



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
POSGRADO EN CIENCIAS BIOLÓGICAS
INSTITUTO DE ECOLOGÍA

**ESTRUCTURA GENÉTICA ESPACIAL ASOCIADA A LA
SELECCIÓN EN RAZAS DE MAÍZ DE MÉXICO**

TESIS

QUE PARA OPTAR POR EL GRADO DE:

MAESTRA EN CIENCIAS BIOLÓGICAS
BIOLOGÍA EVOLUTIVA

PRESENTA:

ALEJANDRA BREÑA OCHOA

TUTOR PRINCIPAL DE TESIS:

DR. DANIEL IGNACIO PIÑERO DALMAU
INSTITUTO DE ECOLOGÍA, UNAM

COMITÉ TUTOR:

DR. RAFAEL LIRA SAADE
FACULTAD DE ESTUDIOS SUPERIORES IZTACALA, UNAM

DR. LUIS ENRIQUE EGUIARTE FRUNS
INSTITUTO DE ECOLOGÍA, UNAM



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
POSGRADO EN CIENCIAS BIOLÓGICAS
INSTITUTO DE ECOLOGÍA

**ESTRUCTURA GENÉTICA ESPACIAL ASOCIADA A LA
SELECCIÓN EN RAZAS DE MAÍZ DE MÉXICO**

TESIS

QUE PARA OPTAR POR EL GRADO DE:

MAESTRA EN CIENCIAS BIOLÓGICAS
BIOLOGÍA EVOLUTIVA

PRESENTA:

ALEJANDRA BREÑA OCHOA

TUTOR PRINCIPAL DE TESIS:

DR. DANIEL IGNACIO PIÑERO DALMAU
INSTITUTO DE ECOLOGÍA, UNAM

COMITÉ TUTOR:

DR. RAFAEL LIRA SAADE
FACULTAD DE ESTUDIOS SUPERIORES IZTACALA, UNAM

DR. LUIS ENRIQUE EGUIARTE FRUNS
INSTITUTO DE ECOLOGÍA, UNAM



Dr. Isidro Ávila Martínez
Director General de Administración Escolar, UNAM
Presente

Me permito informar a usted, que el Subcomité de Biología Evolutiva y Sistemática, en su sesión ordinaria del día 10 de junio de 2013, aprobó el jurado para la presentación de su examen para obtener el grado de **MAESTRA EN CIENCIAS BIOLÓGICAS** del Posgrado en Ciencias Biológicas, de la alumna **BREÑA OCHOA ALEJANDRA** con número de cuenta 405053663 con la tesis titulada **"ESTRUCTURA GENÉTICA ESPACIAL ASOCIADA A LA SELECCIÓN EN RAZAS DE MAÍZ DE MÉXICO"**, bajo la dirección del: DR. DANIEL IGNACIO PIÑERO DALMAU

Presidente: DR. JUAN PABLO JARAMILLO CORREA
Vocal: DRA. ANA LAURA WEGIER BRIUOLO
Secretario: DR. RAFAEL LIRA SAADE
Suplente: DRA. MARÍA FRANCISCA JOSÉ ACEVEDO GASMAN
Suplente: DR. LUIS ENRIQUE EGUIARTE FRUNS

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

ATENTAMENTE
"POR MI RAZA HABLARA EL ESPIRITU"
Cd. Universitaria, D.F., a 13 de septiembre de 2013.

DRA. MARÍA DEL CORO ARIZMENDI ARRIAGA
COORDINADORA DEL PROGRAMA

c.c.p. Expediente de la interesada.

AGRADECIMIENTOS ACADÉMICOS

- ❖ Al Posgrado en Ciencias Biológicas de la Universidad Nacional Autónoma de México (UNAM) por aceptarme en el programa de Maestría en Ciencias Biológicas (Biología Evolutiva) que encaminó este estudio.
- ❖ Al Concejo Nacional de Ciencia y Tecnología por otorgarme una beca (CONACYT 263533), dentro del Programa Nacional de Posgrados de Calidad (PNPC), que proporcionó una manutención mensual durante los dos años de duración de este estudio.
- ❖ A la Coordinación de Estudios de Posgrado de la UNAM por otorgar un apoyo económico, dentro del Programa de Apoyo a los Estudios de Posgrado (PAEP), para presentar este estudio en un congreso nacional.
- ❖ A la Dirección General del Sector Primario y Recursos Naturales Renovables (DGSPRNR) de la Secretaría de Medio Ambiente y Recursos Naturales (SEMARNAT) por otorgar el financiamiento al proyecto denominado "Monitoreo de las razas criollas y de los linajes geográficos de maíz en México usando un enfoque genómico: una estrategia para su uso sustentable y su conservación".
- ❖ A la Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO) por el otorgamiento de un apoyo económico, por parte de Nacional Financiera Fideicomiso "Fondo para la Biodiversidad", que impulsó el inicio de este estudio.
- ❖ Al comité tutorial, integrado por el Dr. Daniel Piñero, el Dr. Rafael Lira y el Dr. Luis Eguiarte, por su evaluación semestral que dirigió este estudio.

AGRADECIMIENTOS PERSONALES

- ❖ Al Dr. Daniel Piñero por darme la oportunidad de estudiar al maíz, guiar mis estudios y brindar consejos que han contribuido a mi crecimiento personal y profesional.
- ❖ A la Dra. Maria Clara Arteaga por organizar la siembra de las semillas, la extracción de ADN y los envíos al INMEGEN de los cientos de muestras de maíz.
- ❖ Al Dr. Rafael Lira y al Dr. Luis Eguiarte por aceptar formar parte de mi comité tutorial y por leer y escuchar cada semestre los avances en este estudio y brindar sus comentarios al respecto.
- ❖ Al Dr. Juan Pablo Jaramillo, a la Dra. Ana Wegier, al Dr. Rafael Lira, a la Dra. Francisca Acevedo y al Dr. Luis Eguiarte por aceptar formar parte de mi jurado de examen y por la evaluación, la revisión, los comentarios y la información que mejoraron esta tesis.
- ❖ A la Dra. Francisca Acevedo y a los integrantes de la Coordinación de Análisis de Riesgo y Bioseguridad de la CONABIO por las facilidades y el apoyo brindado durante este estudio. En particular, al Biól. Oswaldo Oliveros Galindo por la ayuda proporcionada en el diseño de varios mapas en esta tesis.
- ❖ A la M. en C. Fabiola Morales Mandujano y a los integrantes de la Unidad de Genotipificación y Análisis de Expresión Illumina del Instituto Nacional de Medicina Genómica (INMEGEN) por el servicio de genotipificación de los polimorfismos de nucleótido (SNPs).
- ❖ A la Unidad de Recursos Genéticos del Banco de Germoplasma del Centro Experimental Valle de México del Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias (INIFAP) por proporcionar las semillas de maíz para este estudio.
- ❖ Al Instituto de Ecología de la UNAM, particularmente al Departamento de Ecología Evolutiva, por ser la sede de este estudio.
- ❖ A los académicos (aparte de los ya mencionados) que contribuyeron a la culminación de esta tesis. Al Dr. Hugo Rafael Perales Rivera, del Colegio de la Frontera Sur (ECOSUR), por aceptar revisar esta tesis, brindar comentarios al respecto y proporcionar información que la enriqueció. Al Dr. José de Jesús Sánchez González, de la Universidad de Guadalajara, por recibir una versión de esta tesis y brindar comentarios al respecto. Al Dr. Javier Orlando Mijangos Cortés, del Centro

de Investigación Científica de Yucatán (CICY), por aceptar ser guía en una posible visita a una feria de semillas y proporcionar información acerca de las razas cultivadas en la península de Yucatán. Al Dr. Amalio Santacruz Varela y al Dr. Abel Muñoz Orozco, ambos del Colegio de Postgraduados (ColPos), por responder preguntas sobre las razas y proporcionar libros acerca del maíz en México. Al Dr. Andrés Moreno Estrada, de la Universidad de Stanford, por indicar los comandos de los filtros a los que se someten los SNPs.

- ❖ Al Biól. Iván Hernández Baltazar, integrante del Centro de Estudios para el Cambio en el Campo Mexicano (CECCAM), por la plática acerca de los transgénicos y por proporcionar folletos, libros y mapas sobre el maíz en México.
- ❖ A la Dra. Alejandra Vázquez Lobo Yurén por proporcionar el protocolo de extracción de ADN que se utilizó en este estudio.
- ❖ A mis compañeros de laboratorio del Departamento de Ecología Evolutiva. En particular, a la Biól. Laura Figueroa, al Biól. Marco Suárez, a la M. en C. Tania Garrido, al M. en C. Enrique Scheinvar y al Dr. Lev Jardón por responder a mis preguntas y ayudarme en varios análisis.
- ❖ Al Biol. Gabriel Merino Díaz por compartir a los integrantes del comité tutorial y participar en la organización de los tutorales semestrales.

A Leonardo, mi papá, a Emma, mi mamá y a Rodrigo, mi hermano,
por su amor, sus consejos y su apoyo incondicional
ayer, hoy y mañana

ÍNDICE

RESUMEN	1
ABSTRACT	2
INTRODUCCIÓN	3
OBJETIVOS	8
METODOLOGÍA	
Siembra de semillas	9
Extracción de ADN.....	9
Obtención de SNPs	10
Bases de datos de isoenzimas y microsatélites	11
Identificación de SNPs candidatos a la adaptación local	15
Extremos de latitud y altitud	16
Distribuciones geográficas restringidas y extensas	17
Identificación de SNPs candidatos a la selección natural positiva	18
Localización en el genoma de SNPs candidatos	19
Análisis de diversidad y estructura genética espacial	20
RESULTADOS	
SNPs candidatos a la adaptación local	23
Extremos de latitud y altitud	23
Distribuciones geográficas restringidas y extensas	24
SNPs candidatos a la selección natural positiva y su localización en el genoma.....	25
Diversidad genética.....	28
Estructura genética espacial	31
Localización geográfica (altitud, latitud y longitud).....	31
Distancia geográfica	33
Estructura espacial a escala fina.....	34
Estructura espacial a escala regional	36
DISCUSIÓN	
Estructura genética espacial	40
Genes candidatos a la selección natural positiva.....	44

Adaptación local	45
CONCLUSIONES	49
LITERATURA CITADA	51

ANEXOS

I. Protocolo CTAB.....	i
II. Localidades representadas con marcadores	
A. SNPs.....	iii
B. Isoenzimas y microsatélites	xv
III. Distribución geográfica de individuos en extremos de latitud y altitud	xxix
IV. Localización de SNPs en el genoma.....	xxx
V. Análisis con isoenzimas y microsatélites	
A. Diversidad genética	xxxvi
B. AMOVA y StAMOVA con isoenzimas.....	xxxviii
C. AMOVA y StAMOVA con microsatélites	xxxix
D. Prueba de Mantel con isoenzimas	xl
E. Prueba de Mantel con microsatélites	xl
VI. Análisis espacial de componentes principales	
A. Gráfica de barras de los valores eigen de SNPs con $F_{ST} \geq 0.15$	xli
B. Gráfica de barras de los valores eigen de SNPs con $F_{ST} \leq 0.02$	xli
C. Gráficas de los valores eigen en función de la varianza y la autocorrelación.....	xlii

MATERIAL SUPLEMENTARIO

Análisis de isoenzimas con un modelo bayesiano de asignación.....	A
Distribución del índice F_{ST} en isoenzimas y microsatélites	D

CUADROS Y FIGURAS

CUADRO

1. Lista de isoenzimas	12
2. Diferencias de latitud y altitud y la hipotenusa	23
3. Diversidad genética con SNPs.....	29
4. AMOVA y StAMOVA de SNPs con $F_{ST} \geq 0.15$	32
5. AMOVA y StAMOVA de SNPs con $F_{ST} \leq 0.02$	32

FIGURA

1. Distribución geográfica de los individuos con SNPs.....	13
2. Distribución geográfica de los individuos con isoenzimas y microsatélites	14
3. Mapa de desequilibrio de ligamiento	27
4. Prueba de Mantel con SNPs	33
5. Correlogramas.....	35
6. Contribución de los alelos al sPCA.....	38
7. Distribución geográfica de los puntajes regionales del primer componente espacial.....	39
8. Correlograma del tamaño del genoma (Díez <i>et al.</i> 2013)	42
9. Regionalización biogeográfica del maíz (Perales y Golicher)	44
S1. Gráficas del logaritmo de verosimilitud ($L(K)$) y de ΔK	B
S2. Gráfica del coeficiente de pertenencia (Q).....	C
S3. Distribución del índice F_{ST} de isoenzimas	E
S4. Distribución del índice F_{ST} de microsatélites.....	E

RESUMEN

El maíz (*Zea mays* subespecie *mays*) probablemente se domesticó hace alrededor de 9,000 años en la región central de la cuenca del Balsas, en las entidades federativas de Michoacán y Guerrero, a partir del teocintle anual *Zea mays* ssp. *parviglumis*. La dispersión de este cultivo y las técnicas regionales de agricultura tradicional posiblemente han promovido la diversificación en razas y la adaptación local. La selección divergente, una forma de selección positiva que genera la adaptación local, posiblemente ha incrementado la diferenciación genética de loci específicos en el genoma del maíz. Este genoma tiene un tamaño de alrededor de 2.3 gigapares de bases, las cuales están repartidas en alrededor de 10 cromosomas. La identificación de los polimorfismos de nucleótido (SNPs) candidatos a la selección positiva y la determinación de su efecto en la diversidad y estructura genética espacial del maíz son los objetivos de este estudio. Los SNPs con alta diferenciación genética se buscaron entre 41,299 marcadores de 169 individuos de 50 razas de la mayor parte de México. Los SNPs (153) con alta diferenciación ($F_{ST} \geq 0.15$) entre provincias biogeográficas se localizan mayormente en genes y las proteínas que sintetizan se han caracterizado o tienen ortólogos sinténicos con el sorgo (*Sorghum bicolor*). Es posible que algunos de estos genes sean candidatos a la selección positiva o su variación genética esté influida por barridos selectivos. A escala regional, las frecuencias alélicas de estos SNPs se distribuyen en parches que en mayor medida corresponden a las provincias biogeográficas. La interacción entre la latitud y la altitud representa casi la mitad (11.4%) del porcentaje de varianza explicado por estas provincias (23%). Además, la similitud de los genotipos individuales vecinos hasta distancias de alrededor de 300 km, evaluados con SNPs (200) posiblemente neutrales ($F_{ST} \leq 0.02$), sugiere la presencia de un patrón de aislamiento por distancia. Estos resultados concuerdan con los obtenidos con isoenzimas y microsatélites. Las presiones de selección generadas por factores abióticos, bióticos externos y culturales, posiblemente varían entre regiones. Algunos de estos factores que probablemente promueven la adaptación local en el maíz podrían ser la temperatura y la precipitación, el tipo de suelo y la presencia o ausencia de ciertos insectos y patógenos.

Palabras Clave: adaptación local, diferenciación genética, maíz, México, SNPs.

ABSTRACT

Maize (*Zea mays* subspecies *mays*) was probably domesticated around 9,000 years ago in the Central Balsas River Basin (in the states of Michoacán and Guerrero) from the annual teosinte *Zea mays* ssp. *parviglumis*. Maize dispersion and regional techniques of traditional agriculture have probably promoted landrace diversification and local adaptation. Divergent selection, a type of positive selection that generates local adaptation, has probably increased genetic differentiation in specific loci of the maize genome. This genome has a size of around 2.3 gigabase pairs that are distributed in around 10 chromosomes. The objectives of this study include the identification of single nucleotide polymorphisms (SNPs) that are positive selection candidates and the determination of their effect on maize diversity and spatial genetic structure. The search for SNPs with a large genetic differentiation was carried out on a panel of 41,299 markers from 169 individuals of 50 landraces from most of México. SNPs (153) with a large differentiation ($F_{ST} \geq 0.15$) between biogeographic provinces are mostly located in genes and the proteins they synthesize are characterized or have syntenic orthologs with sorghum (*Sorghum bicolor*). It is possible that some of these genes are positive selection candidates or their genetic variation is influenced by selective sweeps. At a regional scale, these SNPs allelic frequencies are distributed in patches that mostly coincide with the biogeographic provinces. Latitude and altitude interaction represents almost half (11.4%) of the variance percentage that is explained by these provinces (23%). A pattern of isolation by distance is suggested by similar individual genotypes until around 300 km that were evaluated with 200 probably neutral SNPs ($F_{ST} \leq 0.02$). These results coincide with those obtained with isozymes and microsatellites. Selective pressures generated by abiotic, external biotic and cultural factors possibly vary between regions. Some of these factors that probably promote local adaptation in maize could be temperature, precipitation, soil type and the presence or absence of certain insects and pathogens.

Key Words: local adaptation, genetic differentiation, maize, México, SNPs

INTRODUCCIÓN

México es el centro de origen y uno de los principales centros de diversificación del maíz (*Zea mays* subespecie *mays*; Kato *et al.* 2009). Este cultivo probablemente se domesticó hace alrededor de 9,000 años en la región central de la cuenca del Balsas, en las entidades federativas de Michoacán y Guerrero, a partir del teocintle *Zea mays* ssp. *parviglumis* (Matsuoka *et al.* 2002a; Piperno *et al.* 2009; Ranere *et al.* 2009; van Heerwaarden *et al.* 2011). Es posible que el maíz se dispersara rápido a otras regiones del país, debido a que se han encontrado polen y fitolitos (restos de origen vegetal mineralizados) antiguos en la costa del Golfo de México (7,300 años; Pohl *et al.* 2007), en el valle de Tehuantepec y en las tierras altas de Oaxaca (6,300 años; Piperno y Flannery 2001). La dispersión en poco tiempo de este cultivo fue probablemente promovida por el uso de las cuencas de los ríos, como la del complejo hidrológico Tepalcatepec-Balsas-Mexcala, como corredores biológicos y culturales para los grupos humanos durante el Holoceno Temprano (que comenzó hace 11,700 años; Zizumbo-Villarreal y Colunga-GarcíaMarín 2010). En la actualidad, el maíz es cultivado en la mayor parte de México y es central en la alimentación, la cultura y la economía de su sociedad (Kato *et al.* 2009).

La domesticación es un proceso evolutivo continuo, mediado por la selección humana, que en las plantas genera fenotipos favorables para el consumo humano y la agricultura pero disminuye o elimina la capacidad de éstas para sobrevivir en condiciones naturales, volviéndolas dependientes del hombre (Harlan 1992). La pérdida de diversidad genética del cultivo en comparación a su pariente silvestre, una consecuencia de este proceso, es el resultado de la selección artificial y del cuello de botella genético (Tenailon *et al.* 2004).

La selección artificial ha reducido la diversidad en loci (lugares) específicos del genoma del maíz como aquellos en donde se localizan genes que regulan los caracteres fenotípicos distintivos, la calidad nutricional, la madurez y la productividad (Yamasaki *et al.* 2005; Wright *et al.* 2005). Algunos de los caracteres fenotípicos que distinguen a este cultivo de los teocintles, sus parientes silvestres, y que forman parte del síndrome de domesticación (Hancock 2004) son el tamaño, la morfología y el número tanto de granos en cada mazorca como de mazorcas en cada planta, la longitud de las ramificaciones y el número y localización de las espigas (Kato *et al.* 2009; Zizumbo-Villarreal y Colunga-GarcíaMarín 2010). En cambio, el cuello de botella

genético disminuyó la diversidad en todo el genoma del maíz en comparación a los teocintles y esto probablemente fue consecuencia de que este cultivo se domesticara a partir de semillas que eran muestreadas y seleccionadas a criterio del hombre (Tenailon *et al.* 2004).

El genoma del maíz tiene un tamaño de alrededor de 2.3 gigapares de bases repartidas en 10 cromosomas que son distintos estructuralmente (Schnable *et al.* 2009). Casi el 85% de su extensión la constituyen elementos transponibles, ya que las regiones eucromáticas se caracterizan por retrotransposones que usualmente se distribuyen de manera anidada (SanMiguel *et al.* 1996; Schnable *et al.* 2009). Además, alrededor de un 8% está constituido por nudos cromosómicos (aglomeraciones de heterocromatina), que incluyen secuencias satelitales grandes (de varias megabases) alternadas por retrotransposones (Ananiev *et al.* 1998). Este genoma presenta loci con variación estructural frecuente, como aquellos donde se localizan genes candidatos a las respuestas al estrés y a los estímulos, y loci con poca variación o inexistente, como ocurre en bloques de más de 10 megapares de bases que no han tenido recombinación reciente (Chia *et al.* 2012). Los loci candidatos a la domesticación cubren alrededor del 7.6% del genoma, tienen un tamaño promedio de 322 kilopares de bases e incluyen un promedio de 3.4 genes (Hufford *et al.* 2012a). En este estudio publicado, se concluyó que la domesticación probablemente ha influido en cientos de genes con diversas funciones biológicas que posiblemente han afectado aspectos aún no estudiados del fenotipo.

La dispersión del maíz fuera de la cuenca del Balsas contribuyó a la evolución de razas; éstas están integradas por poblaciones nativas distinguibles por un conjunto heredable y común de caracteres fenotípicos y agronómicos (Hernández y Alanís 1970). Sin embargo, en las comunidades de México, los individuos suelen mostrar variación continua en caracteres cuantitativos como la altura de la planta, las dimensiones de la mazorca, el número de hileras de granos y el tiempo de floración, por lo que cada población nativa suele estar integrada por variedades (Ortega Paczka 2003; Vielle-Calzada y Padilla 2009). Esta variación fenotípica continua probablemente es en parte el resultado de que el maíz sea una planta de polinización abierta. El polen se dispersa por viento en un área de alrededor de 100 a 1,000 m, en cambio, las semillas permanecen unidas al olote y no se dispersan de manera natural (Kato *et al.* 2009). La rigurosa selección que los agricultores hacen de sus semillas posiblemente regula el intercambio genético continuo entre las variedades y las poblaciones nativas de otras razas, lo cual permite

que las primeras (las poblaciones) mantengan las características fenotípicas que permiten distinguirlas de las segundas (otras razas; Ortega Paczka 2003; Perales *et al.* 2003b).

Los sistemas mexicanos de agricultura tradicional del maíz tienen una distribución espacial local (Brush y Perales 2007; Bellon *et al.* 2011). Los intercambios de semillas ocurren con frecuencia entre agricultores de una comunidad o de redes sociales asociadas que se localizan en un radio de alrededor de 10 km y una altitud de más o menos 50 metros sobre el nivel del mar (msnm; Bellon *et al.* 2011). El manejo en las comunidades se caracteriza por la persistencia de las razas a pesar de la presencia de líneas mejoradas endogámicas, la dominancia relativa de una variedad, el cultivo de otras variedades en menor grado, la sustitución frecuente entre las variedades que se usan para las tortillas, la selección de las semillas basada en un ideotipo local y el mantenimiento de variedades en regiones de mayor altitud (Louette *et al.* 1997; Pressoir y Berthaud 2004; Brush y Perales 2007). Las razas de maíz, en este sistema local de manejo, posiblemente están sujetas a presiones de selección consistente y estable no sólo por parte de los agricultores sino también por parte de los ambientes en los que se cultivan, lo cual genera la adaptación local (Bellon *et al.* 2011).

La adaptación local en un conjunto de poblaciones conectadas por flujo génico probablemente es el resultado de selección natural en proceso o muy reciente que varía espacialmente, ya que se debe a las distintas condiciones ambientales que experimenta cada población (Kawecki y Ebert 2004). Esta selección divergente (especialmente heterogénea) genera la evolución de caracteres que en conjunto integran genotipos locales y éstos tienen, en promedio, una adecuación relativa más alta en su ambiente local que los genotipos originados en otros ambientes (Williams 1966; Kawecki y Ebert 2004).

Las razas de maíz muestran evidencias de distintos niveles de adaptación local a la elevación de origen (Perales *et al.* 2003a; Mercer *et al.* 2008). En México, este cultivo se siembra de cero a 2,900 msnm, en regiones con una temperatura media anual de 11.3 a 26.6°C y una precipitación media anual de 426 a 4,245 mm (Ruiz *et al.* 2008). Sin embargo, la mayor diversidad de razas se localiza a menos de 2,000 msnm, en regiones con temperaturas de subtropicales a templadas (18 a 26°C) y con un régimen de precipitación anual de subhúmedo a semiárido (menos de 580 mm; Brush y Perales 2007; Ruiz *et al.* 2008). Entre las razas adaptadas a los ambientes más fríos (menos de 18°C en la temporada de crecimiento), con heladas más

tempranas y a más de 2,000 msnm, se encuentran las integrantes de los complejos Cónico y Sierra de Chihuahua (Sánchez y Goodman 1992; Perales *et al.* 2003a). La adaptación local a estos ambientes posiblemente fue promovida por la hibridación e introgresión del teocintle *Zea mays ssp. mexicana* al maíz, ya que alrededor del 20% del genoma de estas razas se deriva de este pariente silvestre (van Heerwaarden *et al.* 2011; Hufford *et al.* 2013). Además, la introgresión de este teocintle al cultivo parece estar restringida en los loci del genoma candidatos a la domesticación pero es frecuente en loci candidatos a la adaptación a las tierras altas (Hufford *et al.* 2013). Es posible que este proceso de adaptación involucre cambios fisiológicos, fenológicos y morfológicos costosos, debido a que la adecuación de estas razas disminuye a menor altitud (Mercer *et al.* 2008).

