. G. Aboites, F. Martínez, G. Torres, El negocio de la producción de semillas mejoradas y su rol en el proceso de privatización de la agricultura mexicana. *Espiral Estudios sobre Estado y Sociedad* **6**, 151–185 (1999).
32. A. Espinosa-Calderón *et al.*, Ley de Semillas y Ley Federal de Variedades Vegetales y transgénicos de maíz en México. *Rev. Mex. Cienc. Agríc.* **5**, 293–308 (2014).
33. SIAP, Uso de tecnología y de servicios en el campo Cuadros tabulares 2016 (Servicio de Información Agroalimentaria y Pesquera, 2016). <https://www.gob.mx/siap/documentos/tecnificación>. Accessed 5 June 2018.
34. M. L. Donnet *et al.*, *El potencial de mercado de semillas mejoradas de maíz en México* (CIMMYT, 2012).
35. M. R. Bellon *et al.*, Evolutionary and food supply implications of ongoing maize domestication by Mexican campesinos. *Proc. Biol. Sci.* **285**, 20181049 (2018).
36. R. H. Perales, S. B. Brush, C. O. Qualset, Landraces of Maize in Central Mexico: An altitudinal transect. *Econ. Bot.* **57**, 7–20 (2003).
37. M. R. Bellon, J. E. Taylor, “Folk” soil taxonomy and the partial adoption of new seed varieties. *Econ. Dev. Cult. Change* **41**, 763–786 (1993).
38. J. A. Zarazua-Escobar, G. Almaguer-Vargas, J. G. Ocampo-Ledesma, The direct field support program (PROCAMPO) from and its impact on agricultural productive and commercial knowledge management in Estado de México. *Agric. Soc. Desarrollo* **8**, 89–105 (2011).
39. R. H. Perales, S. B. Brush, C. O. Qualset, Dynamic management of maize landraces in Central Mexico. *Econ. Bot.* **57**, 21–34 (2003).
40. C. U. Miramontes-Piña, *Situación actual y perspectivas del maíz en México 1996–2010* (SIAP Distrito Federal, México, 2006).
41. M. R. Bellon, J. Berthaud, Transgenic maize and the evolution of landrace diversity in Mexico. The importance of farmers’ behavior. *Plant Physiol.* **134**, 883–888 (2004).
42. Y. Matsuoka *et al.*, A single domestication for maize shown by multilocus microsatellite genotyping. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 6080–6084 (2002).
43. D. R. Piperno, A. J. Ranere, I. Holst, J. Iriarte, R. Dickau, Starch grain and phytolith evidence for early ninth millennium B.P. maize from the Central Balsas River Valley, Mexico. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 5019–5024 (2009).
44. A. J. Ranere, D. R. Piperno, I. Holst, R. Dickau, J. Iriarte, The cultural and chronological context of early Holocene maize and squash domestication in the Central Balsas River Valley, Mexico. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 5014–5018 (2009).
45. J. J. Sánchez González *et al.*, Ecogeography of teosinte. *PLoS One* **13**, e0192676 (2018).
46. E. Anderson, H. C. Cutler, Races of Zea Mays: I. Their recognition and classification. *Ann. Mo. Bot. Gard.* **29**, 69–88 (1942).
47. R. J. Elshire *et al.*, A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One* **6**, e19379 (2011).
48. CONABIO, Proyecto global de maíces nativos (Comisión Nacional Para el Conocimiento y Uso de la Biodiversidad, 2011). <https://www.biodiversidad.gob.mx/genes/proyectoMaices.html>. Accessed 23 April 2018.
49. K. Fukunaga *et al.*, Genetic diversity and population structure of teosinte. *Genetics* **169**, 2241–2254 (2005).
50. K. Herten, M. S. Hestand, J. R. Vermeesch, J. K. J. Van Houdt, GBSX: A toolkit for experimental design and demultiplexing genotyping by sequencing experiments. *BMC Bioinformatics* **16**, 73 (2015).
51. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: A flexible trimmer for Illumina sequencing data. *Bioinformatics* **30**, 2114–2120 (2014).
52. F. J. Sedlazeck, P. Rescheneder, A. von Haeseler, NextGenMap: Fast and accurate read mapping in highly polymorphic genomes. *Bioinformatics* **29**, 2790–2791 (2013).
53. H. Li *et al.*, 1000 Genome Project Data Processing Subgroup, The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
54. A. McKenna *et al.*, The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
55. P. Danecek *et al.*, 1000 Genomes Project Analysis Group, The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).
56. I. C. Rojas-Barrera *et al.*, Contemporary evolution of maize landraces and their wild relatives influenced by gene flow with modern maize varieties. Open Science Framework (OSF). <https://osf.io/pqvt4/>. Deposited 9 September 2019.
57. D. H. Alexander, J. Novembre, K. Lange, Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–1664 (2009).
58. T. S. Korneliusson, A. Albrechtsen, R. Nielsen, ANGSD: Analysis of next generation sequencing data. *BMC Bioinformatics* **15**, 356 (2014).
59. E. Y. Durand, N. Patterson, D. Reich, M. Slatkin, Testing for ancient admixture between closely related populations. *Mol. Biol. Evol.* **28**, 2239–2252 (2011).
60. R. J. Kulathinal, L. S. Steivon, M. A. F. Noor, The genomics of speciation in *Drosophila*: Diversity, divergence, and introgression estimated using low-coverage genome sequencing. *PLoS Genet.* **5**, e1000550 (2009).
61. R. E. Green *et al.*, A draft sequence of the Neandertal genome. *Science* **328**, 710–722 (2010).
62. J. Yan *et al.*, Genetic characterization and linkage disequilibrium estimation of a global maize collection using SNP markers. *PLoS One* **4**, e8451 (2009).
63. H. Wickham, *ggplot2: Elegant Graphics for Data Analysis* (Springer, 2016).
64. P. J. Bradbury *et al.*, TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics* **23**, 2633–2635 (2007).
65. M. C. Romay *et al.*, Comprehensive genotyping of the USA national maize inbred seed bank. *Genome Biol.* **14**, R55 (2013).
66. X. Zheng *et al.*, A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* **28**, 3326–3328 (2012).
67. M. N. Price, P. S. Dehal, A. P. Arkin, FastTree: Computing large minimum evolution trees with profiles instead of a distance matrix. *Mol. Biol. Evol.* **26**, 1641–1650 (2009).
68. J. d. J. Sánchez González *et al.*, Distribución geográfica del teocinte (*Zea* spp.) en México y situación actual de las poblaciones (CONABIO, 2008). [https://www.biodiversidad.gob.mx/genes/pdf/proyecto/Anexo8\\_ResultadosProyectos/FZ002/Informe%20final/Teocinte/Informe-Fina%20Teocinte\\_FZ002.pdf](https://www.biodiversidad.gob.mx/genes/pdf/proyecto/Anexo8_ResultadosProyectos/FZ002/Informe%20final/Teocinte/Informe-Fina%20Teocinte_FZ002.pdf). Accessed 27 June 2018.