La identificación de loci en el genoma del maíz candidatos a la adaptación local permitiría conocer la magnitud del efecto de los distintos ambientes en la evolución de las razas y contribuiría a plantear estrategias de conservación en México (Brush y Perales 2007; Mercer *et al.* 2008; Landguth y Balkenhol 2012). El genoma de referencia de este cultivo obtenido a partir de la línea pura endogámica B73 permite la búsqueda de evidencias de selección natural sin el conocimiento previo de los determinantes genéticos de los caracteres de interés (Schnable *et al.* 2009; Siol *et al.* 2010). Esta búsqueda podría llevarse a cabo con los polimorfismos de nucleótido (SNPs), ya que son marcadores abundantes en el genoma, su variación entre los individuos suele estar representada por dos de los cuatro nucleótidos posibles (bialélicos) y evolucionan de acuerdo a un modelo sencillo de mutación de sitios infinitos (Schlötterer 2004).

La diversidad y la estructura genética de las razas de maíz, sin embargo, se han caracterizado con otros marcadores, como las isoenzimas y los microsatélites (Doebley *et al.* 1985; Sánchez *et al.* 2000; Matsuoka *et al.* 2002a; Reif *et al.* 2006; Vigouroux *et al.* 2008). Las isoenzimas son proteínas que difieren en la secuencia de aminoácidos pero catalizan la misma reacción química y, debido a que representan genes, es posible que su variación pudiera estar asociada a la selección natural (Sánchez *et al.* 2000; Pyhäjärvi *et al.* 2011). Los microsatélites son secuencias constituidas por un patrón de uno a seis nucleótidos repetidos en tándem cinco o más veces y se caracterizan por ser altamente polimórficos y por estar usualmente influenciados por el flujo génico, la deriva génica y la mutación (fuerzas evolutivas neutrales; Matsuoka *et al.* 2002b). Por otro lado, la abundancia de los SNPs proporciona una muestra representativa del genoma, reduce la varianza de muestreo entre loci y permite identificar loci candidatos a la selección

(Morin *et al.* 2004). En consecuencia, los estudios del genoma del maíz con los SNPs han permitido identificar loci candidatos a la introgresión con *Z. m. ssp. mexicana*, a la domesticación y al mejoramiento (Hufford *et al.* 2012a, 2013).

Los métodos moleculares más antiguos para la detección del efecto de la selección natural en el genoma se basan en el índice de fijación de Wright (F_{ST} ; Wright 1951; Lewontin y Krakauer 1973). Este parámetro se define como la probabilidad de que dos genes en una subpoblación compartan un ancestro común de esa subpoblación (Wright 1951; Beaumont 2005). El principal supuesto de estos métodos es que si la selección natural ocurre en un locus (lugar) específico del genoma, la diferenciación, evaluada con el índice F_{ST} , será mayor en este locus en comparación a lo que se espera teóricamente en los loci neutrales (Lewontin y Krakauer 1973). Sin embargo, los cambios esperados en los patrones de diferenciación de las frecuencias alélicas debidos a la selección natural incluyen tanto el incremento como el decremento en la diferenciación entre subpoblaciones en loci específicos del genoma (Nielsen 2005).

La principal fuerza evolutiva que impulsa la adaptación local es la selección divergente (Kawecki y Ebert 2004). Este tipo de selección se considera positiva ya que se caracteriza por incrementar la frecuencia de nuevas mutaciones ventajosas en las subpoblaciones (Nielsen 2005). Además, la base genética de la adaptación local podría explicarse con dos formas de interacción entre el genotipo y el ambiente (Anderson *et al.* 2012). La pleiotropía antagónica se caracteriza por el efecto contrario en la adecuación que generan los alelos de un locus del genoma en ambientes distintos (Kawecki y Ebert 2004). En cambio, la neutralidad condicional se identifica por la ventaja adaptativa en un ambiente y la neutralidad en otro distinto que muestra un alelo (Anderson *et al.* 2012). Por lo tanto, los métodos basados en el índice F_{ST} podrían detectar el efecto de la adaptación local en el genoma (Lewontin y Krakauer 1973).

La distinción entre los patrones de adaptación local y de otros tipos de selección u otras fuerzas evolutivas requiere del análisis de representantes de distintas regiones geográficas (Kawecki y Ebert 2004; Morrell *et al.* 2012). En el maíz cultivado en México, las evidencias de adaptación local se han encontrado entre las razas de ciertas regiones (Perales *et al.* 2003a; Mercer *et al.* 2008). En este estudio, se buscan SNPs candidatos a la adaptación local en el genoma de individuos de razas cultivadas a lo largo del territorio nacional para determinar el efecto de la selección natural en la diversidad y estructura genética espacial del maíz.

OBJETIVOS

1. Identificar SNPs candidatos a la adaptación local a lo largo del genoma del maíz cultivado en México.
 - a) Usar el índice de fijación poblacional (F_{ST}) para identificar SNPs con alta diferenciación genética entre individuos de una misma raza en distintas regiones geográficas.
 - b) Establecer si estos marcadores con alta diferenciación genética se comparten entre razas.
2. Determinar el efecto de los SNPs candidatos a la selección natural (positiva) en la diversidad y la estructura genética espacial del maíz en México.
 - a) Comparar la diversidad genética de SNPs candidatos a la selección positiva y de marcadores posiblemente neutrales.
 - b) Identificar patrones de estructura genética con ambos conjuntos de SNPs.
 - c) Comparar los resultados obtenidos con los SNPs con los obtenidos previamente con isoenzimas y microsatélites.

METODOLOGÍA

Siembra de semillas

Las semillas para este estudio fueron proporcionadas por la Unidad de Recursos Genéticos del Banco de Germoplasma del Centro Experimental Valle de México del Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias (INIFAP). Estas semillas fueron solicitadas por la Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO) como parte del Proyecto Global de Maíces Nativos. En total se recibieron 209 muestras de 51 razas de maíz (1-19 muestras por raza); la mayoría (192) colectadas, entre 2006 y 2010, en 139 localidades de 19 entidades federativas en distintas regiones de México (CONABIO 2010). Cada muestra está integrada por 100 semillas de 20 mazorcas que fueron recolectadas en una misma milpa de una raza.

Las semillas se germinaron en el invernadero del Instituto de Ecología de la Universidad Nacional Autónoma de México (UNAM). Dos a cinco semillas de cada muestra se sembraron en tierra negra con vermiculita en proporción 3:1, respectivamente, dentro de almácigos de 50 pozos o en macetas de 500 g. Los almácigos se colocaron en una cámara de crecimiento Conviron, con una temperatura entre 28 y 32°C y una humedad entre 35 y 40%, durante semana y media o dos semanas con riego diario. Durante julio a noviembre del 2011, se sembraron un total de 190 muestras (1-17 muestras por raza) que representan a todas las razas proporcionadas de las 19 entidades federativas. Las semillas de siete de estas muestras, que incluyen a la única de la raza Negrito (colectada en 2009), no germinaron o produjeron plántulas endebles. Las plántulas se mantuvieron en el almácigo o la maceta hasta que alcanzaron una altura de alrededor de 12 cm. Un fragmento de 1/3 de hoja tierna se cortó de la plántula al momento de llevar a cabo la extracción de ácido desoxirribonucleico (ADN). El resto de la plántula se guardó en refrigeración a -70°C.

Extracción de ADN

La extracción de ADN genómico se hizo en el Departamento de Ecología Evolutiva del Instituto de Ecología de la UNAM siguiendo un protocolo modificado que usa el detergente no iónico

bromuro de cetilmetilamonio (CTAB; Anexo I; Doyle y Doyle 1987; Vázquez-Lobo 1996). Este procedimiento se repitió por separado con dos plántulas de cada muestra buscando la mejor calidad en el extracto de ADN. La purificación del extracto se llevó a cabo siguiendo el protocolo del kit de purificación de productos de PCR MinElute de Qiagen. El ADN se visualizó mediante electroforesis en gel de agarosa al 1% teñido con bromuro de etidio en un transiluminador de luz ultravioleta con cámara digital y se determinó su calidad mediante cuantificación en un espectrofotómetro Nanodrop. Si la relación 260/280 nm era 1.6 o mayor, entonces el ADN se consideraba de suficiente calidad para ser procesado. Los extractos de siete muestras no cumplieron con esta relación por lo que no fueron utilizados. El ADN extraído se guardó en refrigeración a -20°C y los extractos de buena calidad (183 muestras) se trasladaron en hieleras de unicel con hielo ($0-4^{\circ}\text{C}$).

Obtención de SNPs

La obtención de los SNPs se llevó a cabo en la Unidad de Genotipificación y Análisis de Expresión Illumina del Instituto Nacional de Medicina Genómica (INMEGEN) a cargo de la M. en C. Fabiola Morales Mandujano y la M. en C. Leticia Sebastián Medina. El protocolo que se siguió fue el del ensayo Infinium HD de Illumina que usa el MaizeSNP50 BeadChip para detectar los alelos de 56,110 SNPs para cada muestra.

Los miles de SNPs fueron seleccionados, durante el desarrollo del chip, tomando en cuenta que el desequilibrio de ligamiento en el genoma del maíz se extiende desde alrededor de 2,000 pares de bases hasta menos de 60 kilopares de bases (Remington *et al.* 2001; Ganai *et al.* 2011). Los SNPs seleccionados se localizan en más del 50% de los 17,520 genes de alta confiabilidad, localizados en los 10 cromosomas, del genoma de referencia de la línea mejorada endogámica B73 (Schnable *et al.* 2009; Ganai *et al.* 2011). En teoría, una distribución homogénea de 50,000 SNPs en el genoma de este cultivo da como resultado en promedio 40 a 50 kilopares de bases entre marcadores, sin embargo, la mayoría de los SNPs vecinos detectados por el chip varían en la cantidad de nucleótidos que los separan de menos de 40 a más de 100 kilopares de bases (Ganai *et al.* 2011). En este estudio publicado, se probó la funcionalidad del chip con 274 líneas mejoradas endogámicas, incluida B73, y se obtuvo que 49,585 de los 56,110 SNPs son informativos, ya que probablemente están separados por más de 20 pares de bases, distancia que

requieren los oligonucleótidos (primers) del ensayo de Illumina para funcionar, y no se localizan en genes parálogos (genes duplicados en loci distintos) en ninguna de estas líneas.

La identificación y el análisis de calidad de los 56,110 SNPs obtenidos de 183 muestras se hizo con el módulo de genotipificación (Genotyping (GT) Module v.1.7.4) del programa GenomeStudio v.2010.1 de Illumina (www.illumina.com/software/genomestudio_software.ilmn). El algoritmo GenCall compara la intensidad de la señal de los dos colores fluorescentes, que representan los dos alelos de un SNP, con agrupaciones de genotipos canónicos. Los genotipos obtenidos son representados gráficamente, de acuerdo a la frecuencia alélica (NormTheta; eje x) en relación a la intensidad de la señal (NormR; eje y), como puntos de color diferente dependiendo de si son homocigotos o heterocigotos. Aquellos que no pudieron ser identificados (CallRate menor a 0.95) y que no se agruparon en las áreas que representan a los genotipos canónicos (GenTrain menor a 0.5) fueron eliminados. En total, 2,009 SNPs no cumplieron con estos requisitos en las muestras analizadas.

Los 54,101 SNPs de 183 muestras pasaron por varios filtros en el programa PLINK v.1.07 (Purcell *et al.* 2007) que consistieron en la exclusión de los genotipos con más del 10% de datos faltantes en cada muestra (--mind) y en cada marcador (--geno) y en la eliminación de los SNPs con una frecuencia del alelo menos frecuente menor a 5% (--maf) y que no estuvieran en equilibrio de Hardy-Weinberg ($p < 0.000005$; --hwe). Cinco muestras fueron eliminadas por datos faltantes y nueve individuos, que representan la primera generación de cruza entre razas, no fueron incluidos en el resto de los análisis. En total, 12,802 SNPs no cumplieron con los requisitos establecidos en los filtros. Finalmente, se obtuvo una base de datos de 41,299 SNPs de 169 individuos de 50 razas de maíz (Figura 1; Anexo II.A).

Bases de datos de isoenzimas y microsatélites

Los genotipos de isoenzimas y microsatélites de individuos de maíz localizados en México se obtuvieron del portal de Panzea (www.panzea.org). Este portal inició como la herramienta bioinformática del proyecto de arquitectura genética del maíz y teocintle que financia la *National Science Foundation* (NSF) de Estados Unidos de América.

Los genotipos de isoenzimas incluyen a 34 marcadores (Cuadro 1) de 3,600 individuos de 259 poblaciones de 59 razas (Doebley *et al.* 1985; Sánchez *et al.* 2000). Doce o 24 individuos

integran a cada población y una a ocho poblaciones representan a cada raza. Los genotipos de microsatélites incluyen a 96 marcadores de 213 individuos de 58 razas (1-6 individuos por raza; Matsuoka *et al.* 2002a; Vigouroux *et al.* 2008). Las poblaciones en ambas bases de datos son prácticamente las mismas; de las 259 en la de isoenzimas, 59 no están representadas en la de microsatélites, y de las 201 en ésta última, dos no están en la primera (Figura 2; Anexo II.B).

Cuadro 1. Abreviatura y nombre de las 34 isoenzimas analizadas en 259 poblaciones de 59 razas.

Locus	Enzima
<i>ACO 1</i>	Aconitasa
<i>ACP 1</i>	Fosfatasa ácida
<i>ADH 1</i>	Alcohol deshidrogenasa
<i>AK</i>	Adenilato cinasa
<i>AMP 1, 3</i>	Arginina aminopeptidasa
<i>CAT 3</i>	Catalasa
<i>DIA 1, 2</i>	Diaforasa
<i>ENP 1</i>	Endopeptidasa
<i>EST 8</i>	Esterasa
<i>GLU 1</i>	β -Glucosidasa
<i>GOT 1, 2, 3</i>	Glutamato-oxaloacetato transaminasa
<i>HEX 2</i>	Hexocinasa
<i>IDH 1, 2</i>	Isocitrato deshidrogenasa
<i>MDH 1, 2, 3, 4, 5</i>	Malato deshidrogenasa
<i>PGD 1, 2</i>	6-Fosfogluconato deshidrogenasa
<i>PGM 1, 2</i>	Fosfoglucomutasa
<i>PHI 1</i>	Fosfohexosa isomerasa
<i>SAD</i>	Ácido shikimico deshidrogenasa
<i>TPI 1, 2, 3, 4, 5</i>	Triosa fosfato isomerasa

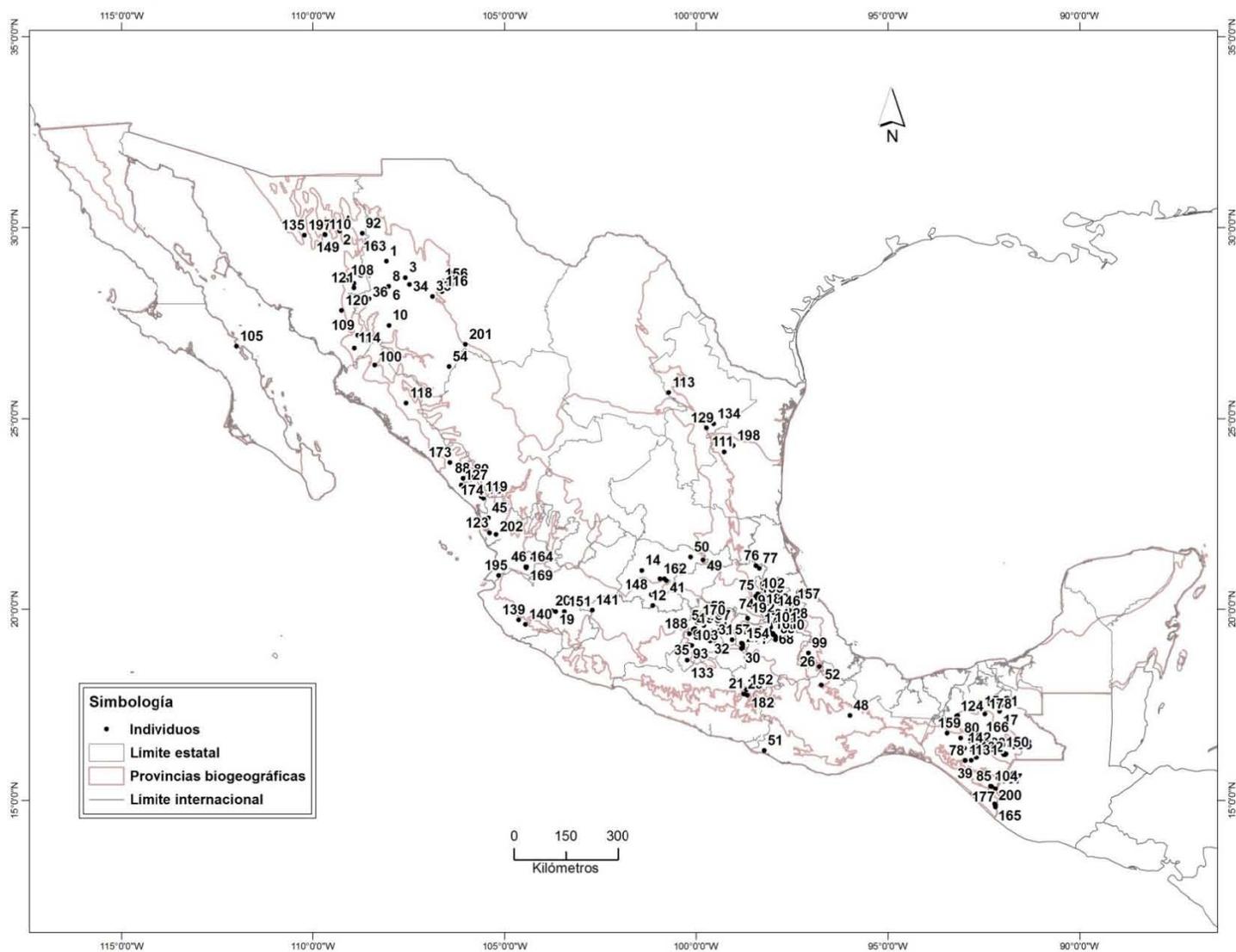


Figura 1. Distribución geográfica de los 169 individuos de 50 razas analizados con 41,299 SNPs. Los números corresponden a la columna Muestra del Anexo II.A. Diseño: Oswaldo Oliveros y Alejandra Breña Ochoa.

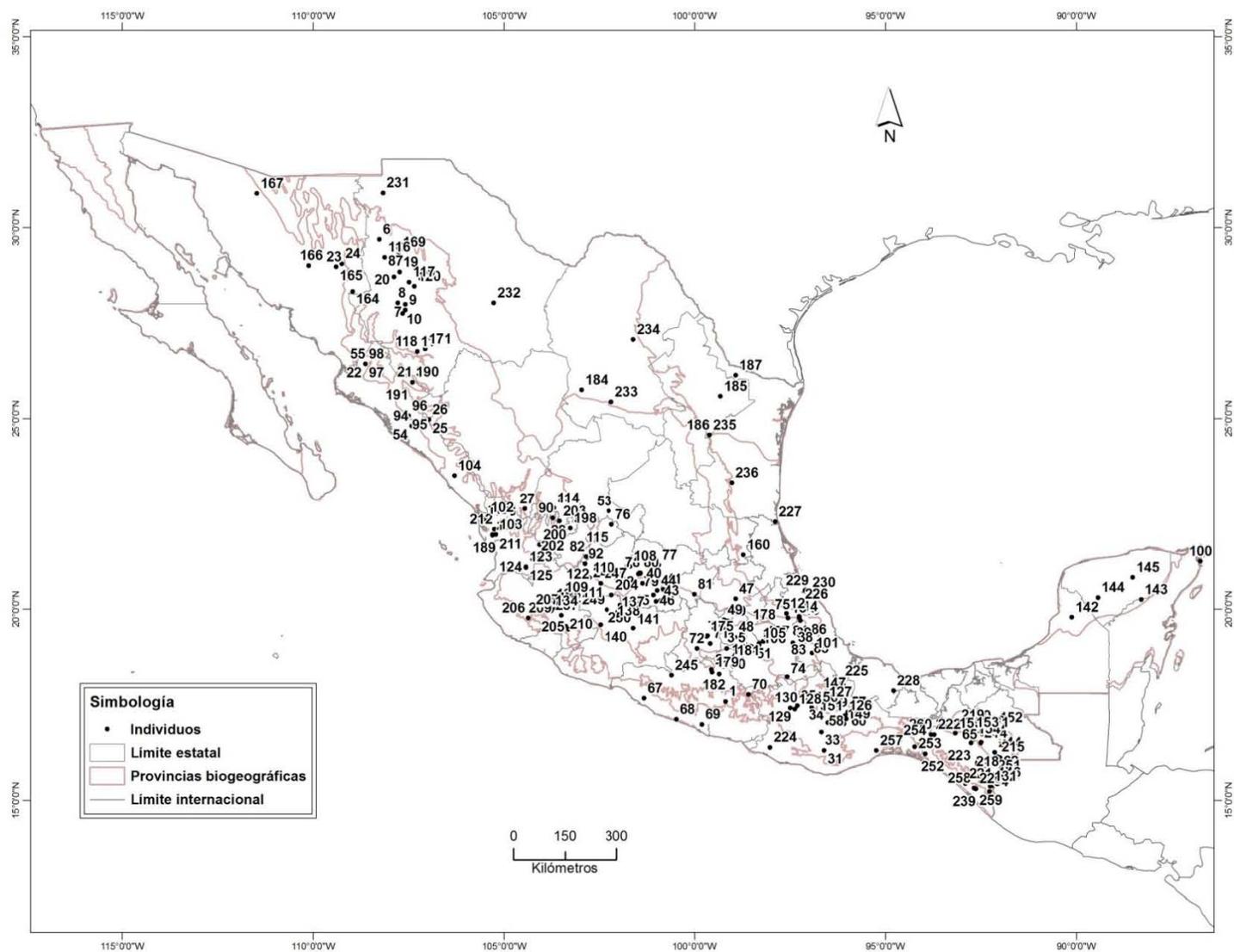


Figura 2. Distribución geográfica de las 261 poblaciones de 59 razas analizadas con 34 isoenzimas y 96 microsatélites. Los números corresponden al Anexo II.B. Diseño: Oswaldo Oliveros y Alejandra Breña Ochoa.

Identificación de SNPs candidatos a la adaptación local

Los métodos usados para identificar SNPs candidatos a la adaptación local se desarrollaron tomando en cuenta los resultados de estructura genética obtenidos con las isoenzimas y los microsatélites. La identificación de estructura genética permite determinar la posible distribución espacial de las frecuencias alélicas (Sjol *et al.* 2010). Entre los patrones que pueden presentar estas frecuencias se incluyen las agrupaciones, las clinas y los de aislamiento por distancia y éstos suelen ser identificados por los análisis que usan modelos bayesianos de asignación (François y Durand 2010). Los microsatélites ya habían sido previamente analizados con esta metodología (Vigouroux *et al.* 2008). En cambio, la relación de los genotipos individuales de las isoenzimas se evaluó con Structure v.2.3 (Pritchard *et al.* 2000). La proporción de pertenencia de estos genotipos a una de las dos agrupaciones sustentadas por la ΔK (Evanno *et al.* 2005) parece mostrar un patrón altitudinal (Figura S2), por lo que este factor podría determinar la distribución espacial de las frecuencias alélicas en el maíz. El análisis se describe con mayor detalle en el Material Suplementario.

El efecto de la localización geográfica en la estructura genética del maíz se evaluó mediante el análisis de varianza molecular con un modelo multivariado lineal general (StAMOVA; Dyer *et al.* 2004). Este análisis se describe con mayor detalle posteriormente (en la sección Análisis de diversidad y estructura genética espacial). Los resultados con los microsatélites muestran que la localización geográfica no tiene un efecto significativo en la repartición de la varianza genética, en contraste, los de las isoenzimas sugieren que este factor si influye en esta repartición (ver Resultados; Anexos V.B y V.C). Los microsatélites se localizan usualmente en loci del genoma donde la distribución de las frecuencias alélicas está influida por procesos neutrales, en cambio, las isoenzimas son proteínas cuya diferenciación podría deberse al efecto de la selección natural (ver Material Suplementario; Sánchez *et al.* 2000; Matsuoka *et al.* 2002b; Pyhäjärvi *et al.* 2011). En el caso de los SNPs con alta diferenciación genética, el StAMOVA muestra un efecto significativo de la interacción entre la latitud y la altitud en la repartición de la varianza genética, mientras que la longitud no parece tener este efecto (ver Resultados; Cuadro 4). Por consiguiente, con base en la literatura y los resultados de estructura genética obtenidos se desarrollaron los métodos que se describen a continuación para identificar SNPs candidatos a la adaptación local en el maíz cultivado en México.

Extremos de latitud y altitud

El primer método tomó en cuenta las diferencias de latitud y altitud de cada par de individuos de una misma raza. Éste consistió en seleccionar las razas, entre las 50 analizadas, que presentan pares de individuos con la máxima diferencia de altitud (de alrededor de 1,000 m). Después, se obtuvo la diferencia de latitud en metros para todas las combinaciones pareadas posibles de cada una de las razas seleccionadas. Los datos de latitud en metros se generaron localizando espacialmente a todos los individuos mediante Google Earth (www.google.com/earth/). El efecto de la longitud se eliminó igualándola entre los pares de individuos al momento de medir la distancia entre ellos. Para esto los individuos a comparar se localizaron en el mapa con la longitud del más norteño y con la latitud que le corresponde a cada uno.

La identificación de los pares de individuos en extremos de latitud y altitud se hizo obteniendo la hipotenusa del triángulo rectángulo cuyos catetos serían la diferencia de altitud y la diferencia de latitud. El valor de esta hipotenusa representa el efecto conjunto de la altitud y la latitud en la distancia espacial de los pares de individuos. La fórmula para el cálculo de la hipotenusa es (Baldor 2003):

$$c = \sqrt{(a^2 + b^2)}$$

Donde c es la hipotenusa, a , la diferencia de altitud (en metros) y b , la diferencia de latitud en metros. La hipotenusa se obtuvo para todas las combinaciones pareadas posibles de cada raza seleccionada y se escogió el par por raza que presentara el valor más alto (de más de 100 km). Estos pares de individuos seleccionados se localizan en extremos de latitud y altitud dentro de la muestra analizada por raza.

El índice F_{ST} entre los pares de individuos se obtuvo para cada uno de los 41,299 SNPs. Los genotipos de cada par se extrajeron de la base de datos de 169 individuos usando PLINK v.1.07 (Purcell *et al.* 2007). El cálculo del índice F_{ST} de cada SNP se llevó a cabo con una serie de comandos en R v.2.15.2 (R Development Core Team). Los marcadores que tienen un valor de F_{ST} igual a uno son los que cuentan con la máxima diferenciación (alelos diferentes a distinta latitud y altitud) entre los pares de individuos analizados. Estos SNPs obtenidos para cada raza se compararon entre sí buscando coincidencias. Los marcadores que coincidan entre razas podrían representar loci en el genoma del maíz candidatos a la adaptación a ambientes diferentes en extremos de latitud y altitud.

Distribuciones geográficas restringidas y extensas

El segundo método tomó en cuenta las distintas extensiones en las distribuciones geográficas de las razas. Aquellas que se encuentran en altitudes intermedias (entre 1,200 y 1,900 msnm) suelen presentar distribuciones geográficas más extensas que aquellas que se localizan a más de 2,000 msnm (Perales *et al.* 2003a).

El método consistió en seleccionar razas, entre las 50 analizadas, que presentan distribuciones geográficas contrastantes, es decir, restringidas o extensas. En primer lugar, se obtuvieron los mapas de distribución geográfica de cada una de las razas en el portal del Proyecto Global de Maíces Nativos de la CONABIO (<http://www.biodiversidad.gob.mx/genes/proyectoMaices.html>). Después, se seleccionaron las razas de acuerdo a su distribución geográfica y a la localización de los individuos de la muestra. El criterio para esto fue que las razas con distribuciones restringidas se localizan en dos entidades federativas y una provincia biogeográfica, mientras que aquellas con distribuciones extensas se encuentran en gran parte de México y los individuos que las representan están en un mínimo de tres entidades federativas y dos provincias biogeográficas.

Una vez que se seleccionaron las razas con distribuciones extensas y restringidas, los individuos de éstas fueron catalogados en dos grupos de acuerdo a la altitud y la latitud. Los límites que definieron a los grupos fueron 2,000 msnm para la altitud y 20°N para la latitud. Los individuos seleccionados fueron aquellos que se encontraran a una altitud y una latitud por debajo o por arriba de estos límites. Aquellos que no cumplieron con estos requisitos fueron eliminados.

El índice F_{ST} se obtuvo, para cada uno de los 41,299 SNPs, entre grupos de individuos definidos por su latitud y altitud dentro de una misma categoría de distribución geográfica. Los genotipos de los individuos seleccionados se extrajeron de la base de datos de 169 individuos usando PLINK v.1.07 (Purcell *et al.* 2007). El cálculo del índice F_{ST} de cada SNP se llevó a cabo con una serie de comandos en R v.2.15.2 (R Development Core Team). Los marcadores con un valor de F_{ST} alto (igual o mayor a 0.4) en cada categoría de distribución geográfica se compararon para identificar coincidencias. Los SNPs que coincidan entre distintas extensiones de distribución geográfica podrían representar loci en el genoma del maíz candidatos a la adaptación a ambientes a distintas latitudes y altitudes.

Identificación de SNPs candidatos a la selección natural positiva

La búsqueda de SNPs con alta F_{ST} compartidos entre agrupaciones de individuos de distintas regiones geográficas no mostró loci en el genoma candidatos a la adaptación a ambientes diferentes en extremos de latitud y altitud (ver Resultados). En consecuencia, los 169 individuos se agruparon de acuerdo a la provincia biogeográfica donde se localizan y se obtuvo el índice F_{ST} entre éstas, para cada uno de los 41,299 SNPs, con el objetivo de identificar el posible efecto de la selección natural (positiva) en la diversidad y estructura genética espacial del maíz.

Las 19 provincias biogeográficas que regionalizan a México se definen por criterios tanto biogeográficos como ecológicos (CONABIO 1997; Morrone 2006). Los patrones similares de distribución geográfica de distintas especies junto con los factores bióticos y abióticos y las características fisiográficas que los definen son la base de esta regionalización (Morrone 2006). En el maíz, la regionalización biogeográfica de las razas sugiere la presencia de complejos caracterizados por morfotipos similares que se distribuyen en áreas con interacciones similares entre el genotipo y el ambiente (Sánchez y Goodman 1992; Brush y Perales 2007). En cambio, los análisis de agrupación genética, como el de asignación bayesiana con las isoenzimas (Figura S2; Material Suplementario), sugieren que posiblemente este cultivo sea un continuo genético sin una clara estructuración a lo largo de su distribución geográfica. Sólo cuando el número de agrupaciones es similar al número de razas analizadas se obtienen grupos genéticos definidos (Reif *et al.* 2006). Por lo tanto, la identificación de SNPs candidatos a la selección natural (positiva) se llevó a cabo agrupando a los individuos por la provincia biogeográfica de procedencia.

Los 169 individuos de 50 razas que cuentan con datos de SNPs se hallan en 15 provincias biogeográficas (Figura 1; Anexo II.A). El número de individuos y de razas por provincia varía de un Mushito en Baja California a 44 individuos de 11 razas en la Faja Volcánica. Las provincias de la Costa del Pacífico y del Golfo de México se subdividieron en norte y sur, debido a que los individuos localizados en éstas se distribuyen en grupos a distintas latitudes (Anexo II.A). En contraste, los individuos de las 59 razas con datos de isoenzimas y microsatélites se localizan en 16 provincias biogeográficas (Figura 2; Anexo II.B). El número de individuos y de razas por provincia varía de 12 o dos individuos de Dzit-Bacal en el Petén a 840 o 49 individuos de 18 razas en la Faja Volcánica, de isoenzimas o microsatélites, respectivamente. Las provincias

biogeográficas con individuos son las mismas para todos los marcadores con excepción de Baja California, representada con los SNPs, y el Petén y Yucatán, representados con los microsatélites y las isoenzimas.

El índice F_{ST} entre provincias biogeográficas se obtuvo para cada uno de los 41,299 SNP buscando identificar un conjunto de marcadores con alta diferenciación genética que pudiera representar el efecto de la selección natural (positiva) en el genoma del maíz. Este estadístico se calculó con una serie de comandos en R v.2.15.2 (R Development Core Team). Los genotipos de los SNPs que tuvieron un valor de F_{ST} igual o mayor a 0.15 y de un conjunto aleatorio de marcadores con valores entre cero y 0.02 se extrajeron de la base de datos de 169 individuos usando PLINK v.1.07 (Purcell *et al.* 2007). Los SNPs con alta F_{ST} (≥ 0.15) entre provincias biogeográficas podrían localizarse en genes candidatos a la selección positiva o en loci influidos por esta fuerza evolutiva, mientras que los marcadores con valores bajos de este parámetro ($0 \leq F_{ST} \leq 0.02$) posiblemente son neutrales. Los análisis de diversidad y estructura genética espacial que se describen posteriormente se llevaron a cabo con estas dos bases de datos para comparar el posible efecto de la selección positiva con el de los procesos neutrales en la diversidad y estructura genética del maíz.

Localización en el genoma de SNPs candidatos

Los SNPs con índice F_{ST} alto identificados con los métodos descritos anteriormente se localizaron en los 10 cromosomas del genoma de referencia del maíz (el de la línea mejorada endogámica B73) buscando confirmar que estos marcadores son candidatos a la selección natural (positiva). El procedimiento para localizar cada SNP se describe a continuación.

El número de identificación (denominado #ss) de la secuencia de referencia, en la base de datos de polimorfismos de nucleótido (dbSNP), se obtuvo del estudio publicado acerca del MaizeSNP50 BeadChip (Ganal *et al.* 2011). Las secuencias de referencia se buscaron (*Batch Query*) en la dbSNP (www.ncbi.nlm.nih.gov/snp) del *National Center for Biotechnology Information* (NCBI) del *National Institute of Health* (NIH) de Estados Unidos de América. La localización en el genoma de referencia (Maize_RefGen_v2) de cada una de estas secuencias se obtuvo usando la *Basic Local Alignment Search Tool* (BLAST) para nucleótidos del portal de Panzea (Altschul *et al.* 1997; www.panzea.org). En cada búsqueda se solicitaron las

coincidencias con un valor-*e* menor a 0.001. Además, esta localización se confirmó mediante búsquedas en la base de datos de exones del conjunto de genes filtrados (*Filtered Gene Set*) del proyecto de secuenciación del genoma del maíz (www.maizesequence.org).

La distancia, en pares de bases, y el desequilibrio de ligamiento entre pares de SNPs en cada cromosoma se estimaron con Tassel v.3.0 (Bradbury *et al.* 2007). El desequilibrio de ligamiento se define como la asociación no aleatoria de alelos en loci distintos del genoma (Flint-Garcia *et al.* 2003). Esta propiedad se midió calculando la correlación (r^2) entre los alelos de dos SNPs para la matriz completa de marcadores. La significancia de este estadístico se obtuvo con una prueba exacta de Fisher de dos colas que calcula el programa. El desequilibrio de ligamiento en el genoma se representó gráficamente con un mapa de calor (*heat map*) basado en los estimados pareados de r^2 y sus valores de significancia (*p*) para cada par de SNPs.

La función de las proteínas sintetizadas por los genes con SNPs con alta F_{ST} y, en particular, con marcadores que generan sustituciones de nucleótido no sinónimas (ocurre una sustitución de aminoácido en la proteína sintetizada), se buscó en la base de datos de modelos de genes del portal de genética y genómica del maíz (MaizeGDB; Andorf *et al.* 2010; www.maizegdb.org). Los genes con SNPs no sinónimos posiblemente son candidatos a la selección natural debido a que la sustitución de aminoácido en la proteína sintetizada podría modificar la función de esta última (Futuyma 2013).

Análisis de diversidad y estructura genética espacial

Los análisis de diversidad y estructura genética espacial que se describen a continuación se llevaron a cabo con los SNPs, las isoenzimas y los microsatélites, con excepción del análisis espacial de componentes principales (sPCA), el cual sólo se hizo con los SNPs. La diversidad genética para cada provincia biogeográfica se midió con la heterocigosis observada (H_O) y la esperada sin sesgo (H_E ; Nei 1978), para todos los marcadores, y con el porcentaje de loci polimórficos ($\%P$), para los SNPs. Estos estadísticos se obtuvieron en GenAIEx v.6.5 (Peakall y Smouse 2006, 2012). El promedio general de alelos por locus (RA), para todos los marcadores, y el promedio de alelos privados (PA), para las isoenzimas y los microsatélites, se calcularon con un método de rarefacción que obtiene estimados estandarizados al tamaño de muestra en ADZE v.1.0 (Szpiech *et al.* 2008).

La repartición de la varianza genética entre la localización geográfica (altitud, latitud y longitud) y las provincias biogeográficas se obtuvo con el StAMOVA. Este análisis se comparó con el de varianza molecular (AMOVA) para determinar el efecto de la localización geográfica en la distribución de las frecuencias alélicas. El AMOVA descompone la varianza genética total en componentes aditivos representados por la varianza explicada y la no explicada (Excoffier *et al.* 1992). El índice Φ_{ST} es un análogo del F_{ST} que resume las agrupaciones de la información genética dentro y entre los grupos. El StAMOVA primero calcula una regresión de las covariables ecológicas o espaciales con la estructura genética y después descompone la varianza genética residual en los componentes aditivos y estima el índice $\Phi_{ST,cov.}$ (Dyer *et al.* 2004). Al obtener la diferencia de los estadísticos obtenidos en ambos análisis ($\Phi_{ST} - \Phi_{ST,cov.}$) se obtiene un estimado de cuanta variación genética observada es explicada por las covariables analizadas. La significancia estadística se estableció con 1,000 permutaciones. Estos análisis y la prueba de Mantel se llevaron a cabo usando el algoritmo Geno del programa Genetic Studio (Dyer *et al.* 2009).

La prueba de Mantel (Mantel 1967) permite comparar los elementos individuales de la matriz de distancias geográficas euclídeas con la matriz de distancias genéticas de Nei (1972). La hipótesis nula establece que las matrices de distancia no están correlacionadas. La porción triangular superior de estas matrices forman dos vectores, cada uno con una longitud $n(n-1)/2$, denominados d_Y y d_X . El estadístico de correlación r_M es el producto cruzado entre las distancias estandarizadas divididas por $(d-1)$, donde d es el número de distancias en la porción triangular superior de cada matriz cuando el estadístico es calculado a partir de matrices simétricas (Legendre y Fortin 2010). Este estadístico es la raíz cuadrada del coeficiente de determinación r_M^2 de la regresión lineal de d_Y y d_X , varía entre -1 y 1 y su significancia fue calculada con 1,000 permutaciones.

El análisis de autocorrelación genética espacial de una población establece la distribución de los genotipos en el espacio. Los intervalos de distancia (24) con tamaños de muestra iguales se determinaron en GenAlEx v.6.5 (Peakall y Smouse 2006, 2012). La hipótesis nula de este análisis establece que los genotipos se distribuyen aleatoriamente en el espacio y se cumple cuando el coeficiente de autocorrelación (r) es igual a cero. En cambio, la hipótesis alternativa establece que los genotipos presentan un patrón de distribución identificable ($r \neq 0$). La significancia del

estadístico se calculó mediante permutaciones aleatorias y bootstrap (ambos con 10,000 repeticiones), mientras que la del correlograma se determinó con la prueba de heterogeneidad (Smouse *et al.* 2008).

El sPCA resume la variabilidad genética entre los individuos en unos cuantos componentes informativos y encuentra los posibles patrones espaciales que estos componentes pudieran mostrar. Estos objetivos se consiguen al encontrar unas cuantas variables sintéticas independientes que optimizan el producto entre la varianza del puntaje de entidad y el estadístico I de Moran (Moran 1948, 1950). Los puntajes con una varianza grande y alta autocorrelación espacial positiva (estructuras regionales) corresponden a altos valores eigen positivos, mientras que los puntajes con una varianza grande y alta autocorrelación espacial negativa (estructuras locales) corresponden a altos valores eigen negativos. Las estructuras regionales significativas representan parches, clinas o patrones intermedios, mientras que las estructuras locales significativas indican que hay una alta diferenciación genética entre individuos vecinos.

Este análisis se corrió en el paquete *adeigenet* v.1.3-4 en R (Jombart *et al.* 2008). La información espacial necesaria para el cálculo del estadístico I de Moran se proporciona al sPCA a partir de una red de conexión que puede ser una gráfica de vecindario o una matriz de distancias euclidianas (Jombart *et al.* 2008). En este caso, se usaron las distancias inversas en donde los pesos espaciales de los sitios vecinos son directamente proporcionales al inverso de las distancias espaciales. Esta opción, en donde todos los sitios son vecinos, es adecuada para casos, como el del maíz, donde se cuenta con un muestreo irregular y alta conectividad entre las poblaciones (Jombart *et al.* 2008; Robinson *et al.* 2012). Los pares de nucleótidos faltantes fueron reemplazados por la media de las frecuencias alélicas. La identificación de los componentes que proporcionan estructuras regionales y locales biológicamente interpretables se hizo revisando tres tipos de resultados: las gráficas de barras de los valores eigen, las de descomposición de estos valores en su varianza y autocorrelación espacial y las pruebas estadísticas regionales y locales. La hipótesis nula de estas pruebas multivariadas de Monte Carlo establece que las frecuencias alélicas individuales se distribuyen aleatoriamente en la red de conexión. La significancia del estadístico $\max(t)$ obtenido se comprobó con 1,000 permutaciones.

RESULTADOS

SNPs candidatos a la adaptación local

Extremos de latitud y altitud

Los 169 individuos de 50 razas se distribuyen en un intervalo de 5 a 2,769 msnm entre los 29°54'26''N y los 14°49'13''N (Figura 1; Anexo II.A). La máxima diferencia de altitud entre pares de individuos de una raza es de 2,495 m, la cual se registró entre uno de Baja California y otro de Guanajuato de Mushito. Además, 13 razas adicionales presentan pares de individuos con diferencia de altitud en su mayoría entre 1,000 y 1,500 m. Todas las combinaciones pareadas posibles de cada una de estas 14 razas se usaron para obtener la diferencia de latitud en metros y para calcular la hipotenusa (a partir de la diferencia de altitud y de latitud). La hipotenusa entre algunos pares de individuos de la mayoría de estas razas es mayor a 100 km. Las excepciones incluyeron a todas las combinaciones de Comiteco, Zamorano Amarillo y Dulcillo del Noroeste, por lo que estas tres razas fueron eliminadas. Uno de los pares de individuos con hipotenusa mayor a 100 km de las 11 razas restantes fue seleccionado aleatoriamente para obtener el índice F_{ST} (Cuadro 2; Anexo III).

Cuadro 2. Diferencias de latitud y altitud y la hipotenusa de ambas entre pares de individuos de 11 razas. Las localidades, la altitud y la latitud de cada individuo se encuentran en el Anexo II.A.

Muestras	Raza	Latitud (m)	Altitud (m)	Hipotenusa (km)
35-70	Chalqueño	172867.06	678	172.868
157-26	Coscomatepec	180054.94	75	180.055
16-173	Elotero de Sinaloa	105985.48	38	105.985
105-49	Mushito	621364.03	1845	621.367
99-84	Nal-tel de Altura	291566.48	632	291.567
167-166	Olotón	148723.59	1079	148.728
119-121	Tabloncillo	605086.63	1457	605.088
177-17	Tepecintle	276596.58	20	276.596
134-133	Tuxpeño Norteño	686610.02	1056	686.611
102-160	Elotes Cónicos	148589.33	902	148.592
111-113	Ratón	172118.54	770	172.120

Las 11 razas de maíz que presentan individuos en extremos de latitud y altitud no comparten SNPs con un valor de F_{ST} igual a uno. Este resultado se repite si se incluyen en las comparaciones de marcadores a todos los pares de individuos de cada una de estas razas con hipotenusa mayor a 100 km. Debido a esto, no se encontró un conjunto de SNPs candidatos a la adaptación local que mostraran loci en el genoma del maíz sujetos a la adaptación a distintos ambientes en extremos de latitud y altitud.

Distribuciones geográficas restringidas y extensas

Las 50 razas que cuentan con datos de SNPs presentan distribuciones geográficas contrastantes. Aquellas seleccionadas por su distribución geográfica fueron 13 para la categoría de restringida y 14 para la de extensa. Entre las razas con distribución geográfica extensa están las 11 mencionadas en la sección anterior con excepción de Coscomatepec, Mushito, Nal-tel de Altura y Olotón. Éstas no cuentan con una distribución geográfica que cumpla con los requisitos establecidos en la metodología.

Los individuos de razas con distribución geográfica extensa se localizan en su mayoría en la Faja Volcánica, por lo que su altitud es mayor y su latitud menor a los límites establecidos (2,000 msnm y 20°N). Debido a esto, los dos grupos para obtener el índice F_{ST} en esta categoría de distribución consistieron en individuos que se localizan a una altitud baja y una latitud alta y viceversa. Ningún individuo de Cónico Norteño cumplió con estos requisitos. El índice F_{ST} se obtuvo de un total de 51 individuos seleccionados de 13 razas (que incluyen a Celaya, Cónico, Elotes Occidentales, Olotillo, Tuxpeño y Vandeño), repartidos en grupos de altitud y latitud de 24 y 27 integrantes.

Los individuos de razas con distribución geográfica restringida se localizan en su mayoría en la Sierra Madre Occidental, por lo que tanto su altitud como su latitud son mayores a los límites establecidos (2,000 msnm y 20°N). Debido a esto, los dos grupos para obtener el índice F_{ST} en esta categoría de distribución consistieron en individuos que se localizan a una altitud y una latitud alta o baja. Los individuos de Harinoso de Ocho, Jala, Tehua y Mountain Yellow no cumplieron con estos requisitos. El índice F_{ST} se obtuvo de un total de 16 individuos de nueve razas (Chiquito, Dzit-Bacal, Zapalote Chico, Mixeño, Gordo, Apachito, Cristalino de Chihuahua,

Azul y Palomero de Chihuahua), repartidos en grupos de altitud y latitud de siete y nueve integrantes. El límite altitudinal se modificó a 1,500 msnm para que ambas agrupaciones quedaran con un número similar de individuos.

A un valor de F_{ST} igual o mayor a 0.4, dos SNPs, SYN36384 y PZE-105063649, que generan sustituciones no sinónimas y sinónimas en los genes GRMZM2G458824 y -088737, respectivamente, se comparten entre las dos categorías de distribución geográfica. La proteína que se sintetiza a partir del gen GRMZM2G458824 no está caracterizada ni se ha encontrado un ortólogo sinténico (misma localización en el cromosoma) con el sorgo (*Sorghum bicolor*; Andorf *et al.* 2010; www.maizegdb.org). En cambio, en la síntesis de la proteína a partir del gen GRMZM2G088737 participa un ácido ribonucleico mensajero (ARNm) registrado en el *Transcriptome Shotgun Assembly* (TSA) de la NCBI del NIH (www.ncbi.nlm.nih.gov/genbank/tsa). El SNP (PZE-105063649) presente en este gen también tuvo alta diferenciación genética ($F_{ST}=0.187$) entre provincias biogeográficas (siguiente sección).

SNPs candidatos a la selección natural positiva y su localización en el genoma

Los 41,299 SNPs presentan un valor de F_{ST} promedio entre provincias biogeográficas de 0.025, con un intervalo de -0.049 a 0.232 . En total, 154 SNPs tienen un valor de F_{ST} igual o mayor a 0.15. Debido a su alta diferenciación genética, es posible que algunos de estos marcadores se localicen en genes candidatos a la selección natural (positiva) o en loci cuya variación está influida por esta fuerza evolutiva y, por consiguiente, en el texto se les denomina de esta manera. Aleatoriamente se seleccionaron 200 SNPs que presentan una F_{ST} entre cero y 0.02. La baja diferenciación genética en estos marcadores sugiere que posiblemente son neutrales. Con estos dos conjuntos de SNPs se llevaron a cabo los análisis de diversidad y estructura genética espacial.

La localización de los 154 SNPs con valor de F_{ST} igual o mayor a 0.15 en el genoma de referencia de la línea mejorada endogámica B73 muestra que se distribuyen en los 10 cromosomas, con una cobertura que varía de seis en el sexto cromosoma a 21 en el primer, tercer y quinto cromosomas (Anexo IV). Dos SNPs (PZA00417.2 y PZE-108038903) se localizan en la misma posición en el octavo cromosoma y representan la misma sustitución de nucleótido por lo que uno fue eliminado. La posición del marcador PUT-163a-74235050-3609 en el tercer

cromosoma no pudo ser determinada con exactitud debido a que los dos extremos de la secuencia de referencia se alinean perfectamente en lugares separados por 138 posiciones.

La distancia promedio entre pares de SNPs en un mismo cromosoma es de 62 megapares de bases, con un intervalo de 23 pares de bases a 284 megapares de bases. El 27.6% (234), más de la cuarta parte, de las comparaciones pareadas en el mismo cromosoma (un total de 848) muestran una correlación significativa ($p < 0.01$) entre sus alelos, con una variación en r^2 de 0.052 a 1.00 (Figura 3). La media general de r^2 entre pares de SNPs en un cromosoma es 0.089. Además, el 16.7% (1,214) de las comparaciones pareadas entre marcadores en distintos cromosomas (un total de 7,279) tienen una correlación significativa ($p < 0.01$) entre sus alelos, con una variación en r^2 de 0.046 a 0.262 (Figura 3). Por lo tanto, los cambios en las frecuencias alélicas de estos pares de SNPs están asociados, es decir, presentan desequilibrio de ligamiento.

El 64% (98) de los marcadores con alta F_{ST} se localizan en genes (91 en total) y los restantes (55) se hallan en loci no codificantes del genoma. De los 98 SNPs en genes, 39 (40%) se localizan en intrones y 30 (31%) generan una sustitución sinónima (no ocurre una sustitución de aminoácido en la proteína sintetizada). No fue posible determinar el tipo de sustitución (sinónima o no sinónima) de 20 marcadores en exones debido a que no se encontró información al respecto.

Seis genes (GRMZM2G171324, -157127, -359746, -012213, -125294, -107306) cuentan con SNPs (uno por gen) que generan sustituciones no sinónimas. Además, cuatro genes (GRMZM2G113911, -081554, -130348, -016660) presentan dos a cuatro marcadores, cada uno, con alta F_{ST} . Las proteínas que se sintetizan a partir de estos genes no se han caracterizado. La excepción es el transportador de hierro vacuolar 1 (VIT1) que es sintetizado a partir del gen GRMZM2G107306. Esta proteína es una transportadora de hierro necesaria para el almacenamiento de este mineral en vacuolas, se localiza en la membrana vacuolar y se expresa durante el desarrollo de la semilla y el embrión (Kim *et al.* 2006). Este gen ya había sido reportado anteriormente como candidato a la selección positiva direccional (van Heerwaarden *et al.* 2012).

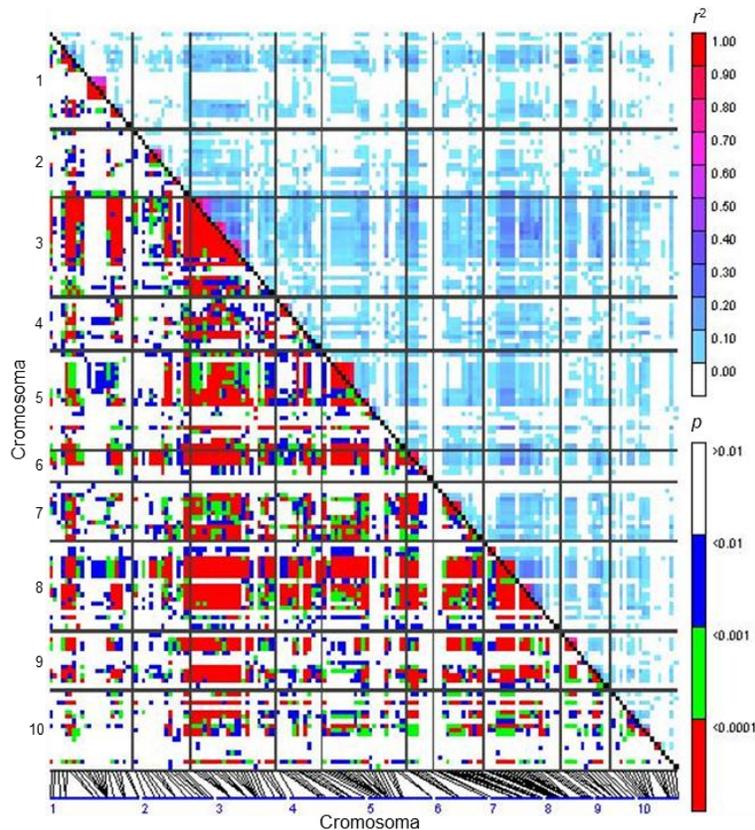


Figura 3. Mapa de disequilibrio de ligamiento por cromosoma entre todos los pares posibles de marcadores de un total de 153 SNPs con índice $F_{ST} \geq 0.15$ de 169 individuos de 50 razas. La diagonal negra representa las comparaciones de un marcador con sí mismo. La parte superior muestra el estadístico de disequilibrio de ligamiento (r^2) y la inferior, su significancia (p). En la parte inferior se muestra la posición de cada SNP en el cromosoma correspondiente.

Siete genes más (GRMZM2G110913, -096171, -029323, -016660, -048170, -025242, -035899), que tienen SNPs con alta F_{ST} entre provincias biogeográficas, ya habían sido detectados previamente como posibles candidatos a la selección por domesticación o mejoramiento (Hufford *et al.* 2012a). Las proteínas caracterizadas y sintetizadas a partir de estos genes presentan funciones relacionadas con la transcripción (GRMZM2G029323 y -096171) o con modificaciones del producto durante la síntesis proteica (GRMZM2G035899 y -016660). Otros dos genes (GRMZM2G079089, -317406), con SNPs diferenciados (PZE-103172227, PUT-163a-78116163-4360, respectivamente), también codifican proteínas caracterizadas (www.maizesequence.org). El gen GRMZM2G079089 codifica para un homólogo de

tiorredoxina *h* (*trh1*), una proteína con una estructura y función similar a la tiorredoxina (una enzima oxidorreductasa) presente en las levaduras y las plantas (Ishiwatari *et al.* 1998). El gen GRMZM2G317406 codifica para una proteína específica del gametofito masculino (*mgs1*) localizada en el citoplasma del grano de polen y el tubo polínico (Stinson *et al.* 1987).

Diversidad genética

Los SNPs analizados presentan dos alelos por locus. La proporción de alelos por locus es ligeramente menor para los 153 SNPs candidatos a la selección positiva que para aquellos posiblemente neutrales, sin embargo, ésta es similar en todas las provincias biogeográficas en ambos casos. La riqueza alélica (*RA*) promedio ajustada al tamaño de muestra es 1.282 (± 0.016), con una variación entre 1.136 a 1.379, para los candidatos a la selección positiva y 1.343 (± 0.010), con una variación entre 1.186 a 1.380, para los posiblemente neutrales (Cuadro 3). La proporción de SNPs con dos alelos en la muestra analizada es cercana al 100%, para los marcadores candidatos a la selección positiva y principalmente para aquellos posiblemente neutrales, en la mayoría de las provincias biogeográficas. El porcentaje de SNPs polimórficos (*P*) varía entre 13.6 y 98%, con una media de 70.3%, para los candidatos a la selección positiva y entre 18.5 y 100%, con una media de 79.9% para los posiblemente neutrales (Cuadro 3).

Las isoenzimas y los microsatélites analizados cuentan con proporciones similares de alelos para cada marcador en todas las provincias biogeográficas. Los microsatélites tienen una mayor proporción de alelos por locus que los SNPs y las isoenzimas debido a que son marcadores polimórficos (más de dos alelos en un locus). La riqueza alélica (*RA*) promedio ajustada al tamaño de muestra ($n=2$) para cada provincia biogeográfica es 1.268 (± 0.007), con una variación entre 1.266 y 1.30, para las isoenzimas y 1.786 (± 0.007), con una variación entre 1.778 y 1.823, para los microsatélites (Anexo V.A). A diferencia de los SNPs, las isoenzimas y los microsatélites cuentan con alelos únicos en cada provincia biogeográfica. Los alelos privados (*PA*) por provincia varían entre 0.027 y 0.064, con una media de 0.048 (± 0.002), para las isoenzimas y entre 0.384 y 0.428, con una media de 0.390 (± 0.005), para los microsatélites (Anexo V.A).

Cuadro 3. Estadísticos de diversidad genética por provincia biogeográfica de 153 SNPs con índice $F_{ST} \geq 0.15$ y de 200 SNPs con índice $F_{ST} \leq 0.02$ (sombreado) de 169 individuos de 50 razas. Los totales se calcularon a partir de las medias de cada provincia. El error estándar está entre paréntesis. En las provincias con un individuo se muestran los promedios de variación entre todos los marcadores de ese individuo.

Provincia biogeográfica	N^1 (R^2)	RA^3	H_O^4	H_E^5	$\%P^6$
Altiplano Norte (Chihuahuense)	2 (2)	1.325 (0.023)	0.266 (0.027)	0.325 (0.023)	57.1
		1.352 (0.020)	0.298 (0.023)	0.353 (0.020)	62.5
Altiplano Sur (Zacatecano-Potosino)	4 (4)	1.156 (0.017)	0.106 (0.014)	0.156 (0.017)	40.3
		1.343 (0.015)	0.282 (0.018)	0.345 (0.015)	79.0
Altos de Chiapas	11 (6)	1.293 (0.015)	0.240 (0.015)	0.293 (0.015)	83.1
		1.349 (0.011)	0.302 (0.012)	0.350 (0.011)	96.5
Baja California	1 (1)	1.136 (0.028)	0.136 (0.028)	0.136 (0.028)	13.6
		1.186 (0.028)	0.185 (0.028)	0.185 (0.028)	18.5
Depresión del Balsas	4 (3)	1.379 (0.016)	0.260 (0.019)	0.379 (0.016)	83.8
		1.380 (0.014)	0.335 (0.018)	0.381 (0.014)	85.5
Golfo de México (Norte)	6 (3)	1.357 (0.015)	0.254 (0.016)	0.357 (0.015)	83.8
		1.346 (0.012)	0.308 (0.014)	0.347 (0.012)	90.0
Golfo de México (Sur)	5 (4)	1.310 (0.017)	0.206 (0.015)	0.310 (0.017)	73.4
		1.354 (0.013)	0.305 (0.016)	0.355 (0.013)	88.5
Faja Volcánica	45 (11)	1.228 (0.012)	0.180 (0.010)	0.228 (0.012)	98.0
		1.365 (0.009)	0.325 (0.009)	0.365 (0.009)	100.0
Costa del Pacífico (Norte)	19 (9)	1.314 (0.014)	0.256 (0.012)	0.314 (0.014)	96.8
		1.359 (0.010)	0.312 (0.011)	0.360 (0.010)	98.5
Costa del Pacífico (Sur)	18 (10)	1.334 (0.014)	0.250 (0.012)	0.334 (0.014)	91.6
		1.344 (0.010)	0.288 (0.011)	0.344 (0.010)	99.5
Sierra de Oaxaca	1 (1)	1.312 (0.037)	0.312 (0.037)	0.312 (0.037)	31.2
		1.327 (0.033)	0.325 (0.033)	0.325 (0.033)	73.0
Sierra Madre Occidental	18 (11)	1.264 (0.014)	0.193 (0.012)	0.264 (0.014)	90.3
		1.352 (0.010)	0.299 (0.011)	0.353 (0.010)	98.0
Sierra Madre Oriental	18 (7)	1.241 (0.013)	0.193 (0.012)	0.241 (0.013)	89.0
		1.369 (0.010)	0.329 (0.010)	0.370 (0.010)	99.5

Provincia biogeográfica	N ¹ (R ²)	RA ³	H _O ⁴	H _E ⁵	%P ⁶
Sierra Madre del Sur	3 (2)	1.337 (0.020)	0.264 (0.021)	0.337 (0.020)	68.2
		1.345 (0.016)	0.273 (0.017)	0.344 (0.016)	73.0
Soconusco	5 (3)	1.227 (0.017)	0.196 (0.018)	0.227 (0.017)	60.4
		1.371 (0.013)	0.329 (0.016)	0.372 (0.013)	87.0
Sonorense	7 (5)	1.325 (0.014)	0.260 (0.014)	0.325 (0.014)	90.3
		1.361 (0.012)	0.311 (0.014)	0.362 (0.012)	90.5
Tamaulipeca	2 (2)	1.254 (0.024)	0.269 (0.029)	0.254 (0.024)	44.2
		1.333 (0.020)	0.315 (0.024)	0.333 (0.020)	60.0
Total	169	1.282 (0.016)	0.226 (0.005)	0.282 (0.005)	70.3 (6.1)
		1.343 (0.010)	0.301 (0.004)	0.344 (0.004)	79.9 (5.8)

¹N: número de individuos; ²R: razas por provincia; ³RA: riqueza alélica; ⁴H_O: heterocigosis observada; ⁵H_E: heterocigosis esperada; ⁶%P: porcentaje de loci polimórficos.

La diversidad genética, medida con la heterocigosis, es menor entre los SNPs candidatos a la selección positiva que entre los posiblemente neutrales, aunque en los dos casos es similar en cada provincia biogeográfica. La heterocigosis esperada sin sesgo (H_E) promedio es 0.282 (± 0.005), con una variación entre 0.136 y 0.379, para los candidatos a la selección positiva y 0.344 (± 0.004), con una variación entre 0.185 y 0.381, para los posiblemente neutrales (Cuadro 3). La heterocigosis observada (H_O) es menor o igual a la esperada, y varía entre 0.106 y 0.312, con una media de 0.226 (± 0.005), para los candidatos a la selección positiva y entre 0.185 y 0.335, con una media de 0.301 (± 0.004), para los posiblemente neutrales (Cuadro 3).

Los valores de diversidad genética, en relación con los SNPs, son menores aunque similares en las isoenzimas y mayores en los microsatélites. La heterocigosis esperada sin sesgo (H_E) promedio es 0.265 (± 0.010), con una variación entre 0.188 y 0.298, para las isoenzimas y 0.782 (± 0.004), con una variación entre 0.719 y 0.818, para los microsatélites (Anexo V.A). La heterocigosis observada (H_O) es menor o igual a la esperada, y varía entre 0.134 y 0.236, con una media de 0.202 (± 0.008), para las isoenzimas y entre 0.457 y 0.629, con una media de 0.554 (± 0.006), para los microsatélites (Anexo V.A).

Estructura genética espacial

Localización geográfica (altitud, latitud y longitud)

La localización geográfica, evaluada con la altitud, la latitud y la longitud, determina la distribución espacial de la variación genética medida con los SNPs candidatos a la selección positiva y con las isoenzimas. Para los SNPs, las provincias biogeográficas explican significativamente 23% ($\Phi_{ST}=0.230$; $p<0.001$) de la varianza genética total (Cuadro 4). De este porcentaje, 11.4%, casi la mitad, se debe significativamente ($p<0.001$) a la interacción entre la latitud y la altitud, de acuerdo a la diferencia entre los estadísticos Φ_{ST} obtenidos con el AMOVA y el StAMOVA ($\Phi_{ST,cov.}=0.116$; $p<0.001$). En el caso de las isoenzimas, las provincias biogeográficas explican significativamente 5.5% ($\Phi_{ST}=0.055$; $p<0.001$) de la varianza genética total (Anexo V.B). De este porcentaje, 2.2%, un poco menos de la mitad, se debe significativamente a la altitud ($p<0.01$), a la interacción entre la latitud y la altitud ($p<0.001$) y a la interacción entre la latitud y la longitud ($p<0.05$), de acuerdo al StAMOVA ($\Phi_{ST,cov.}=0.033$; $p<0.001$).

La localización geográfica no parece ser un factor que determine la distribución espacial de la variación genética medida con los SNPs posiblemente neutrales y con los microsatélites. Para los SNPs, las provincias biogeográficas explican significativamente 0.5% ($\Phi_{ST}=0.005$; $p<0.001$) de la varianza genética total (Cuadro 5). En el caso de los microsatélites, las provincias explican significativamente 2.3% ($\Phi_{ST}=0.023$; $p<0.001$) de la varianza genética total (Anexo V.C). En ambas situaciones, la altitud, la latitud, la longitud y sus interacciones pareadas no influyen de manera significativa en la varianza genética explicada por las provincias biogeográficas, de acuerdo a lo obtenido en los StAMOVAs respectivos (Cuadro 5; Anexo V.C).

Cuadro 4. Análisis de varianza molecular (AMOVA y StAMOVA) entre provincias biogeográficas de 153 SNPs con índice $F_{ST} \geq 0.15$ de 169 individuos de 50 razas.

Fuente	df¹	SS²	MS³	p⁴	Φ_{ST}
AMOVA					
Provincia Biogeográfica	16	3115.492	194.718	0.0001	0.230
Error	152	7900.729	51.979		
Total	168	11016.221			
StAMOVA					
Altitud	1	308.519	308.519	0.068	
Latitud	1	11.490	11.490	0.473	
Longitud	1	-1.743	-1.743	0.504	
Altitud-Latitud	1	770.961	770.961	0.001	
Altitud-Longitud	1	92.992	92.992	0.342	
Latitud-Longitud	1	172.838	172.838	0.221	
Provincia Biogeográfica	16	1880.204	117.513	0.0001	0.116
Error	146	7780.959	53.294		
Total	168	11016.221	65.573		

¹df: grados de libertad; ²SS: suma de cuadrados; ³MS: cuadrados medios; ⁴p: significancia.

Cuadro 5. Análisis de varianza molecular (AMOVA y StAMOVA) entre provincias biogeográficas de 200 SNPs con índice $F_{ST} \leq 0.02$ de 169 individuos de 50 razas.

Fuente	df¹	SS²	MS³	p⁴	Φ_{ST}
AMOVA					
Provincia Biogeográfica	16	1359.230	84.952	0.0001	0.005
Error	152	12348.809	81.242		
Total	168	13708.040			
StAMOVA					
Altitud	1	47.448	47.448	0.311	
Latitud	1	3.565	3.565	0.466	
Longitud	1	-0.710	-0.710	0.482	
Altitud-Latitud	1	122.093	122.093	0.159	
Altitud-Longitud	1	14.485	14.485	0.432	
Latitud-Longitud	1	89.522	89.522	0.224	
Provincia Biogeográfica	16	1119.040	69.940	0.0001	-0.019
Error	146	12312.597	84.333		
Total	168	13708.040	81.596		

¹df: grados de libertad; ²SS: suma de cuadrados; ³MS: cuadrados medios; ⁴p: significancia.

Distancia geográfica

La distancia geográfica, de acuerdo a las pruebas de Mantel llevadas a cabo con los tres tipos de marcadores, no influye en la variación de las distancias genéticas. Las correlaciones entre las distancias geográficas y las genéticas de los SNPs candidatos a la selección positiva ($r_M=0.021$, $p=0.177$; Figura 4A), de los SNPs posiblemente neutrales ($r_M=-0.026$, $p=0.923$; Figura 4B), de las isoenzimas ($r_M=0.039$, $p=0.023$; Anexo V.D) y de los microsatélites ($r_M=0.002$, $p=0.442$; Anexo V.E) no son significativas. En estos casos, los intervalos de distancia genética varían en cada marcador, siendo de 0.30 a 0.872 para los SNPs candidatos a la selección positiva, de 0.122 a 0.431 para los SNPs posiblemente neutrales, de 0.014 a 0.574 para las isoenzimas y de 0.276 a 2.401 para los microsatélites. En cambio, las distancias geográficas están entre cero y 2,764 km, para los SNPs, o 2,652 km, para las isoenzimas y los microsatélites; una diferencia de 112 km.

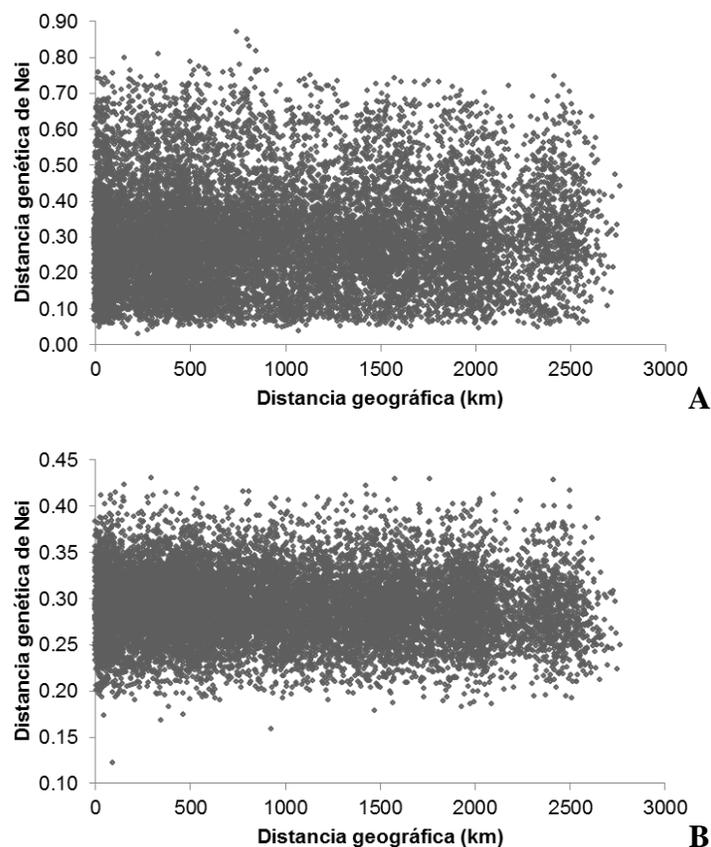


Figura 4. Correlación entre la distancia geográfica (en kilómetros) y la distancia genética de Nei de 169 individuos de 50 razas. A. 153 SNPs con índice $F_{ST} \geq 0.15$ ($r_M=0.021$; $p=0.177$). B. 200 SNPs con índice $F_{ST} \leq 0.02$ ($r_M=-0.026$; $p=0.923$).

Estructura espacial a escala fina

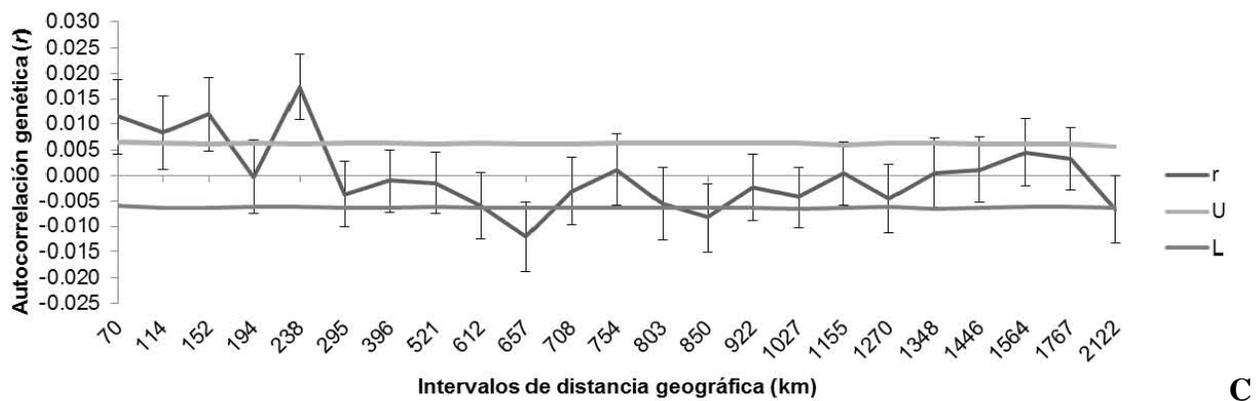
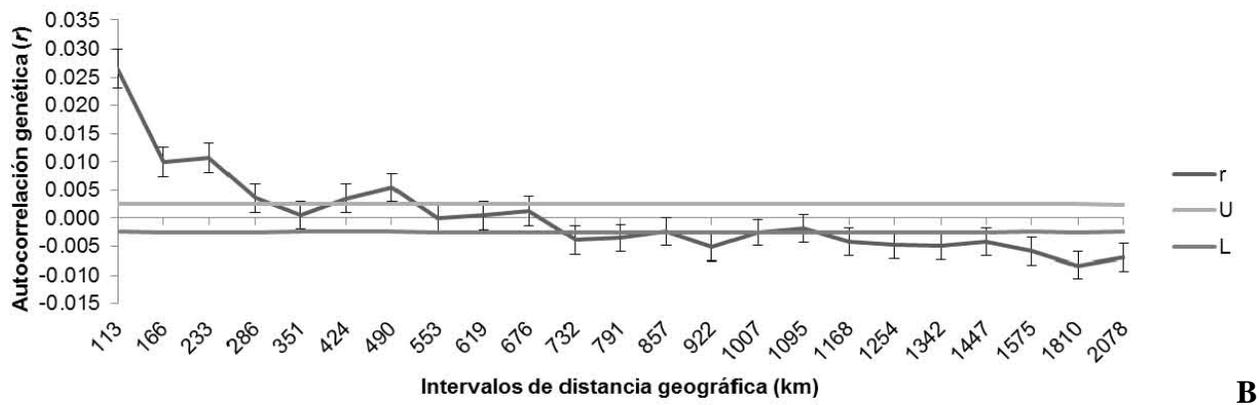
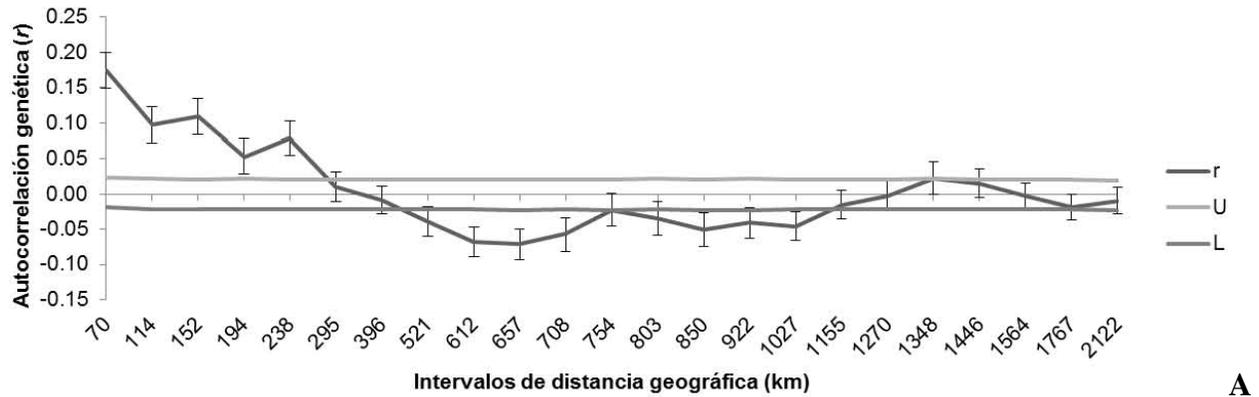
Los genotipos individuales, estén o no posiblemente influidos por la selección natural, muestran patrones de distribución espacial que varían a distintos intervalos de distancia geográfica. El correlograma de los SNPs candidatos a la selección positiva (Figura 5A) indica que los genotipos vecinos son significativamente similares hasta alrededor de 295 km. El coeficiente de autocorrelación genética (r), en este intervalo de distancia geográfica, es positivo y disminuye de 0.174 a 0.011. En cambio, estos genotipos son significativamente distintos entre distancias de alrededor de 521 a 1,027 km, en donde r es negativo y varía entre -0.022 y -0.071 . A partir de una distancia de alrededor de 1,155 km, los genotipos individuales se distribuyen de manera aleatoria, ya que r varía entre -0.018 y 0.023 dentro del intervalo de no significancia.

El patrón de distribución espacial de los genotipos individuales de los microsatélites es similar al obtenido con los SNPs candidatos a la selección positiva. El correlograma de los microsatélites (Figura 5B) muestra que los genotipos vecinos son significativamente similares hasta alrededor de 286 o 490 km, en donde r es positivo y disminuye de 0.026 a 0.003. Entre distancias de alrededor de 553 a 676 km, estos genotipos se distribuyen de manera aleatoria, ya que r varía entre cero y 0.001 dentro del intervalo de no significancia. A partir de una distancia de alrededor de 732 km, los genotipos individuales son significativamente distintos, debido a que r es negativo y varía entre -0.003 y -0.008 .

El correlograma de los SNPs posiblemente neutrales (Figura 5C) indica que los genotipos vecinos son significativamente similares, aunque de manera no tan evidente como en los casos anteriores, hasta alrededor de 238 km. En este intervalo de distancia geográfica, la variación de r tiene un patrón irregular entre cero y 0.017. A partir de una distancia de alrededor de 295 km, los genotipos se distribuyen de manera aleatoria, ya que r varía entre 0.005 y -0.007 dentro del intervalo de no significancia.

Los genotipos individuales de las isoenzimas muestran un patrón de distribución espacial similar al obtenido con los SNPs posiblemente neutrales. El correlograma de las isoenzimas (Figura 5D), calculado con 259 individuos escogidos al azar (un individuo por población), muestra que los genotipos vecinos son significativamente similares hasta alrededor de 276 km. En este intervalo de distancia geográfica, la variación de r tiene un patrón irregular entre -0.017

y 0.044. A partir de una distancia de alrededor de 332 km, los genotipos individuales se distribuyen de manera aleatoria, ya que r varía entre 0.011 y -0.027 dentro del intervalo de no significancia.



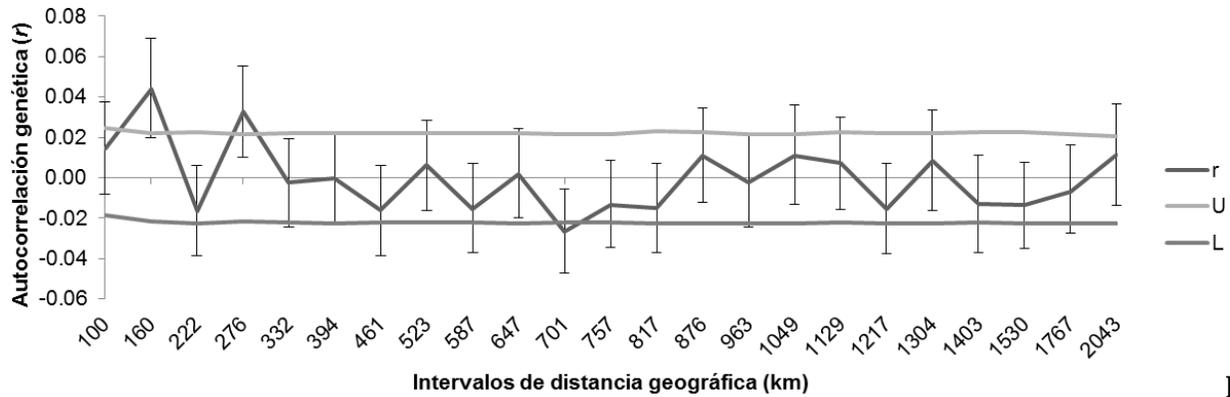


Figura 5. Correlogramas del coeficiente de autocorrelación genética (r) en función de 24 intervalos de distancia geográfica de individuos de 50-59 razas. Las distancias mostradas corresponden al valor más alto del intervalo. Límite superior (U) e inferior (L) del intervalo de confianza del 95% de no significancia. A. 153 SNPs con índice $F_{ST} \geq 0.15$. B. 96 microsatélites. C. 200 SNPs con índice $F_{ST} \leq 0.02$. D. 34 isoenzimas.

Estructura espacial a escala regional

Los SNPs candidatos a la selección positiva presentan estructura regional. Las evidencias de esto, de acuerdo al sPCA, se describen a continuación. El primer valor eigen es 2.118, el segundo 0.680 y los restantes valores positivos disminuyen gradualmente de 0.392 a cero (Anexo VI.A). En cambio, los valores eigen negativos varían de manera gradual de cero a -0.253 y la diferencia entre el último y el penúltimo, con un valor de -0.226 , no es considerable. Otra evidencia es que el primer valor eigen positivo tiene la mayor varianza, aunque no la mayor autocorrelación, entre todos estos valores, inclusive la varianza del segundo valor positivo no es muy distinta a la de los restantes (Anexo VI.C.a). En contraste, el último valor eigen negativo presenta una varianza y una autocorrelación similares a las de los demás valores negativos. La última evidencia es la prueba estadística que muestra que las frecuencias alélicas tienen un patrón de distribución no aleatorio a escala regional ($\max(t)=0.025$, $p<0.01$), pero esta distribución sí es aleatoria a escala local ($\max(t)=0.011$, $p=0.999$). En conclusión, el primer componente del sPCA determina la estructura regional que presentan los 153 SNPs con alta F_{ST} y no hay una estructura local significativa.

Los SNPs posiblemente neutrales no cuentan con estructura ni regional ni local, de acuerdo a las evidencias del sPCA. En primer lugar, los valores eigen varían gradualmente de positivos a negativos (Anexo VI.B). El primer valor eigen es 0.476, el segundo 0.394, el penúltimo -0.340 y el último -0.437 . En segundo lugar, el primer valor eigen positivo y el último valor negativo presentan una varianza similar entre sí y ligeramente distinta, aunque con una autocorrelación ni mayor ni menor, a la de los valores eigen restantes (Anexo VI.C.b). Por último, las pruebas estadísticas indican que las frecuencias alélicas tienen un patrón de distribución aleatorio tanto a escala regional ($\max(t)=0.008$, $p=0.30$) como a escala local ($\max(t)=0.008$, $p=0.341$). Por lo tanto, los SNPs con baja F_{ST} no muestran estructura espacial significativa.

Los alelos que contribuyen más del 5% a explicar el primer componente del sPCA corresponden a ocho SNPs candidatos a la selección positiva (Figura 6). Seis SNPs (PZE-105063649, -105063667, -107118935, -108031468, -109038290, -110066819) cuentan con purinas como alelos (adenina y guanina), mientras que los otros dos (PZE-105085845 y SYN32227), presentan purinas y pirimidinas (guanina o adenina y citosina). El marcador PZE-105063649 también tuvo una alta diferenciación genética compartida entre agrupaciones de razas con distribuciones geográficas contrastantes ($F_{ST} \geq 0.4$) y genera una sustitución sinónima en el gen GRMZM2G088737. El ARNm que participa en la transcripción de este gen está registrado en el portal TSA (www.ncbi.nlm.nih.gov/genbank/tsa). El marcador PZE-110066819 genera una sustitución no sinónima en el gen GRMZM2G107306. La proteína VIT1 se sintetiza a partir de este gen que ya había sido reportado como candidato a la selección positiva direccional (van Heerwaarden *et al.* 2012). Los seis marcadores restantes se localizan en loci no codificantes (PZE-105085845, -108031468), en intrones (PZE-105063667, -107118935, -109038290) de tres genes (GRMZM2G389462, -080567, -171031, respectivamente) o generan sustituciones no sinónimas (SYN32227 en el gen GRMZM2G012213). Las proteínas sintetizadas a partir de estos cuatro genes no se han caracterizado aunque sí presentan ortólogos sinténicos con el sorgo (*S. bicolor*).

La localización en un mapa de los puntajes de entidad individuales del primer componente del sPCA de los SNPs candidatos a la selección sugiere que las frecuencias alélicas se distribuyen en parches a escala regional (Figura 7). Los puntajes positivos varían entre 0.011 y 4.293, mientras que los puntajes negativos varían entre -0.374 y -5.795 . Cada grupo de puntajes (cuadros blancos o negros) representa una agrupación y su valor (tamaño de los cuadros) que tan

similares o diferentes son los individuos entre sí, de acuerdo a sus frecuencias alélicas. Todos o la mayoría ($\geq 60\%$) de los individuos del Altiplano Zacatecano-Potosino, de la Faja Volcánica, del sur del Golfo de México, de las Sierras Madre Occidental y Oriental y del Soconusco presentan puntajes positivos. En cambio, todos o la mayoría ($\geq 60\%$) de los individuos del Altiplano Chihuahuense, de los Altos de Chiapas, de Baja California, de la Depresión del Balsas, del norte del Golfo de México, de la Costa del Pacífico, de la Sierra de Oaxaca, de la Sierra Madre del Sur y de las regiones Sonorense y Tamaulipeca presentan puntajes negativos.

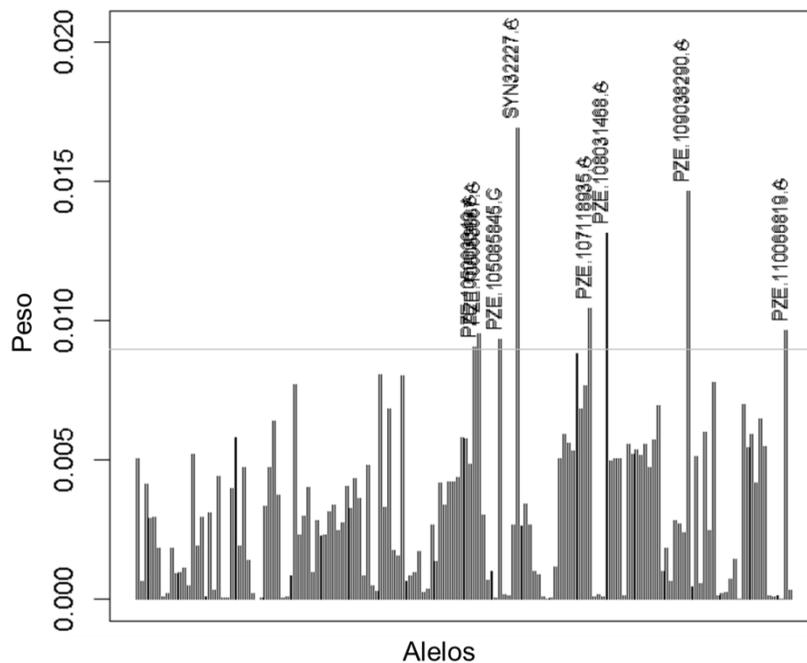


Figura 6. Contribución de los alelos al primer componente del sPCA de 153 SNPs con índice $F_{ST} \geq 0.15$ de 169 individuos de 50 razas. Los alelos con nombre contribuyen más del 5% a explicar este componente.

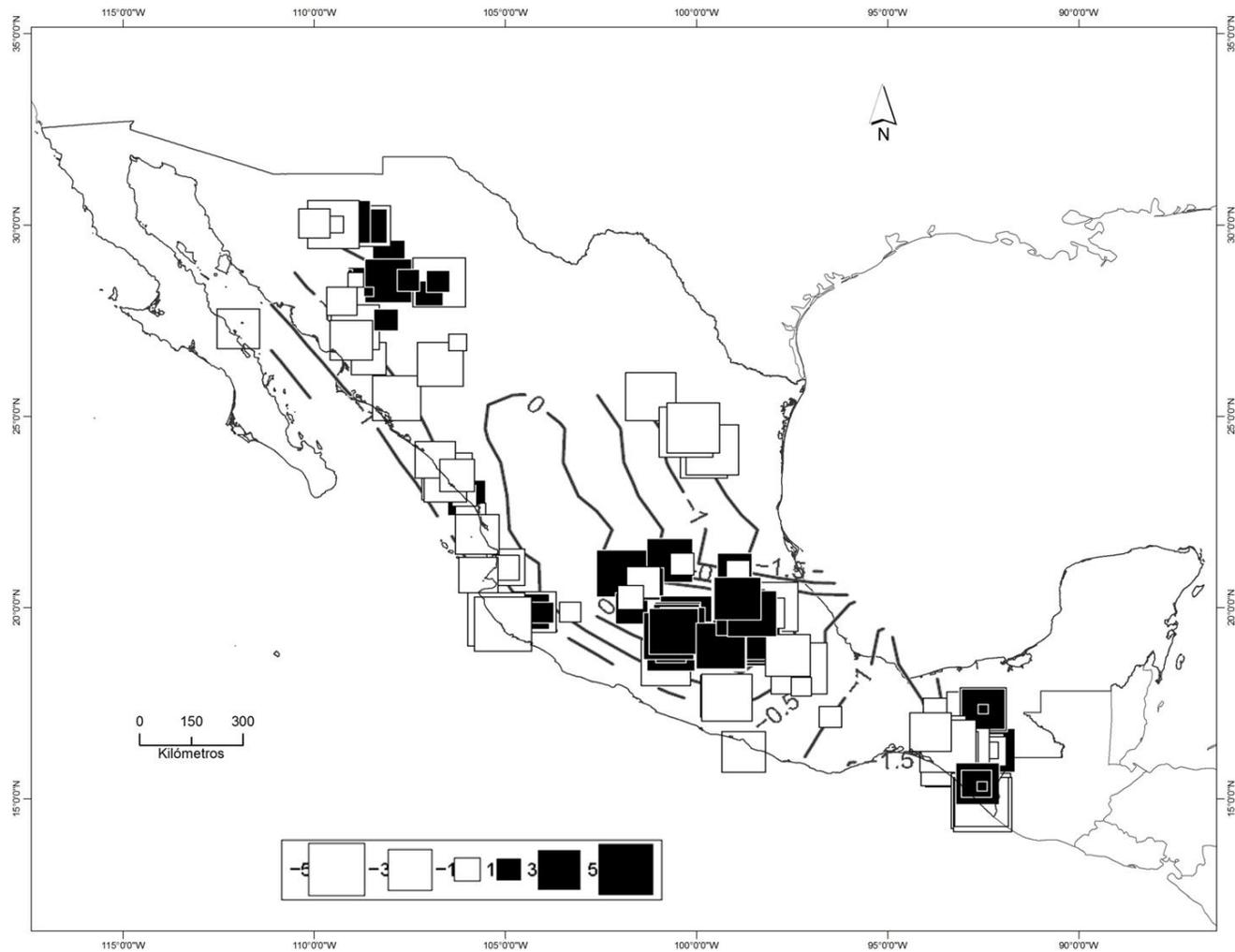


Figura 7. Distribución geográfica de los puntajes regionales del primer componente espacial del sPCA de 153 SNPs con índice $F_{ST} \geq 0.15$ de 169 individuos de 50 razas. Cada cuadro representa un individuo, los negros puntajes positivos, los blancos negativos y el tamaño la magnitud del puntaje en cada componente. Las líneas de contorno grises representan gradientes ortogonales entre conjuntos de individuos diferenciados.

DISCUSIÓN

Estructura genética espacial

El maíz cultivado en México presenta patrones espaciales de estructura genética regional que posiblemente están determinados, al menos en parte, por la selección natural (positiva). Las frecuencias alélicas de los SNPs con alta diferenciación genética tienen un patrón de distribución en parches a lo largo del país. Algunos de estos SNPs se localizan en genes que se han reportado previamente como candidatos a la selección positiva direccional por domesticación o mejoramiento (ver siguiente sección; Hufford *et al.* 2012a; van Heerwaarden *et al.* 2012). Además, otros SNPs con alta diferenciación genética podrían localizarse en loci del genoma cuya variación está influida por barridos selectivos.

Los barridos selectivos pueden ser identificados como loci con reducida variación genética en relación a otras regiones genómicas y por un incremento en el desequilibrio de ligamiento con loci cercanos (Schoville *et al.* 2012). En el genoma de los cultivos, la extensión del desequilibrio de ligamiento suele ser mayor que en sus parientes silvestres, en particular, en loci del genoma sujetos a intensa selección (Hamblin *et al.* 2011). El promedio de decaimiento del desequilibrio de ligamiento en el maíz es de cinco a 10 kilopares de bases (Yan *et al.* 2009). En cambio, el promedio que separa a los SNPs con alta diferenciación genética analizados es de 62 megapares de bases. Sin embargo, un poco más de la cuarta parte (27.6%) de estos marcadores que se localizan en un mismo cromosoma se encuentran en desequilibrio de ligamiento. Además, 16.7% de las comparaciones pareadas entre SNPs en distintos cromosomas también muestran señales de asociación entre sus alelos. Por consiguiente, es posible que algunos de estos marcadores se localicen en loci influidos por los efectos de los barridos selectivos.

La diversidad genética en la muestra analizada es similar a la reportada en otros estudios del maíz en México. La heterocigosis esperada (H_E) entre provincias biogeográficas es similar (0.188-0.298), para las isoenzimas (0.179-0.295; Sánchez *et al.* 2000), y en promedio mayor (0.782 ± 0.004), para los microsatélites (0.61; Reif *et al.* 2006), a la reportada con estos marcadores entre razas mexicanas. Además, el promedio de este estadístico (0.344 ± 0.004), para los SNPs posiblemente neutrales, es ligeramente mayor y con un intervalo de variación más

restringido (0.185-0.381) a los reportados con este marcador para algunas líneas mejoradas endogámicas (0.32, Hamblin *et al.* 2007; promedio 0.26 y entre 0.02-0.50, Jones *et al.* 2007).

La localización geográfica explica parte de la varianza genética del maíz repartida entre provincias biogeográficas con marcadores probablemente informativos acerca del efecto de la selección positiva. La interacción entre la latitud y la altitud representa casi la mitad del porcentaje de varianza explicado por las provincias biogeográficas con los SNPs con alta diferenciación genética (Cuadro 4), mientras que esta interacción, la altitud y la interacción entre la latitud y la longitud explican cerca de la mitad de este porcentaje entre provincias con las isoenzimas (Anexo V.B). Estos resultados sustentan lo reportado en otros estudios acerca del efecto de la altitud y la localización geográfica en la diversidad del cultivo (Brush y Perales 2007; Mercer *et al.* 2008; Hung *et al.* 2012; Díez *et al.* 2013).

La altitud tiene un papel importante en la agrupación racial del maíz (Brush y Perales 2007; Mercer *et al.* 2008), debido posiblemente a que los gradientes de temperatura y precipitación, que suelen asociarse a este factor, determinan la distribución potencial de éstas (Hufford *et al.* 2012b; Ureta *et al.* 2012). En los teocintles anuales, *Z. m. ssp. parviglumis* y *Z. m. ssp. mexicana*, la altitud determina significativamente la estructura poblacional entre ambas subespecies (Pyhäjärvi *et al.* 2012). En cambio, la latitud se asocia al periodo de luz solar diario que regula la floración en las plantas incluido el maíz (Diamond 2002; Hung *et al.* 2012). Sin embargo, las razas cultivadas en México se distribuyen en áreas donde el promedio de la duración del día durante la temporada de crecimiento varía entre 12.46 y 12.98 h (Ruiz *et al.* 2008), un intervalo reducido. Por otro lado, la latitud y la longitud, junto con la altitud, son buenos predictores del tamaño del genoma en el cultivo (Díez *et al.* 2013).

La distancia geográfica determina la distribución de los genotipos individuales del maíz y éstos últimos muestran patrones espaciales a fina escala que posiblemente se deben al efecto de la selección positiva y principalmente al de aislamiento por distancia. Los genotipos individuales vecinos son significativamente similares hasta distancias entre 295 y 286 o 490 km, de acuerdo a los correlogramas de los SNPs con alta diferenciación genética (Figura 5A) y de los microsatélites (Figura 5B), respectivamente, y estos genotipos son significativamente diferentes a más de 732 km con el segundo marcador. Los patrones espaciales obtenidos con estos marcadores son similares, aunque con distancias geográficas mayores, a los reportados para la

variación del tamaño del genoma en 22 razas mexicanas de maíz y en los teocintles anuales *Z. m. ssp. parviglumis* y *Z. m. ssp. mexicana* (Díez *et al.* 2013). En este estudio reciente se encontró que el estadístico *I* de Moran del tamaño del genoma es significativo en vecindarios de 65 km para las razas de maíz y de 75 km para los teocintles anuales (Figura 8).

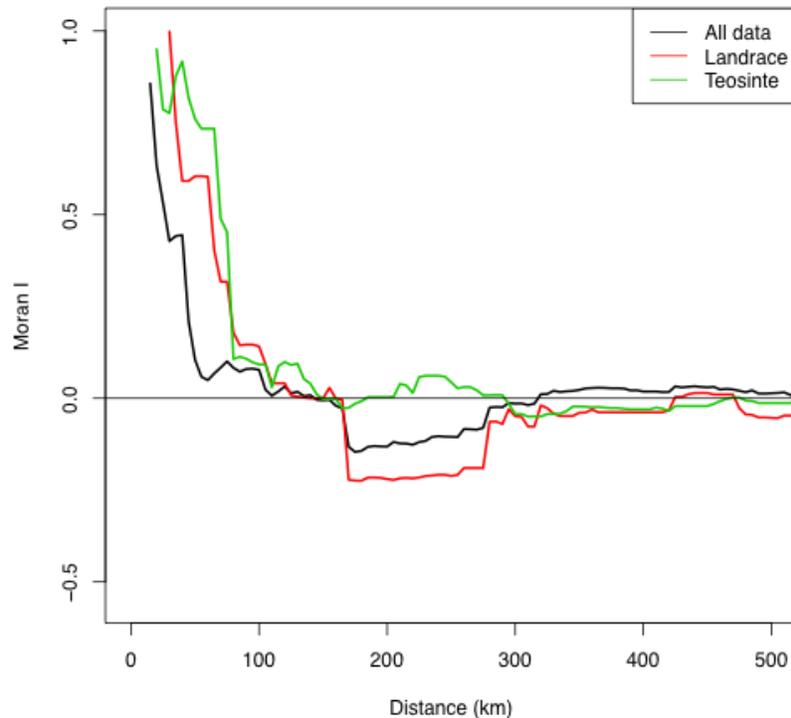


Figura 8. Correlograma del estadístico *I* de Moran del tamaño del genoma frente al tamaño del vecindario, medido en intervalos de distancia de 5 km. Este correlograma corresponde a la Figura S1 del estudio de Díez *et al.* 2013.

La similitud de los correlogramas de los microsatélites (Figura 5B) y del tamaño del genoma (Figura 8) sugieren que los patrones espaciales encontrados en el maíz corresponden al aislamiento por distancia. Este patrón espacial se ha detectado en este cultivo a escala nacional (México; este estudio) y a escala continental (América; Matsuoka *et al.* 2002a; Vigouroux *et al.* 2008), sin embargo, no se ha detectado a escala regional (Pressoir y Berthaud 2004) y esto posiblemente se deba al manejo tradicional. Los intercambios de semillas, que representan la migración de los individuos entre poblaciones, ocurren mayormente en un radio de 10 km, sin

embargo, éstos pueden extenderse, aunque no muy frecuentemente, hasta 50 km en México (Bellon *et al.* 2011) e inclusive hasta 300 km en América (Vigouroux *et al.* 2008). El manejo local de las semillas promueve el flujo génico entre los individuos a distancias de unos cuantos kilómetros y lo limita a mayores distancias, lo cual explicaría el patrón de aislamiento por distancia. Además, este patrón espacial se observa en los teocintles anuales, en donde explica la estructura poblacional dentro y entre las subespecies (Pyhäjärvi *et al.* 2012). En otros cultivos y sus parientes silvestres, como el arroz (*Oryza sativa*) el sorgo (*Sorghum* ssp.) y el betabel (*Beta vulgaris*), también se ha encontrado que los intercambios de semillas o la dispersión del polen son menos frecuentes a mayor distancia geográfica y, por consiguiente, los niveles de flujo génico disminuyen y se genera un patrón de aislamiento por distancia (Arnaud *et al.* 2009; Pusadee *et al.* 2009; Muraya *et al.* 2011).

A escala regional, las frecuencias alélicas de los SNPs con alta diferenciación genética del maíz se distribuyen en parches que en mayor medida corresponden a las provincias biogeográficas (Figura 7). Esta correspondencia podría esperarse ya que el índice F_{ST} a partir del cual se seleccionaron los SNPs se obtuvo entre estas regiones. Sin embargo, las razas se agrupan en complejos regionales que se localizan en áreas con interacciones entre el genotipo y el ambiente similares (Sánchez y Goodman 1992; Brush y Perales 2007). Una alternativa de agrupación espacial para este cultivo en México, distinta aunque similar a las provincias biogeográficas (CONABIO 1997), podría ser la regionalización biogeográfica basada en los modelos de distribución geográfica para cada raza que fue desarrollada por Perales y Golicher (Perales, comunicación personal; Figura 9).

El patrón espacial encontrado en el maíz, a partir de las provincias biogeográficas (CONABIO 1997) y posiblemente también a partir de la regionalización biogeográfica (Figura 9), sugiere que las presiones de selección natural debidas a factores abióticos y bióticos externos son semejantes regionalmente mientras que varían en mayor o menor grado en distintas regiones. En las razas mexicanas se ha encontrado que las diferencias en las condiciones ambientales son la principal fuerza detrás de los patrones generales de diversidad en el paisaje, sin embargo, también hay una influencia cultural (Sánchez *et al.* 2000; Pressoir y Berthaud 2004; Brush y Perales 2007). Algunos de los factores ambientales y culturales que posiblemente generan patrones de adaptación local en este cultivo podrían ser los niveles de temperatura y precipitación, el tipo de

suelo, la presencia o ausencia de ciertos insectos o patógenos y las técnicas de cultivo y de control de plagas (Mercer *et al.* 2008).



Figura 9. Regionalización biogeográfica del maíz de Perales y Golicher (Perales, comunicación personal), basada en los modelos de distribución geográfica para cada raza cultivada en México.

Genes candidatos a la selección natural positiva

La mayoría de los SNPs con alta diferenciación genética se localizan en genes y algunos de éstos podrían ser candidatos a la selección natural (positiva). Un ejemplo es el marcador PZE-110066819, el cual genera una sustitución no sinónima en el gen GRMZM2G107306. La proteína VIT1, sintetizada a partir de este gen, importa el hierro, un elemento esencial para procesos vitales, a las vacuolas de los embriones en las semillas (Kim *et al.* 2006). Además, el marcador PZE-110066819 es uno de los ocho que contribuyen más del 5% a explicar los patrones espaciales de estructura genética regional encontrados con el sPCA.

Otros dos genes (GRMZM2G079089, -317406), que tienen SNPs con alta diferenciación genética en exones (PZE-103172227, PUT-163a-78116163-4360, respectivamente), sintetizan proteínas que ya se han caracterizado. El gen GRMZM2G079089 sintetiza la proteína *trh1*. Las tiorredoxinas son proteínas de poca masa molecular, sintetizadas por una familia de genes, se localizan en el citoplasma y probablemente están involucradas en la activación de varias enzimas del citosol, ya que su función es deshacer los enlaces disulfuro de proteínas blanco (Ishiwatari *et al.* 1998). El gen GRMZM2G317406 sintetiza la proteína *mgs1*, la cual se expresa en el grano del polen y el tubo polínico (Stinson *et al.* 1987). Además, otros cinco genes (GRMZM2G171324, -157127, -359746, -125294, -012213), con SNPs que generan sustituciones no sinónimas (PZE-101015149, -101116836, -102137516, -108042154, SYN32227, respectivamente) codifican para proteínas que no se han caracterizado aunque cuentan con ortólogos sinténicos con el sorgo (*S. bicolor*; Andorf *et al.* 2010; www.maizegdb.org).

La mayoría de las proteínas sintetizadas a partir de genes que cuentan con SNPs con alta diferenciación genética no se han caracterizado, pero presentan ortólogos sinténicos con el sorgo (*S. bicolor*). Los genes ortólogos son aquellos que se originaron de un solo gen ancestral en el último ancestro común de los genomas comparados (Koonin 2005). En el caso del maíz y del sorgo, los progenitores de los genomas de estos cultivos se separaron evolutivamente hace alrededor de 11.9 millones de años (Swigonová *et al.* 2004). Debido a esta cercanía evolutiva, estos genomas pueden incluir bloques sinténicos conservados (conjunto de genes que mantienen el orden ancestral) constituidos por cientos o inclusive miles de genes (Dandekar *et al.* 1998; Koonin 2005). Los ortólogos usualmente tienen funciones equivalentes, sin embargo, es posible que proteínas no ortólogas ni homólogas (genes que comparten un origen común) lleven a cabo estas funciones en organismos diferentes (Koonin 2005). Además, los genes duplicados en el maíz, aunque se encuentren en regiones ortólogas, no necesariamente son ortólogos (Swigonová *et al.* 2004). Por lo tanto, la mayoría de los genes codifican proteínas cuya función aún no se conoce con certeza.

Adaptación local

La búsqueda de SNPs candidatos a la adaptación local con dos métodos distintos (extremos de latitud y altitud y distribuciones geográficas contrastantes) mostró que dos marcadores con alta

diferenciación genética se comparten entre agrupaciones de razas en distintos ambientes. El marcador SYN36384 genera una sustitución no sinónima en el gen GRMZM2G458824, el cual codifica para una proteína que aún no se ha caracterizado ni cuenta con un ortólogo sinténico con el sorgo (*S. bicolor*; Andorf *et al.* 2010; www.maizegdb.org). En cambio, el marcador PZE-105063649 genera una sustitución sinónima en el gen GRMZM2G088737, el cual codifica para una proteína que tampoco se ha caracterizado pero el ARNm que participa en la transcripción está registrado en el TSA (www.ncbi.nlm.nih.gov/genbank/tsa). Además, este marcador también tiene alta diferenciación genética entre provincias biogeográficas y es uno de los ocho que contribuyen más del 5% a explicar los patrones espaciales de estructura genética regional encontrados con el sPCA.

El escaso número de SNPs compartidos entre agrupaciones de razas en distintos ambientes no significa que el proceso de selección natural divergente que genera los patrones de adaptación local no esté operando en el maíz. Este patrón podría no detectarse debido a la presencia de genotipos generalistas favorecidos por altos niveles de flujo génico y por variación temporal en la selección (Kawecki y Ebert 2004), sin embargo, esto posiblemente no ocurre en este cultivo. En las comunidades mexicanas, la integración de las semillas al germoplasma local depende no sólo de la introducción sino también de las prácticas de manejo y ambos procesos varían regionalmente (Dyer y Taylor 2008). La tasa de reemplazo con semillas mejoradas o nativas externas es mayor, en las comunidades del noroeste, noreste y oeste de México, y menor, en las comunidades del centro y sudeste del país, que la tasa de difusión de las semillas nativas (Louette *et al.* 1997; Perales *et al.* 2003b; Pressoir y Berthaud 2004; Dyer y Taylor 2008), lo cual posiblemente genera distintos patrones de flujo génico regional. En cambio, la selección anual de las semillas por parte de los agricultores tradicionales, un evento en general subjetivo, posiblemente promueve un proceso estabilizador que mantiene la diversidad genética y los caracteres particulares de las razas en ciertas poblaciones nativas (Perales *et al.* 2003b; Dyer y Taylor 2008).

Las presiones estables y consistentes de selección con bajos niveles de flujo génico a nivel regional promueven la selección divergente que genera la adaptación local (Bellon *et al.* 2011). Un claro ejemplo de adaptación local en el maíz son las razas del complejo Cónico, distribuidas mayormente en la Faja Volcánica, y las del complejo Sierra de Chihuahua, localizadas en la Sierra Madre Occidental (Sánchez y Goodman 1992; CONABIO 2010). La adecuación de estas

razas disminuye en altitudes menores a los 2,000 msnm y el cultivo de las variedades mejoradas no se ha extendido en altitudes mayores a este límite, debido posiblemente a que su producción disminuye como resultado de interacciones entre el genotipo y el ambiente más intensas (Perales *et al.* 2003a; Mercer *et al.* 2008).

Otra razón de que no se haya detectado un patrón genético consistente de adaptación local entre las razas de maíz en distintos ambientes podría ser que distintas poblaciones nativas han alcanzado picos adaptativos de diferente altura, por ejemplo, si han evolucionado combinaciones alternativas de genes epistáticos que dan como resultado distinta adecuación (Kawecki y Ebert 2004). La adaptación a un nuevo ambiente podría llevarse a cabo a partir de ligeros cambios en las frecuencias alélicas de múltiples SNPs en el genoma localizados en distintos genes (adaptación poligénica; Pritchard y Di Rienzo 2010) y las combinaciones de genes, que generan caracteres fenotípicos similares en ambientes semejantes, podrían ser diferentes (evolución convergente; Arendt y Reznick 2008).

Las razas de maíz se han adaptado a distintos ambientes en un corto tiempo, de alrededor de 9,000 años, desde su domesticación (Matsuoka *et al.* 2002a). La adaptación a partir de la variación presente en el genoma es más rápida que a partir de una nueva mutación, debido a que los alelos potencialmente ventajosos ya están presentes en el momento en que ocurre el cambio ambiental (Barrett y Schluter 2007). El genoma del maíz presenta una amplia variación estructural que promueve los procesos de adaptación a partir de múltiples copias de alelos antiguos (Chia *et al.* 2012). La selección posiblemente ha puesto a prueba estos alelos en el pasado y, éstos últimos, podrían representar distintos cambios genéticos potencialmente ventajosos para un mismo ambiente (Barrett y Schluter 2007). Debido a esto, es posible que en el maíz el principal mecanismo de adaptación local sea mediante el incremento en las frecuencias de alelos ya presentes en el genoma que a partir de nuevas mutaciones (van Heerwaarden *et al.* 2012). Por ejemplo, en las líneas mejoradas endogámicas, que son aún más recientes que las razas, se han detectado genes candidatos que cuentan con patrones de ancestría y de longitud de haplotipo similares al resto del genoma y, en consecuencia, posiblemente no han ocurrido cambios bruscos en las frecuencias alélicas de los loci neutrales asociados (van Heerwaarden *et al.* 2012).

Es posible que las poblaciones nativas de distintas razas de maíz en ambientes similares hayan desarrollado caracteres adaptativos semejantes a partir de genes candidatos diferentes debido a que los alelos ventajosos podrían no dispersarse entre estas regiones. En los cereales, los caracteres asociados al síndrome de domesticación, como la dehiscencia de los granos, suelen estar regulados por conjuntos diferentes de genes en el genoma (Li y Gill 2006). En contraste, los caracteres asociados a la diversificación posterior a la domesticación, como la ausencia o casi ausencia de la amilasa del almidón en el endospermo, y aquellos asociados al mejoramiento, como el color del grano y la espiga, suelen estar regulados por un mismo conjunto de genes (Gross y Olsen 2010). Un caso en el que se ha estudiado la evolución convergente en una especie con amplia distribución geográfica es el del ser humano. Entre las poblaciones de distintos continentes, distintos genes o conjuntos de genes determinan la piel clara en los europeos y los asiáticos orientales o la persistencia de la lactasa en los adultos europeos y africanos (Tishkoff *et al.* 2006; Norton *et al.* 2007; Coop *et al.* 2009).

CONCLUSIONES

La selección divergente, la principal fuerza evolutiva que promueve la adaptación local, posiblemente opera en el maíz mayormente mediante el incremento en las frecuencias de alelos ya presentes en el genoma que a partir de nuevas mutaciones. El reducido número de SNPs con alta diferenciación genética compartidos entre agrupaciones de razas en distintos ambientes sugiere que posiblemente distintas combinaciones de genes han generado caracteres adaptativos similares. No obstante, las razas cultivadas a más de 2,000 msnm muestran evidencias de adaptación local ya que son sensibles a las condiciones ambientales características de altitudes menores y que incluyen cambios en temperatura y precipitación. Sin embargo, el mecanismo genético que regula la adaptación local en el maíz aún no se ha descrito.

El maíz cultivado en México presenta patrones espaciales de estructura genética regional posiblemente determinados por la selección natural (positiva). Algunos de los genes que tienen SNPs con alta diferenciación genética podrían ser candidatos a esta selección o localizarse en loci cuya variación está influida por el efecto de los barridos selectivos. En particular, dos genes (GRMZM2G107306 y -088737) probablemente son candidatos debido a que cuentan con SNPs (PZE-110066819 y -105063649, respectivamente) que ya habían sido reportados en la literatura o que tuvieron alta diferenciación genética con dos métodos en este estudio. Ambos marcadores son informativos para la estructura genética espacial encontrada. Además, la proteína VIT1, sintetizada por el gen GRMZM2G107306, tiene una función posiblemente clave en el desarrollo del embrión, ya que importa el hierro, un elemento esencial para procesos vitales, a las vacuolas de los embriones en las semillas.

Las condiciones ambientales, que incluyen a factores abióticos y bióticos externos, posiblemente sean la principal fuerza que determina la diversidad en el maíz, sin embargo, también hay una influencia cultural. El patrón de aislamiento por distancia encontrado posiblemente se debe a que los intercambios de semillas suelen llevarse a cabo entre comunidades cercanas. Además, la influencia de la localización geográfica detectada y el patrón espacial en parches de las frecuencias alélicas obtenido sugieren que las presiones de selección varían regionalmente. Algunos de los factores ambientales y culturales que posiblemente generan patrones de adaptación local en este cultivo podrían ser los niveles de temperatura y

precipitación, el tipo de suelo, la presencia o ausencia de ciertos insectos o patógenos y las técnicas de cultivo y de control de plagas.

LITERATURA CITADA

- Altschul S. F., T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller, D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* 25: 3389-3402.
- Ananiev E. V., R. L. Phillips, H. W. Rines. 1998. A knob-associated tandem repeat in maize capable of forming fold-back DNA segments: are chromosome knobs megatransposons? *Proceedings of the National Academy of Science* 95: 10785-10790.
- Anderson J. T., C.-R. Lee, C. A. Rushworth, R. I. Colautti, T. Mitchell-Olds. 2013. Genetic trade-offs and conditional neutrality contribute to local adaptation. *Molecular Ecology* 22: 699-708.
- Andorf C. M., C. J. Lawrence, L. C. Harper, M. L. Schaeffer, D. A. Campbell, T. Z. Sen. 2010. The Locus Lookup tool at MaizeGDB: identification of genomic regions in maize by integrating sequence information with physical and genetic maps. *Bioinformatics* 26: 434-436.
- Arendt J., D. Reznick. 2008. Convergence and parallelism reconsidered: what have we learned about the genetics of adaptation? *Trends in Ecology and Evolution* 23: 26-32.
- Arnaud J.-F., S. Fénart, C. Godé, S. Deledicque, P. Touzet, J. Cuguen. 2009. Fine-scale geographical structure of genetic diversity in inland wild beet populations. *Molecular Ecology* 18: 3201-3215.
- Baldor J. A. 2003. Geometría plana y del espacio. Introducción a la trigonometría. 19° Ed. Publicaciones Cultural. México.
- Barrett R. D. H., D. Schluter. 2007. Adaptation from standing genetic variation. *Trends in Ecology and Evolution* 23: 38-44.
- Beaumont M. A. 2005. Adaptation and speciation: what can F_{ST} tell us? *Trends in Ecology and Evolution* 20: 435-440.
- Bellon M. R., D. Hodson, J. Hellin. 2011. Assessing the vulnerability of traditional maize seed systems in Mexico to climate change. *Proceedings of the National Academy of Science* 108: 13432-13437.

- Bradbury P. J., Z. Zhang, D. E. Kroon, T. M. Casstevens, Y. Ramdoss, E. S. Buckler. 2007. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23: 2633-2635.
- Brush S. B., H. R. Perales. 2007. A maize landscape: ethnicity and agro-biodiversity in Chiapas Mexico. *Agriculture, Ecosystems and Environment* 121: 211-221.
- Chia J.-M., C. Song, P. J. Bradbury, D. Costich, N. de Leon, J. Doebley, R. J. Elshire, B. Gaut, L. Geller, J. C. Glaubitz, M. Gore, K. E. Guill, J. Holland, M. B. Hufford, J. Lai, M. Li, X. Liu, Y. Lu, R. McCombie, R. Nelson, J. Poland, B. M. Prasanna, T. Pyhäjärvi, T. Rong, R. S. Sekhon, Q. Sun, M. I. Tenailon, F. Tian, J. Wang, X. Xu, Z. Zhang, S. M. Kaeppler, J. Ross-Ibarra, M. D. McMullen, E. S. Buckler, G. Zhang, Y. Xu, D. Ware. 2012. Maize HapMap2 identifies extant variation from a genome in flux. *Nature Genetics* 44: 803-807.
- CONABIO. 2010. Base de datos del proyecto global de maíces nativos “Recopilación, generación, actualización y análisis de información acerca de la diversidad genética de maíces y sus parientes silvestres en México”. Octubre de 2010. Comisión Nacional para el Conocimiento y Uso de la Biodiversidad. México D. F.
- CONABIO. 1997. Provincias biogeográficas de México. Comisión Nacional para el Conocimiento y Uso de la Biodiversidad. Escala: 1:4000000.
- Coop G., J. K. Pickrell, J. Novembre, S. Kudaravalli, J. Li, D. Absher, R. M. Myers, L. L. Cavalli-Sforza, M. W. Feldman, J. K. Pritchard. 2009. The role of geography in human adaptation. *PLoS Genetics* 5: e1000500. doi:10.1371/journal.pgen.1000500.
- Dandekar T., B. Snel, M. Huynen, P. Bork. 1998. Conservation of gene order: a fingerprint of proteins that physically interact. *Trends in Biochemical Sciences* 23: 324-328.
- Diamond J. 2002. Evolution, consequences, and future of plant and animal domestication. *Nature* 418: 700-707.
- Díez C. M., B. S. Gaut, E. Meca, E. Scheinvar, S. Montes-Hernández, L. E. Eguiarte, M. I. Tenailon. 2013. Genome size variation in wild and cultivated maize along altitudinal gradients. *New Phytologist*: doi: 10.1111/nph.12247.
- Doebley J. F., M. M. Goodman y C. W. Stuber. 1985. Isozyme variation in the races of maize from Mexico. *American Journal of Botany* 72: 629-639.
- Doebley J. F., M. M. Goodman y C. W. Stuber. 1984. Isoenzymatic variation in *Zea* (Gramineae). *Systematic Botany* 9: 203-218.

- Doyle J. J., J. L. Doyle. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19: 11-15.
- Dyer G. A., J. E. Taylor. 2008. A crop population perspective on maize seed systems in Mexico. *Proceedings of the National Academy of Science* 105: 470-475.
- Dyer R. J. 2009. GeneticStudio: a suite of programs for spatial analysis of genetic-marker data. *Molecular Ecology Resources* 9: 110-113.
- Dyer R. J., R. D. Westfall, V. L. Sork, P. E. Smouse. 2004. Two-generation analysis of pollen flow across a landscape V: a stepwise approach for extracting factors contributing to pollen structure. *Heredity* 92: 204-211.
- Evanno G., S. Regnaut y J. Goudet. 2005. Detecting the number of clusters of individuals using the software Structure: A simulation study. *Molecular Ecology* 14: 2611-2620.
- Excoffier L., P. E. Smouse, J. M. Quattro. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131: 479-491.
- Flint-Garcia S. A., J. M. Thornsberry, E. S. Buckler. 2003. Structure of linkage disequilibrium in plants. *Annual Review of Plant Biology* 54: 357-374.
- François O., E. Durand. 2010. Spatially explicit Bayesian clustering models in population genetics. *Molecular Ecology Resources* 10: 773-784.
- Fukunaga K., J. Hill, Y. Vigouroux, Y. Matsuoka, J. Sánchez G., K. Liu, E. S. Buckler, J. Doebley. 2005. Genetic diversity and population structure of teosinte. *Genetics* 169: 2241-2254.
- Futuyma D. J. 2013. *Evolution*. 3a Ed. Sinauer Associates, Inc. Estados Unidos de América.
- Ganal M. W., G. Durstewitz, A. Polley, A. Bérard, E. S. Buckler, A. Charcosset, J. D. Clarke, E.-M. Graner, M. Hansen, J. Joets, M.-C. Le Paslier, M. D. McMullen, P. Montalent, M. Rose, C.-C. Shön, Q. Sun, H. Walter, O. C. Martin, M. Falque. 2011. A large maize (*Zea mays* L.) SNP genotyping array: development and germplasm genotyping, and genetic mapping to compare with the B73 reference genome. *PLoS ONE* 6: e28334. doi:10.1371/journal.pone.0028334.
- Gross B. L., K. M. Olsen. 2010. Genetic perspectives on crop domestication. *Trends in Plant Science* 15: 529-537.

- Hamblin M. T., E. S. Buckler, J.-L. Jannink. 2011. Population genetics of genomics-based crop improvement methods. *Trends in Genetics* 27: 98-106.
- Hamblin M. T., M. L. Warburton, E. S. Buckler. 2007. Empirical comparison of simple sequence repeats and single nucleotide polymorphisms in assessment of maize diversity and relatedness. *PLoS ONE* 12: e1367. doi:10.1371/journal.pone.0001367.t001.
- Hancock J. F. 2004. *Plant evolution and the origin of crop species*. 2^a Ed. Publicaciones CABI. Cambridge, Massachusetts, Estados Unidos de América.
- Harlan J. R. 1992. *Crops and man*. 2^a Ed. American Society of Agronomy, Crop Science Society. Madison, Wisconsin, Estados Unidos de América.
- Hernández X. E., G. Alanís F. 1970. Estudio morfológico de cinco nuevas razas de maíz de la Sierra Madre Occidental de México: Implicaciones citogenéticas y filogenéticas. *Agrociencia* 5: 3-30.
- Hufford M. B., P. Lubinsky, T. Pyhäjärvi, M. T. Devengenzo, N. C. Ellstrand, J. Ross-Ibarra. 2013. The genomic signature of crop-wild introgression in maize. *PLoS Genetics* 9: e1003477. doi:10.1371/journal.pgen.1003477.
- Hufford M. B., X. Xu, J. van Heerwaarden, T. Pyhäjärvi, J. M. Chia, R. A. Cartwright, R. J. Elshire, J. C. Glaubitz, K. E. Guill, S. M. Kaeppler, J. Lai, P. L. Morrell, L. M. Shannon, C. Song, N. M. Springer, R. A. Swanson-Wagner, P. Tiffin, J. Wang, G. Zhang, J. Doebley, M. D. McMullen, D. Ware, E. S. Buckler, S. Yang, J. Ross-Ibarra. 2012a. Comparative population genomics of maize domestication and improvement. *Nature Genetics* 44: 808-811.
- Hufford M. B., E. Martínez-Meyer, B. S. Gaut, L. E. Eguiarte, M. I. Tenailon. 2012b. Inferences from the historical distribution of wild and domesticated maize provide ecological and evolutionary insight. *PLoS ONE* 7: e47659. doi:10.1371/journal.pone.0047659.
- Hung H.-Y., L. M. Shannon, F. Tian, P. J. Bradbury, C. Chen, S. A. Flint-Garcia, M. D. McMullen, D. Ware, E. S. Buckler, J. F. Doebley, J. B. Holland. 2012. ZmCTT and the genetic basis of day-length adaptation underlying the postdomestication spread of maize. *Proceedings of the National Academy of Science* 109: E1913-E1921.
- Ishiwatari Y., T. Fujiwara, K. C. McFarland, K. Nemoto, H. Hayashi, M. Chino, W. J. Lucas. 1998. Rice phloem thioredoxin h has the capacity to mediate its own cell-to-cell transport through plasmodesmata. *Planta* 205: 12-22.

- Jombart T., S. Devillard, A.-B. Dufour, D. Pontier. 2008. Revealing cryptic spatial patterns in genetic variability by a new multivariate method. *Heredity* 101: 92-103.
- Jones E. S., H. Sullivan, D. Bhatramakki, J. S. C. Smith. 2007. A comparison of simple sequence repeat and single nucleotide polymorphism marker technologies for the genotypic analysis of maize (*Zea mays* L.). *Theoretical and Applied Genetics* 115: 361-371.
- Kato T. A., C. Mapes, L. M. Mera, J. A. Serratos, R. A. Bye. 2009. Origen y diversificación del maíz: una revisión analítica. Universidad Nacional Autónoma de México, Comisión Nacional para el Conocimiento y Uso de la Biodiversidad. México.
- Koonin E. V. 2005. Orthologs, paralogs, and evolutionary genomics. *Annual Review of Genetics* 39: 309-338.
- Kawecki T. J., D. Ebert. 2004. Conceptual issues in local adaptation. *Ecology Letters* 7: 1225-1241.
- Kim S. A., T. Punshon, A. Lanzirotti, L. Li, J. A. Alonso, J. R. Ecker, J. Kaplan, M. L. Guerinot. 2006. Localization of iron in *Arabidopsis* seed requires the vacuolar membrane transporter VIT1. *Science* 314: 1295-1298.
- Landguth E. L., N. Balkenhol. 2012. Relative sensitivity of neutral versus adaptive genetic data for assessing population differentiation. *Conservation Genetics* 13: 1421-1426.
- Legendre P., M.-J. Fortin. 2010. Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data. *Molecular Ecology Resources* 10: 831-844.
- Lewontin R. C., J. Krakauer. 1973. Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics* 74: 175-195.
- Li W., B. Gill. 2006. Multiple genetic pathways for seed shattering in the grasses. *Functional and Integrative Genomics* 6: 300-309.
- Louette D., A. Charrier, J. Berthaud. 1997. *In situ* conservation of maize in Mexico: genetic diversity and maize seed management in a traditional community. *Economic Botany* 51: 20-38.
- Mantel N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Research* 27: 209-220.

- Matsuoka Y., Y. Vigouroux, M. M. Goodman, J. J. Sánchez G., E. Buckler, J. Doebley. 2002a. A single domestication for maize shown by multilocus microsatellite genotyping. *Proceedings of the National Academy of Science* 99: 6080-6084.
- Matsuoka Y., S. E. Mitchell, S. Kresovich, M. Goodman, J. Doebley. 2002b. Microsatellites in *Zea* – variability, patterns of mutations, and use for evolutionary studies. *Theoretical and Applied Genetics* 104: 436-450.
- Mercer K., A. Martínez-Vásquez, H. R. Perales. 2008. Asymmetrical local adaptation of maize landraces along an altitudinal gradient. *Evolutionary Applications* 1: 489-500.
- Moran P. 1950. Notes on continuous stochastic phenomena. *Biometrika* 37: 17-23.
- Moran P. 1948. The interpretation of statistical maps. *Journal of the Royal Statistical Society, Series B* 10: 243-251.
- Morin P. A., G. Luikart, R. K. Wayne, SNP workshop group. 2004. SNPs in ecology, evolution and conservation. *Trends in Ecology and Evolution* 19: 208-216.
- Morrell P. L., E. S. Buckler, J. Ross-Ibarra. 2012. Crop genomics: advances and applications. *Nature Reviews Genetics* 13: 85-96.
- Morrone J. J. 2006. Biogeographic areas and transition zones of Latin America and the Caribbean Islands based on panbiogeographic and cladistics analyses of the entomofauna. *Annual Review of Entomology* 51: 467-494.
- Muraya M. M., S. de Villiers, H. K. Parzies, E. Mutegi, F. Sagnard, B. M. Kanyenji, D. Kiambi, H. H. Geiger. 2011. Genetic structure and diversity of wild sorghum populations (*Sorghum* spp.) from different eco-geographical regions of Kenya. *Theoretical and Applied Genetics* 123: 571-583.
- Nei M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89: 583-590.
- Nei M. 1972. Genetic distance between populations. *The American Naturalist* 106: 283-292.
- Nielsen R. 2005. Molecular signatures of natural selection. *Annual Reviews of Genetics* 39: 197-218.
- Norton H. L., R. A. Kittles, E. Parra, P. McKeigue, X. Mao, K. Cheng, V. A. Canfield, D. G. Bradley, B. McEvoy, M. D. Shriver. 2007. Genetic evidence for the convergent evolution of light skin in Europeans and East Asians. *Molecular Biology and Evolution* 24: 710-722.

- Ortega Paczka, R. 2003. La diversidad del maíz en México. pp. 123-154. In: Esteva G., C. Marielle (coords.). Sin maíz no hay país. Consejo Nacional para la Cultura y las Artes, Dirección General de Culturas Populares e Indígenas, Museo Nacional de Culturas Populares. México.
- Peakall R., P. E. Smouse. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research – an update. *Bioinformatics* 28: 2537-2539.
- Peakall R., P. E. Smouse. 2006. GenAlEx 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* 6: 288-295.
- Perales R. H., S. B. Brush, C. O. Qualset. 2003a. Landraces of maize in central Mexico: an altitudinal transect. *Economic Botany* 57: 7-20.
- Perales R. H., S. B. Brush, C. O. Qualset. 2003b. Dynamic management of maize landraces in central Mexico. *Economic Botany* 57: 21-34.
- Piperno D. R., A. J. Ranere, I. Holst, J. Iriarte y R. Dickau. 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 Science* 106: 5019-5024.
- Piperno D. R., K. V. Flannery. 2001. The earliest archaeological maize (*Zea mays* L.) from highland Mexico: New accelerator mass spectrometry dates and their implications. *Proceedings of the National Academy of Science* 98: 2101-2103.
- Pohl M. E. D., D. R. Piperno, K. O. Pope, J. G. Jones. 2007. Microfossil evidence for pre-Columbian maize dispersals in the neotropics from San Andrés, Tabasco, Mexico. *Proceedings of the National Academy of Science* 104: 6870-6875.
- Pressoir G., J. Berthaud. 2004. Patterns of population structure in maize landraces from the Central Valleys of Oaxaca in Mexico. *Heredity* 92: 88-94.
- Pritchard J., A. Di Rienzo. 2010. Adaptation – not by sweeps alone. *Nature Reviews Genetics* 11: 665-667.
- Pritchard J. K., M. Stephens y P. Donnelly. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155: 945-959.
- Purcell S., B. Neale, K. Todd-Brown, L. Thomas, M. A. R. Ferreira, D. Bender, J. Maller, P. Sklar, P. I. W. de Bakker, M. J. Daly, P. C. Sham. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics* 81: 559-575.

- Pusadee T., S. Jamjod, Y.-C. Chiang, B. Rerkasem, B. A. Schaal. 2009. Genetic structure and isolation by distance in a landrace of Thai rice. *Proceedings of the National Academy of Science* 106: 13880-13885.
- Pyhäjärvi T., M. B. Hufford, S. Mezouk, J. Ross-Ibarra. 2012. Complex patterns of local adaptation in teosinte. arXiv:1208.0634v1 (<http://arxiv.org/abs/1208.0634>).
- Pyhäjärvi T., S. T. Kujala, O. Savolainen. 2011. Revisiting protein heterozygosity in plants – nucleotide diversity in allozyme coding genes of conifer *Pinus sylvestris*. *Tree Genetics and Genomes* 7: 385-397.
- Ranere A. J., D. R. Piperno, I. Holst, R. Dickau, J. Iriarte. 2009. The cultural and chronological context of early Holocene maize and squash domestication in the Central Balsas River Valley, Mexico. *Proceedings of the National Academy of Science* 106: 5014-5018.
- R Development Core Team. 2008. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- Reif J. C., M. L. Warburton, X. C. Xia, D. A. Hoisington, J. Crossa, S. Taba, J. Muminovic, M. Bohn, M. Frisch, A. E. Melchinger. 2006. Grouping of accessions of Mexican races of maize revisited with SSR markers. *Theoretical and Applied Genetics* 113: 177-185.
- Robinson S. J., M. D. Samuel, D. L. Lopez, P. Shelton. 2012. The walk is never random: subtle landscape effects shape gene flow in a continuous white-tailed deer population in the Midwestern United States. *Molecular Ecology* 21: 4190-4205.
- Ruiz C. J. A., N. Durán P., J. J. Sánchez G., J. Ron P., D. R. González E., J. B. Holland, G. Medina G. 2008. Climatic adaptation and ecological descriptors of 42 Mexican maize races. *Crop Science* 48: 1502-1512.
- Sánchez G. J. J., M. M. Goodman y C. W. Stuber. 2000. Isozymatic and morphological diversity in the races of maize of Mexico. *Economic Botany* 54: 43-59.
- Sánchez G. J. J., M. M. Goodman. 1992. Relationships among the Mexican races of maize. *Economic Botany* 46: 72-85.
- SanMiguel P., A. Tikhonov, Y. K. Jin, N. Motchoulskaia, D. Zakbarov, A. Melake-Berhan, P. S. Springer, K. J. Edwards, M. Lee, Z. Avramova, J. L. Bennetzen. 1996. Nested retrotransposons in the intergenic regions of the maize genome. *Science* 274: 765-768.
- Schnable *et al.* 2009. The B73 maize genome: complexity, diversity, and dynamics. *Science* 326: 1112-1115.

- Schlötterer C. 2004. The evolution of molecular markers – just a matter of fashion? *Nature Reviews Genetics* 5: 63-69.
- Schoville S. D., A. Bonin, O. François, S. Lobreaux, C. Melodelima, S. Manel. 2012. Adaptive genetic variation on the landscape: methods and cases. *Annual Review of Ecology, Evolution and Systematics* 43: 23-43.
- Siol M., S. I. Wright, S. C. H. Barrett. 2010. The population genomics of plant adaptation. *New Phytologist* 188: 313-332.
- Smouse P. E., R. Peakall, E. Gonzalez. 2008. A heterogeneity test for fine-scale genetic structure. *Molecular Ecology* 17: 3389-3400.
- Stinson J. R., A. J. Eisenberg, R. P. Willing, M. E. Pe, D. D. Hanson, J. P. Mascarenhas. 1987. Genes expressed in the male gametophyte of flowering plants and their isolation. *Plant Physiology* 83: 442-447.
- Swigonová Z., J. Lai, J. Ma, W. Ramakrishna, V. Llaca, J. L. Bennetzen, J. Messing. 2004. Close split of sorghum and maize genome progenitors. *Genome Research* 14: 1916-1923.
- Szpiech Z. A., M. Jakobsson, N. A. Rosenberg. 2008. ADZE: a rarefaction approach for counting alleles private to combinations of populations. *Bioinformatics* 24: 2498-2504.
- Tenaillon M. I., J. U'Ren, O. Tenaillon, B. S. Gaut. 2004. Selection versus demography: a multilocus investigation of the domestication process in maize. *Molecular Biology and Evolution* 21: 1214-1225.
- Tishkoff S. A., F. A. Reed, A. Ranciaro, B. F. Voight, C. C. Babbitt, J. S. Silverman, K. Powell, H. M. Mortensen, J. B. Hirbo, M. Osman, M. Ibrahim, S. A. Omar, G. Lema, T. B. Nyambo, J. Ghoris, S. Bumpstead, J. K. Pritchard, G. A. Wray, P. Deloukas. 2007. Convergent adaptation of human lactase persistence in Africa and Europe. *Nature Genetics* 39: 31-40.
- Ureta C., E. Martínez-Meyer, H. R. Perales, E. R. Álvarez-Buylla. 2012. Projecting the effects of climate change on the distribution of maize races and their wild relatives in Mexico. *Global Change Biology* 18: 1073-1082.
- Van Heerwaarden J., M. B. Hufford, J. Ross-Ibarra. 2012. Historical genomics of North American maize. *Proceedings of the National Academy of Science* 109: 12420-12425.

- Van Heerwaarden J., J. Doebley, W. H. Briggs, J. C. Glaubitz, M. M. Goodman, J. J. Sánchez G., J. Ross-Ibarra. 2011. Genetic signals of origin, spread, and introgression in a large sample of maize landraces. *Proceedings of the National Academy of Science* 108: 1088-1092.
- Vazquez-Lobo Y. A. 1996. Filogenia de hongos endófitos del género *Pinus* L. implementación de técnicas moleculares y resultados preliminares. Tesis de Licenciatura. Universidad Nacional Autónoma de México.
- Vielle-Calzada J.-P., J. Padilla. 2009. The Mexican landraces: description, classification, and diversity. Pp.:543-561. In: Bennetzen J. L. y S. C. Hake (eds.). *Handbook of maize: its biology*. Springer Science, Business Media. Nueva York, Estados Unidos de América.
- Vigouroux Y., J. C. Glaubitz, Y. Matsuoka, M. M. Goodman, J. Sánchez G., J. Doebley. 2008. Population structure and genetic diversity of New World maize races assessed by DNA microsatellites. *American Journal of Botany* 95: 1240-1253.
- Williams G. C. 1966. *Adaptation and natural selection*. Princeton University Press. Nueva Jersey, Estados Unidos de América.
- Wright S. I., I. V. Bi, S. G. Schroeder, M. Yamasaki, J. F. Doebley, M. D. McMullen, B. S. Gaut. 2005. The effects of artificial selection on the maize genome. *Science* 308: 1310-1314.
- Wright S. 1951. The genetical structure of populations. *Annals of Eugenics* 15: 323-354.
- Yamasaki M., M. I. Tenailon, I. V. Bi, S. G. Schroeder, H. Sanchez-Villeda, J. F. Doebley, B. S. Gaut, M. D. McMullen. 2005. A large-scale screen for artificial selection in maize identifies candidate agronomic loci for domestication and crop improvement. *The Plant Cell* 17: 2859-2872.
- Yan J., T. Shah, M. L. Warburton, E. S. Buckler, M. D. McMullen, J. Crouch. 2009. Genetic characterization and linkage disequilibrium estimation of a global maize collection using SNP markers. *PLoS ONE* 4: e8451. doi:10.1371/journal.pone.0008451.
- Zizumbo-Villarreal D., P. Colunga-GarcíaMarín. Origin of agriculture and plant domestication in West Mesoamerica. *Genetic Resources and Crop Evolution* 57: 813-825.

ANEXOS

ANEXO I. Protocolo CTAB modificado para la extracción de ADN (Doyle y Doyle 1987; Vázquez-Lobo 1996).

Sustancias por muestra

2 ml de nitrógeno líquido
2 ml de buffer CTAB 2X
2 μ l de β -mercaptoetanol
4 μ l de RNAsa
600 μ l de cloroformo-octanol 24:1
400 μ l de isopropanol frío
1 ml de etanol 70% frío
200 μ l de agua desionizada

Material

mortero con pistilo estéril por muestra
2 tubos de 1.5 ml por muestra
centrifugadora de 13,000 rpm a 4°C
incubadora a 60°C
refrigerador a -20°C

Procedimiento

1. En un mortero frío moler de 30 a 100 mg de tejido congelado con nitrógeno líquido.
2. Cuando se evapore el nitrógeno líquido y el tejido sea un polvo fino agregar inmediatamente 1 ml de buffer CTAB 2X y 4 μ l de RNAsa. Revolver la muestra hasta que licue.
3. Verter la muestra en un tubo de 1.5 ml y centrifugar a 13,000 rpm durante 10 min a temperatura ambiente.
4. Eliminar el sobrenadante e incorporar la muestra en 600 μ l de buffer CTAB 2X. Incubarla en un baño María a 60°C por 10 min y agitarla repetidamente durante este periodo.
5. Agregar 600 μ l de cloroformo-octanol 24:1, agitar hasta homogeneizar y centrifugar a 13,000 rpm durante 10 min a 4°C (o hasta que el sobrenadante quede transparente).
6. Trasladar el sobrenadante a un nuevo tubo de 1.5 ml (si el sobrenadante está muy turbio se pueden repetir los pasos 5 y 6)

7. Agregar 400 μ l de isopropanol frío e inmediatamente mezclar volteando suavemente el tubo. Incubar de 30 min a 24 h a -20°C .
8. Centrifugar a 13,000 rpm durante 10 min a 4°C y eliminar el sobrenadante teniendo cuidado de no tirar el pellet de ADN.
9. Agregar 500 μ l de etanol 70% frío y centrifugar a 13,000 rpm durante 2 min a 4°C . Eliminar el sobrenadante teniendo cuidado de no tirar el pellet de ADN y repetir este paso.
10. Secar en una campana hasta que no queden rastros de etanol en el tubo. Incorporar el pellet de ADN en 200 μ l de ddH₂O estéril (la cantidad puede variar dependiendo del tamaño del pellet) y almacenar a -20°C .

ANEXO II

Anexo II.A. Localidades de los 169 individuos de 50 razas analizados con SNPs. Las provincias biogeográficas corresponden a CONABIO 1997, sin embargo, las de la Costa del Pacífico y del Golfo de México se subdividieron en Norte y Sur.

Muestra	No. colecta	Raza	Localidad	Provincia biogeográfica	Altitud	Latitud	Longitud
27	2009-232	Ancho	Carretera San Juan Tepeoculco, Atlautla, Estado de México	Faja Volcánica	2226	18.998	-98.779
29	2009-235	Ancho	Carretera San Juan Tepeoculco, Atlautla, Estado de México	Faja Volcánica	2226	18.998	-98.779
30	2009-236	Ancho	San Andres Tlalamac, Atlautla, Estado de México	Faja Volcánica	2073	18.968	-98.809
1	2	Apachito	Tres Ojitos, Madera, Chihuahua	Sierra Madre Occidental	2106	29.113	-108.080
3	16	Apachito	Santo Tomas, Guerrero, Chihuahua	Sierra Madre Occidental	1975	28.686	-107.583
5	717	Arrocillo	Eloxotitlán, Zacatlán, Puebla	Sierra Madre Oriental	2159	19.934	-97.981
97	731	Arrocillo	Foxcoatzingo, Zacatlán, Puebla	Sierra Madre Oriental	2104	19.956	-97.979
146	734	Arrocillo	Eloxotitlán, Zacatlán, Puebla	Sierra Madre Oriental	2131	19.943	-97.980
185	732	Arrocillo	Foxcoatzingo, Zacatlán, Puebla	Sierra Madre Oriental	2104	19.956	-97.979
6	415	Azul	Las Varas, Temosachic, Chihuahua	Sierra Madre Occidental	2215	28.455	-108.015
8	416	Azul	Las Varas, Temosachic, Chihuahua	Sierra Madre Occidental	2220	28.457	-108.016
100	SIN-2007-087	Blando de Sonora	Chinobampo, El Fuerte, Sinaloa	Costa del Pacífico (Norte)	299	26.395	-108.383
10	38	Bofo	Cuiteco, Urique, Chihuahua	Sierra Madre Occidental	1734	27.435	-108.008

Muestra	No. colecta	Raza	Localidad	Provincia biogeográfica	Altitud	Latitud	Longitud
31	2009-041	Cacahuacintle	San Pedro Tlaltizapán, Santiago Tianguistenco, Estado de México	Faja Volcánica	2585	19.196	-99.509
32	2009-049	Cacahuacintle	Calimaya, Calimaya, Estado de México	Faja Volcánica	2700	19.168	-99.618
63	41	Cacahuacintle	San Pablo Zitlaltepec, Trinidad Sánchez Santos, Tlaxcala	Faja Volcánica	2554	19.222	-97.901
64	159	Cacahuacintle	5 km E Atotonilco, Tlaxco, Tlaxcala	Faja Volcánica	2745	19.600	-98.057
69	85	Cacahuacintle	Loma Ancha, Agua Blanca, Hidalgo	Sierra Madre Oriental	2133	20.339	-98.431
14	2009-REPO- 118	Celaya	Ejido Tuna Mansa, Silao, Guanajuato	Altiplano Sur (Zacatecano-Potosino)	1823	21.009	-101.405
12	2009-REPO- 114	Celaya	El Ahuacate, Uriangato, Guanajuato	Faja Volcánica	1877	20.091	-101.119
148	2009-REPO- 101	Celaya	San José de Araceo, Valle de Santiago, Guanajuato	Faja Volcánica	1750	20.374	-101.161
35	2009-006	Chalqueño	La Cofradía, Amatepec, Estado de México	Depresión del Balsas	1390	18.660	-100.228
37	2009-150	Chalqueño	San Lorenzo Oyamel, Temoaya, Estado de México	Faja Volcánica	2595	19.418	-99.595
67	92	Chalqueño	500 km NW Lazaro Cardenas, Huamantla, Tlaxcala	Faja Volcánica	2530	19.359	-97.988
68	22	Chalqueño	Ignacio Zaragoza, Cuapiaxlta, Tlaxcala	Faja Volcánica	2497	19.291	-97.929
184	2009-263	Chalqueño	Pahuacán, Ayapango, Estado de México	Faja Volcánica	2536	19.104	-98.804
70	111	Chalqueño	Rancho San Antonio Texas, Metepec, Hidalgo	Sierra Madre Oriental	2068	20.221	-98.337

Muestra	No. colecta	Raza	Localidad	Provincia biogeográfica	Altitud	Latitud	Longitud
71	112	Chalqueño	San Juan Hueyapan, Cuauhtepac de Hinojosa, Hidalgo	Sierra Madre Oriental	2136	20.030	-98.281
2	SON2007-064	Chapalote	Huasabas, Huasabas, Sonora	Sonorense	542	29.907	-109.301
4	SON2007-065	Chapalote	Huasabas, Huasabas, Sonora	Sonorense	537	29.904	-109.300
149	SON2007-067	Chapalote	Moctezuma, Moctezuma, Sonora	Sonorense	596	29.805	-109.674
15	Chis-2009-79 R	Chiquito	Ranchería San Nicolás, Salto de Agua, Chiapas	Golfo de México (Sur)	120	17.322	-92.089
81	79	Chiquito	Ranchería San Nicolás, Salto de Agua, Chiapas	Golfo de México (Sur)	120	17.322	-92.089
82	121	Comiteco	San Marcos, Independencia, Chiapas	Altos de Chiapas	1553	16.196	-91.977
150	122	Comiteco	La Libertad, Independencia, Chiapas	Altos de Chiapas	1550	16.249	-92.020
183	127	Comiteco	Carranza, Independencia, Chiapas	Altos de Chiapas	1550	16.195	-91.937
186	133	Comiteco	Guayamuses	Altos de Chiapas	1550	16.233	-91.975
80	39	Comiteco	Ejido Suchiapa, Suchiapa, Chiapas	Costa del Pacífico (Sur)	455	16.623	-93.100
21	12	Conejo	Olinalá, Olinalá, Guerrero	Depresión del Balsas	1312	17.780	-98.743
23	25	Conejo	San Gabriel, Olinalá, Guerrero	Depresión del Balsas	1685	17.778	-98.742
152	30	Conejo	Tamalakatzingo, Olinalá, Guerrero	Sierra Madre del Sur	1493	17.886	-98.681
182	45	Conejo	Cualac, Cualac, Guerrero	Sierra Madre del Sur	1402	17.745	-98.656
192	2009-182	Cónico	Xala, Axapusco, Estado de México	Altiplano Sur (Zacatecano-Potosino)	2479	19.759	-98.647

Muestra	No. colecta	Raza	Localidad	Provincia biogeográfica	Altitud	Latitud	Longitud
43	2009-046	Cónico	Colonia Libertad, Santa Cruz Atizapán, Estado de México	Faja Volcánica	2585	19.185	-99.510
44	2009-078	Cónico	El Espinal, Villa Victoria, Estado de México	Faja Volcánica	2605	19.461	-100.063
60	6	Cónico	El Carmen Tequexquitla, El Carmen Tequexquitla, Tlaxcala	Faja Volcánica	2358	19.319	-97.652
101	142	Cónico	Capula, Tetla, Tlaxcala	Faja Volcánica	2557	19.519	-98.048
154	2009-244	Cónico	San Antonio Soyatzingo, Amecameca, Estado de México	Faja Volcánica	2400	19.091	-98.782
179	2009-097	Cónico	San Antonio Mextepec, San Felipe del Progreso, Estado de México	Faja Volcánica	2663	19.651	-99.956
181	139	Cónico	Xalostoc, Xalostoc, Tlaxcala	Faja Volcánica	2469	19.394	-98.069
187	2009-084	Cónico	Emiliano Zapata, Ixtlahuaca, Estado de México	Faja Volcánica	2538	19.569	-99.791
188	2009-133	Cónico	San Pablo Malacatepec, Villa de Allende, Estado de México	Faja Volcánica	2452	19.352	-100.168
190	2009-080	Cónico	San Diego Suchitepec, Villa Victoria, Estado de México	Faja Volcánica	2576	19.491	-100.026
72	89	Cónico	Ejido Cerro Alto, Agua Blanca, Hidalgo	Sierra Madre Oriental	2059	20.332	-98.379
73	105	Cónico	Nopalillo, Metepec, Hidalgo	Sierra Madre Oriental	2144	20.243	-98.291
153	99	Cónico	Cañada de Flores, Metepec, Hidalgo	Sierra Madre Oriental	2143	20.287	-98.405

Muestra	No. colecta	Raza	Localidad	Provincia biogeográfica	Altitud	Latitud	Longitud
180	116	Cónico	El Capulín, Cuauhtemoc de Hinojosa, Hidalgo	Sierra Madre Oriental	2297	19.999	-98.275
189	82	Cónico	Ranchería Chichicaxtle, Agua Blanca, Hidalgo	Sierra Madre Oriental	2104	20.359	-98.374
24	3	Cónico Norteño	Bustillos, Cuauhtemoc, Chihuahua	Sierra Madre Occidental	2010	28.497	-106.632
25	4	Cónico Norteño	Bustillos, Cuauhtemoc, Chihuahua	Sierra Madre Occidental	2010	28.495	-106.629
156	5	Cónico Norteño	Zamaloapan, Cuauhtemoc, Chihuahua	Sierra Madre Occidental	2019	28.519	-106.656
98	263	Coscomatepec	San Antonio Rayon, Xonotla, Puebla	Golfo de México (Norte)	167	20.110	-97.485
157	291	Coscomatepec	Buena Vista, Ayotoxco de Guerrero, Puebla	Golfo de México (Norte)	126	20.127	-97.456
26	VER-54A	Coscomatepec	El Palmar, Tezonapa, Veracruz	Golfo de México (Sur)	201	18.500	-96.783
28	255	Coscomatepec	Equimita, Cuetzalán, Puebla	Sierra Madre Oriental	1604	19.965	-97.565
33	14	Cristalino de Chihuahua	La Cantera, Cusihuirachi, Chihuahua	Sierra Madre Occidental	2094	28.193	-106.879
34	17	Cristalino de Chihuahua	Guadalupe, Guerrero, Chihuahua	Sierra Madre Occidental	2034	28.503	-107.475
36	CHIH-235	Dulce	Moris, Moris, Chihuahua	Sierra Madre Occidental	950	28.133	-108.533
91	SON2007-033	Dulcillo del Noroeste	Agua Blanca, Yécora, Sonora	Sierra Madre Occidental	1435	28.537	-108.925
108	SON2007-049	Dulcillo del Noroeste	Huisamopa, Sahuaripa, Sonora	Sierra Madre Occidental	708	28.613	-109.106
39	Repetida 18	Dzit-Bacal	Nuevo Vicente Guerrero, Villacorzo, Chiapas	Costa del Pacífico (Sur)	618	16.032	-92.980
40	Repetida 31	Dzit-Bacal	Ejido El Parral, Villacorzo, Chiapas	Costa del Pacífico (Sur)	654	16.364	-93.002

Muestra	No. colecta	Raza	Localidad	Provincia biogeográfica	Altitud	Latitud	Longitud
159	67	Dzit-Bacal	Rivera El Gavilán, Ocozocoautla de Espinosa, Chiapas	Costa del Pacífico (Sur)	779	16.752	-93.457
16	SIN-2007-054	Elotero de Sinaloa	Agua Zarca, El Rosario, Sinaloa	Costa del Pacífico (Norte)	220	22.890	-105.546
88	SIN-2007-026	Elotero de Sinaloa	Santa Catarina, Concordia, Sinaloa	Costa del Pacífico (Norte)	209	23.431	-106.082
89	SIN-2007-031	Elotero de Sinaloa	La Guayanera, Concordia, Sinaloa	Costa del Pacífico (Norte)	929	23.406	-105.892
173	SIN-2007-083	Elotero de Sinaloa	La Labor, San Ignacio, Sinaloa	Costa del Pacífico (Norte)	258	23.846	-106.426
174	SIN-2007-038	Elotero de Sinaloa	Potrerrillo, Concordia, Sinaloa	Costa del Pacífico (Norte)	1526	23.458	-105.831
55	2009-081	Elotes Cónicos	San Diego Suchitepec, Villa Victoria, Estado de México	Faja Volcánica	2576	19.491	-100.026
56	2009-085	Elotes Cónicos	Emiliano Zapata, Ixtlahuaca, Estado de México	Faja Volcánica	2538	19.569	-99.791
57	2009-13	Elotes Cónicos	San Pedro Atocpan, Milpa Alta, Distrito Federal	Faja Volcánica	2530	19.192	-99.050
65	39	Elotes Cónicos	Francisco Javier Mina, Trinidad Sánchez Santos, Tlaxcala	Faja Volcánica	2640	19.191	-97.925
66	58	Elotes Cónicos	5 km E Huamantla, Huamantla, Tlaxcala	Faja Volcánica	2621	19.322	-97.973
160	2009-240	Elotes Cónicos	San Juan Tehuixtitlan, Atlautla, Estado de México	Faja Volcánica	2600	19.056	-98.772
161	128	Elotes Cónicos	San José Teacalco-Quinta Sección, San José Tecalco, Tlaxcala	Faja Volcánica	2620	19.339	-98.047

Muestra	No. colecta	Raza	Localidad	Provincia biogeográfica	Altitud	Latitud	Longitud
170	2009-098	Elotes Cónicos	San Juan Jalpa, San Felipe del Progreso, Estado de México	Faja Volcánica	2535	19.706	-99.929
171	183	Elotes Cónicos	Santiago Tlalpan, Hueyotlipan, Tlaxcala	Faja Volcánica	2584	19.488	-98.316
191	2009-266	Elotes Cónicos	Pahuacán, Ayapango, Estado de México	Faja Volcánica	2560	19.109	-98.807
74	81	Elotes Cónicos	Ranchería Chichicaxtle, Agua Blanca, Hidalgo	Sierra Madre Oriental	2104	20.359	-98.374
75	87	Elotes Cónicos	Calabazas-Primera Seccion, Agua Blanca, Hidalgo	Sierra Madre Oriental	2081	20.355	-98.376
102	73	Elotes Cónicos	Ranchería Potrero de Monroy, Agua Blanca, Hidalgo	Sierra Madre Oriental	1698	20.398	-98.383
172	86	Elotes Cónicos	Calabazas-Primera Seccion, Agua Blanca, Hidalgo	Sierra Madre Oriental	2081	20.355	-98.376
162	2009-REPO-015	Elotes Occidentales	Miraflores, Comonfort, Guanajuato	Altiplano Sur (Zacatecano-Potosino)	2071	20.787	-100.936
41	2009-REPO-002	Elotes Occidentales	Comonfort, Comonfort, Guanajuato	Faja Volcánica	1800	20.744	-100.766
42	2009-REPO-009	Elotes Occidentales	Palmillas de San Juan, Comonfort, Guanajuato	Faja Volcánica	1854	20.784	-100.815
103	2009-063	Elotes Occidentales	El Peñón, Temascaltepec, Estado de México	Faja Volcánica	1779	19.044	-100.108
92	SON2007-084	Gordo	Mesa de Tres Ríos, Nacori Chico, Sonora	Sierra Madre Occidental	1880	29.840	-108.710
163	SON2007-085	Gordo	Mesa de Tres Ríos, Nacori Chico, Sonora	Sierra Madre Occidental	1880	29.840	-108.710
45	Nay-2010-91	Harinoso de Ocho	San Felipe Aztatán, Tecuala, Nayarit	Costa del Pacífico (Norte)	43	22.384	-105.419

Muestra	No. colecta	Raza	Localidad	Provincia biogeográfica	Altitud	Latitud	Longitud
46	NAY-129	Jala	Jala, Jala, Nayarit	Costa del Pacífico (Norte)	1060	21.100	-104.441
47	2003-2	Jala	Jomulco, Jala, Nayarit	Costa del Pacífico (Norte)	1153	21.102	-104.429
164	2003-6	Jala	Jomulco, Jala, Nayarit	Costa del Pacífico (Norte)	1153	21.102	-104.429
169	2004-17	Jala	Jomulco, Jala, Nayarit	Costa del Pacífico (Norte)	1057	21.078	-104.440
48	OAX-190	Mixeño	La Aguilera, Totontepec, Oaxaca	Sierra Madre del Sur	1500	17.217	-95.983
104	190	Mountain Yellow	Ejido Benito Juárez, Motozintla, Chiapas	Soconusco	1967	15.363	-92.308
50	2009-JAAG-062	Mushito	Derramaderos, Victoria, Guanajuato	Altiplano Sur (Zacatecano-Potosino)	2500	21.358	-100.137
105	2009-09	Mushito	Mulegé, Mulegé, Baja California Sur	Baja California	5	26.890	-111.986
49	2009-JAAG-049	Mushito	Cerro Prieto, Aarjea, Guanajuato	Sierra Madre Oriental	1850	21.278	-99.805
83	135	Nal-Tel de Altura	Guayamuses, Independencia, Chiapas	Altos de Chiapas	1551	16.233	-91.977
84	136	Nal-Tel de Altura	Campamento, Independencia, Chiapas	Altos de Chiapas	1552	16.216	-91.916
51	OAX-165	Nal-Tel de Altura	El Ciruelo, Santiago Pinotepa Nacional, Oaxaca	Costa del Pacífico (Sur)	50	16.300	-98.217
52	OAX-400	Nal-Tel de Altura	Finca Carlota, San Bartolomé Ayautla, Oaxaca	Golfo de México (Sur)	1250	18.017	-96.733
99	VER-115	Nal-Tel de Altura	Ixtaczoquitlán, Ixtaczoquitlán, Veracruz	Sierra de Oaxaca	920	18.850	-97.067
78	18	Olotillo	Nuevo Vicente Guerrero, Villacorzo, Chiapas	Costa del Pacífico (Sur)	618	16.032	-92.980
79	31	Olotillo	Ejido El Parral, Villacorzo, Chiapas	Costa del Pacífico (Sur)	654	16.364	-93.002

Muestra	No. colecta	Raza	Localidad	Provincia biogeográfica	Altitud	Latitud	Longitud
165	260	Olotillo	Medio Monte-Primera Sección, Tuxtla Chico, Chiapas	Costa del Pacífico (Sur)	191	14.902	-92.217
200	264	Olotillo	Medio Monte, Tuxtla Chico, Chiapas	Costa del Pacífico (Sur)	159	14.874	-92.205
76	136	Olotillo	Ixtlahuatempa, Huejutla, Hidalgo	Golfo de México (Norte)	225	21.133	-98.436
77	152	Olotillo	Tenexco I, Atlapexco, Hidalgo	Golfo de México (Norte)	194	21.064	-98.346
166	228	Olotón	Mitzitón, San Cristobal de las Casas, Chiapas	Altos de Chiapas	2299	16.640	-92.542
85	192	Olotón	Ejido Benito Juárez, Motozintla, Chiapas	Soconusco	2057	15.366	-92.315
86	197	Olotón	Barrio Tenan, Motozintla, Chiapas	Soconusco	1828	15.360	-92.294
167	269	Olotón	Puente Maravillas, Motozintla, Chiapas	Soconusco	1220	15.296	-92.188
109	SON2007-003	Onaveño	Cochibampo, Álamos, Sonora	Costa del Pacífico (Norte)	461	27.173	-108.824
110	SON2007-058	Onaveño	Moctezuma, Moctezuma, Sonora	Sonorense	626	29.809	-109.680
54	220	Palomero de Chihuahua	Puerto Los Chiles, Balleza, Chihuahua	Sierra Madre Occidental	2769	26.360	-106.444
58	2009-169	Palomero Toluqueño	San Marcos Tlazalpan, San Bartolo Morelos, Estado de México	Faja Volcánica	2688	19.798	-99.717
93	2009-059	Pepitilla	San Pedro Temeyac, Temascaltepec, Estado de México	Faja Volcánica	1830	19.053	-100.169

Muestra	No. colecta	Raza	Localidad	Provincia biogeográfica	Altitud	Latitud	Longitud
94	2009-060	Pepitilla	San Pedro Temeyac, Temascaltepec, Estado de México	Faja Volcánica	1830	19.053	-100.169
106	2009-061	Pepitilla	El Peñón, Temascaltepec, Estado de México	Faja Volcánica	1779	19.044	-100.108
176	2009-066	Pepitilla	El Peñón, Temascaltepec, Estado de México	Faja Volcánica	1779	19.044	-100.108
113	NL-2008-054	Ratón	Rinconada Nuevo León, Villa de García, Nuevo León	Altiplano Norte (Chihuahuense)	1001	25.681	-100.704
111	TMS-35	Ratón	Oyama, Hidalgo, Tamaulipas	Golfo de México (Norte)	231	24.128	-99.255
198	63	Ratón	Ejido Nuevo San Antonio, San Carlos, Tamaulipas	Golfo de México (Norte)	199	24.290	-99.021
114	SON2007-009	Reventador	La Isleta, Álamos, Sonora	Costa del Pacífico (Norte)	204	26.844	-108.912
197	SON2007-072	Reventador	La Estancia, Aconchi, Sonora	Sonorense	598	29.794	-110.210
151	JAL-829	Serrano de Jalisco	Rancho El Tepopote, Atoyac, Jalisco	Faja Volcánica	2000	19.929	-103.438
19	JAL-825	Serrano de Jalisco	Llano Ojo Zarco, Tapalpa, Jalisco	Faja Volcánica	2160	19.933	-103.667
20	JAL-827	Serrano de Jalisco	Tapalpa, Tapalpa, Jalisco	Faja Volcánica	1800	19.950	-103.767
116	131	Tablilla de Ocho	La Concepción, Gran Morelos, Chihuahua	Altiplano Norte (Chihuahuense)	1791	28.320	-106.632
201	323	Tablilla de Ocho	La Noria y Anexos, Huejotitán, Chihuahua	Sierra Madre Occidental	1775	26.942	-106.016
118	SIN-2006-010	Tabloncillo	Boca de Arroyo, Badiraguato, Sinaloa	Costa del Pacífico (Norte)	226	25.401	-107.562
119	SIN-2006-020	Tabloncillo	Maloya, El Rosario, Sinaloa	Costa del Pacífico (Norte)	113	22.949	-105.607

Muestra	No. colecta	Raza	Localidad	Provincia biogeográfica	Altitud	Latitud	Longitud
121	SON2007-022	Tabloncillo	Los Vallecitos, Yécora, Sonora	Sierra Madre Occidental	1570	28.411	-108.926
120	SON2007-020	Tabloncillo	La Estrella, Rosario, Sonora	Sonorense	503	27.827	-109.245
123	85	Tabloncillo Perla	Palma Grande, Tuxpan, Nayarit	Costa del Pacífico (Norte)	24	21.991	-105.390
195	51	Tabloncillo Perla	El Colomo, Bahía de Banderas, Nayarit	Costa del Pacífico (Norte)	54	20.874	-105.149
202	80	Tabloncillo Perla	Peñas, Del Nayar, Nayarit	Costa del Pacífico (Norte)	20	21.946	-105.222
124	175	Tehua	San Pablo Huacana, Ocoatepec, Chiapas	Altos de Chiapas	1545	17.197	-93.206
125	177	Tehua	San Pablo Huacana, Ocoatepec, Chiapas	Altos de Chiapas	1525	17.216	-93.175
178	275	Tepecintle	Monterrey Segunda, Tila, Chiapas	Altos de Chiapas	899	17.251	-92.467
177	266	Tepecintle	Cacaoatales, Metapa de Oguez, Chiapas	Costa del Pacífico (Sur)	100	14.820	-92.197
17	Chis-2009-80 R	Tepecintle	Vicente Guerrero, Salto de Agua, Chiapas	Golfo de México (Sur)	120	17.322	-92.089
126	195	Tepecintle	Ejido Benito Juárez, Motozintla, Chiapas	Soconusco	2050	15.367	-92.319
127	SIN-2007-003	Tuxpeño	Malpica, Concordia, Sinaloa	Costa del Pacífico (Norte)	86	23.254	-106.132
131	1	Tuxpeño	Benito Juárez, La Concordia, Chiapas	Costa del Pacífico (Sur)	575	16.044	-92.819
132	2	Tuxpeño	Ejido La Concordia, La Concordia, Chiapas	Costa del Pacífico (Sur)	552	16.119	-92.689
129	NL-2008-017	Tuxpeño	Las Crucitas, Linares, Nuevo León	Tamaulipeca	557	24.756	-99.723
133	2009-006	Tuxpeño Norteño	La Cofradía, Amatepec, Estado de México	Depresión del Balsas	1390	18.660	-100.228

Muestra	No. colecta	Raza	Localidad	Provincia biogeográfica	Altitud	Latitud	Longitud
134	NL-2008-002	Tuxpeño Norteño	Ejido La Petaca, Linares, Nuevo León	Tamaulipeca	334	24.859	-99.534
193	277	Vandeño	Monterrey Segunda, Tila, Chiapas	Altos de Chiapas	899	17.251	-92.467
137	5	Vandeño	Ejido La Concordia, La Concordia, Chiapas	Costa del Pacífico (Sur)	551	16.121	-92.696
138	14	Vandeño	Nuevo Vicente Guerrero, Villacorzo, Chiapas	Costa del Pacífico (Sur)	612	16.042	-92.978
135	SON2007-070	Vandeño	La Estancia, Aconchi, Sonora	Sonorense	610	29.796	-110.212
139	Jal-445	Zamorano Amarillo	Villa Purificación, Purificación, Jalisco	Costa del Pacífico (Sur)	450	19.717	-104.633
140	Jal-377	Zamorano Amarillo	Zapotillo, Casimiro Castillo, Jalisco	Costa del Pacífico (Sur)	300	19.600	-104.450
141	Mich-18	Zamorano Amarillo	Juiquilpan, Juiquilpan, Michoacán	Faja Volcánica	1550	19.967	-102.717
142	35	Zapalote Chico	Ejido El Parral, Villacorzo, Chiapas	Costa del Pacífico (Sur)	654	16.364	-93.002
144	35	Zapalote Grande	Ejido El Parral, Villacorzo, Chiapas	Costa del Pacífico (Sur)	654	16.364	-93.002

Anexo II.B. Localidades de 261 poblaciones de 59 razas cuyos individuos fueron analizados con isoenzimas y microsatélites. En negritas son poblaciones representadas sólo con isoenzimas y subrayadas, sólo con microsatélites. Las poblaciones representan accesiones del Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias (INIFAP) o del Centro Internacional de Mejoramiento del Maíz y Trigo (CIMMYT). Las provincias biogeográficas corresponden a CONABIO 1997.

No.	Población	Raza	Localidad	Provincia biogeográfica	Altitud	Latitud	Longitud
1	GRO-326	Ancho	Chilapa, Guerrero	Sierra Madre del Sur	1300	17.583	-99.183
2	GRO-383	Ancho	Iguala-Campuzano, Guerrero	Depresión del Balsas	1200	18.417	-99.550
3	MOR-154	Ancho	Totolapan, Morelos	Faja Volcánica	1900	18.983	-98.930
4	MOR-46	Ancho	Santa Catarina, Morelos	Faja Volcánica	1700	18.967	-99.150
5	MOR-62	Ancho	San Andrés Cautempan, Morelos	Faja Volcánica	1630	18.983	-98.950
6	CHIH-138	Apachito	Ejido Largo, Chihuahua	Sierra Madre Occidental	2210	29.683	-108.267
7	CHIH-166	Apachito	San Juanito, Chihuahua	Sierra Madre Occidental	2400	27.983	-107.583
8	CHIH-177	Apachito	Basagochic, Chihuahua	Sierra Madre Occidental	2000	28.017	-107.783
9	CHIH-180	Apachito	Bocoyna, Chihuahua	Sierra Madre Occidental	2180	27.830	-107.583
10	CHIH-182	Apachito	Estación Creel, Chihuahua	Sierra Madre Occidental	2310	27.750	-107.650
11	CHIH-207	Apachito	Agua Blanca, Chihuahua	Sierra Madre Occidental	2510	26.750	-107.267
12	PUE-864	Arrocillo Amarillo	Plan de Guadalupe, Puebla	Sierra Madre Oriental	2260	19.883	-97.583
13	PUE-91	Arrocillo Amarillo	Zaragoza, Puebla	Faja Volcánica	2317	19.767	-97.550
14	VER-311	Arrocillo Amarillo	Tepichulco, Veracruz	Sierra Madre Oriental	2220	19.800	-97.250

No.	Población	Raza	Localidad	Provincia biogeográfica	Altitud	Latitud	Longitud
15	VER-342	Arrocillo Amarillo	Altotonga, Veracruz	Sierra Madre Oriental	2410	19.767	-97.250
16	VER-359	Arrocillo Amarillo	Estanzuela, Veracruz	Faja Volcánica	2200	19.700	-97.217
17	CHIH-147	Azul	Gómez Farías, Chihuahua	Sierra Madre Occidental	2140	29.350	-107.750
18	CHIH-158	Azul	Guerrero, Chihuahua	Sierra Madre Occidental	2040	28.567	-107.483
19	CHIH-218	Azul	Matachic, Chihuahua	Sierra Madre Occidental	1970	28.830	-107.730
20	CHIH-220	Azul	Tocomoro chic, Chihuahua	Sierra Madre Occidental	1900	28.700	-107.883
21	SIN-61	Blando de Sonora	San Javier de Abajo, Sinaloa	Costa del Pacífico	750	25.950	-107.400
22	SIN-7	Blando de Sonora	Las Amelas, Sinaloa	Sonorense	115	26.430	-108.630
23	SON-117	Blando de Sonora	Bacanora, Sonora	Sonorense	410	28.967	-109.400
24	SON-29	Blando de Sonora	Sahuaripa, Sonora	Sonorense	400	29.050	-109.250
25	DGO-110	Bofo	San José del Palmar, Durango	Costa del Pacífico	500	24.967	-106.967
26	DGO-123	Bofo	San José del Palmar, Durango	Costa del Pacífico	500	24.967	-106.967
27	DGO-95	Bofo	Huazamota, Durango	Sierra Madre Occidental	510	22.630	-104.450
28	JAL-289	Bofo	Mexquitic, Jalisco	Costa del Pacífico	1400	22.383	-103.730
29	NAY-191	Bofo	Amatlán de Jora, Nayarit	Sierra Madre Occidental	1000	21.383	-104.130
30	NAY-203	Bofo	Hda. Rosas Blancas, Nayarit	Costa del Pacífico	1850	21.100	-104.430
31	OAX-221	Bolita	Agua del Sol, Oaxaca	Sierra Madre del Sur	1549	16.300	-96.600
32	OAX-28	Bolita	Etla, Oaxaca	Sierra Madre del Sur	1646	17.200	-96.800
33	OAX-40	Bolita	San Juan Chilateca, Oaxaca	Sierra Madre del Sur	1520	16.783	-96.670

No.	Población	Raza	Localidad	Provincia biogeográfica	Altitud	Latitud	Longitud
34	OAX-44	Bolita	Teotitlán del Valle, Oaxaca	Sierra Madre del Sur	793	17.030	-96.500
35	OAX-68	Bolita	Nochixtlán, Oaxaca	Sierra Madre del Sur	1646	17.467	-97.300
36	MEX-212	Cacahuacintle	Estado de México	Faja Volcánica	2620	19.300	-99.650
37	MEX-7	Cacahuacintle	Toluca, Estado de México	Faja Volcánica	2652	19.283	-99.670
38	PUE-552	Cacahuacintle	San Martín Ojo de Agua, Puebla	Faja Volcánica	2600	19.000	-97.383
39	PUE-591	Cacahuacintle	Tlachichuca, Puebla	Faja Volcánica	2600	19.117	-97.417
40	GTO-101	Celaya	Irapuato, Guanajuato	Faja Volcánica	1700	20.670	-101.350
41	GTO-265	Celaya	Celaya, Guanajuato	Altiplano Sur (Zacatecano-Potosino)	1750	20.517	-100.817
42	GTO-29	Celaya	San Juan, Guanajuato	Altiplano Sur (Zacatecano-Potosino)	1799	20.917	-101.467
43	GTO-36	Celaya	Salvatierra, Guanajuato	Faja Volcánica	1799	20.217	-100.883
44	GTO-69	Celaya	Cortázar, Guanajuato	Faja Volcánica	1812	20.483	-100.967
45	GTO-84	Celaya	Progreso (Jaral del), Guanajuato	Faja Volcánica	1738	20.367	-101.067
46	GTO-88	Celaya	Sta. Cruz, Guanajuato	Faja Volcánica	1768	20.200	-101.000
47	HGO-7	Chalqueño	El Rincón, Hidalgo	Altiplano Sur (Zacatecano-Potosino)	2100	20.267	-98.917
48	MEX-208	Chalqueño	Chalco, Estado de México	Faja Volcánica	2250	19.267	-98.930
49	MEX-37	Chalqueño	Tepotztlán, Estado de México	Faja Volcánica	2273	19.717	-99.217
50	MEX-48	Chalqueño	Tultepec, Estado de México	Faja Volcánica	2256	19.683	-99.130
51	PUE-82	Chalqueño	San Pedro Atlixco, Puebla	Faja Volcánica	1829	18.967	-98.467
52	PUE-87	Chalqueño	Cholula, Puebla	Faja Volcánica	2181	19.067	-98.300

No.	Población	Raza	Localidad	Provincia biogeográfica	Altitud	Latitud	Longitud
53	ZAC-4	Chalqueño	Ojo Caliente, Zacatecas	Altiplano Sur (Zacatecano-Potosino)	2104	22.567	-102.250
54	SIN-2	Chapalote	Culiacán, Sinaloa	Sonorense	61	24.800	-107.417
55	SIN-6	Chapalote	Las Amelas, Sinaloa	Sonorense	115	26.430	-108.630
56	SIN-65	Chapalote	San Javier de Abajo, Sinaloa	Costa del Pacífico	750	25.950	-107.400
57	OAX-571	Choapaneco	San Juan Comaltepec, Oaxaca	Sierra Madre del Sur	760	17.300	-95.983
58	OAX-572	Choapaneco	San Juan Comaltepec, Oaxaca	Sierra Madre del Sur	760	17.300	-95.983
59	OAX-574	Choapaneco	San Juan Comaltepec, Oaxaca	Sierra Madre del Sur	760	17.300	-95.983
60	OAX-575	Choapaneco	San Juan Comaltepec, Oaxaca	Sierra Madre del Sur	760	17.300	-95.983
61	OAX-576	Choapaneco	San Juan Comaltepec, Oaxaca	Sierra Madre del Sur	760	17.300	-95.983
62	CHIS-352	Comiteco	Santa Rita, Chiapas	Altos de Chiapas	1530	16.117	-92.000
63	CHIS-38	Comiteco	San Jerónimo, Chiapas	Costa del Pacífico	2043	15.683	-92.200
64	CHIS-39	Comiteco	El Retiro, Chiapas	Soconusco	1400	15.230	-92.267
65	CHIS-609	Comiteco	Finca La Campana, Chiapas	Altos de Chiapas	1700	16.517	-92.483
66	CHIS-86	Comiteco	Comitán, Chiapas	Altos de Chiapas	1555	16.250	-92.130
67	GRO-129	Conejo	Murga, Guerrero	Costa del Pacífico	100	17.670	-101.317
68	GRO-157	Conejo	San Jerónimo, Guerrero	Costa del Pacífico	30	17.117	-100.467
69	GRO-17	Conejo	El Treinta, Guerrero	Costa del Pacífico	100	16.983	-99.800
70	GRO-176	Conejo	Zacualpa, Guerrero	Sierra Madre del Sur	300	17.767	-98.583

No.	Población	Raza	Localidad	Provincia biogeográfica	Altitud	Latitud	Longitud
71	MEX-108	Cónico	Ran. Guadalupe, Estado de México	Faja Volcánica	2520	19.100	-99.583
72	MEX-58	Cónico	Tlacotepec, Estado de México	Faja Volcánica	2580	18.967	-99.930
73	MEX-72	Cónico	Toluca, Estado de México	Faja Volcánica	2652	19.283	-99.670
74	PUE-109	Cónico	Acatepec, Puebla	Faja Volcánica	2104	18.230	-97.567
75	PUE-32	Cónico	Ixtlahuaca, Puebla	Sierra Madre Oriental	2250	19.850	-97.983
<u>76</u>	<u>AGS-7</u>	<u>Cónico Norteño</u>	<u>Tepezala, Aguascalientes</u>	<u>Altiplano Sur (Zacatecano-Potosino)</u>	<u>2165</u>	<u>22.217</u>	<u>-102.167</u>
77	GTO-165	Cónico Norteño	Cerrito, Guanajuato	Altiplano Sur (Zacatecano-Potosino)	2000	21.150	-100.930
78	GTO-22	Cónico Norteño	El Jicote, Guanajuato	Altiplano Sur (Zacatecano-Potosino)	1780	20.950	-101.430
79	GTO-23	Cónico Norteño	Marroquín, Guanajuato	Altiplano Sur (Zacatecano-Potosino)	1780	20.930	-101.417
80	GTO-34	Cónico Norteño	Buena Vista, Guanajuato	Altiplano Sur (Zacatecano-Potosino)	1780	20.950	-101.417
81	QRO-2	Cónico Norteño	San Juan del Río, Querétaro	Altiplano Sur (Zacatecano-Potosino)	1921	20.383	-100.000
82	ZAC-12	Cónico Norteño	Nochistlán, Zacatecas	Altiplano Sur (Zacatecano-Potosino)	1951	21.367	-102.850
83	VER-110	Coscomatepec	La Cuchilla / Huatusco, Veracruz	Sierra Madre Oriental	1320	19.167	-96.967
84	VER-113	Coscomatepec	La Cuchilla / Huatusco, Veracruz	Sierra Madre Oriental	1340	19.167	-96.967
85	VER-404	Coscomatepec	San Diego / Huatusco, Veracruz	Sierra Madre Oriental	1515	19.167	-96.967

No.	Población	Raza	Localidad	Provincia biogeográfica	Altitud	Latitud	Longitud
86	VER-457	Coscomatepec	Elotepec, Veracruz	Faja Volcánica	1800	19.200	-97.030
87	CHIH-128	Cristalino de Chihuahua	Madera, Chihuahua	Sierra Madre Occidental	2095	29.217	-108.130
88	CHIH-154	Cristalino de Chihuahua	Gómez Farías, Chihuahua	Sierra Madre Occidental	2140	29.350	-107.750
89	CHIH-254	Cristalino de Chihuahua	La Junta, Chihuahua	Sierra Madre Occidental	1900	28.450	-107.350
90	JAL-300	Dulce de Jalisco	Mexquitic, Jalisco	Costa del Pacífico	1400	22.383	-103.730
91	JAL-304	Dulce de Jalisco	Colotlán, Jalisco	Sierra Madre Occidental	1700	22.117	-103.267
92	JAL-78	Dulce de Jalisco	Yahualica, Jalisco	Altiplano Sur (Zacatecano-Potosino)	1880	21.183	-102.883
93	ZAC-182	Dulce de Jalisco	Potrero de Gallegos, Zacatecas	Sierra Madre Occidental	2050	22.650	-103.700
94	SIN-25	Dulcillo del Noroeste	Alcoyonque, Sinaloa	Sonorense	70	24.800	-107.400
95	SIN-33	Dulcillo del Noroeste	La Noria / Jesús María, Sinaloa	Costa del Pacífico	500	25.067	-107.500
96	SIN-34	Dulcillo del Noroeste	La Noria / Jesús María, Sinaloa	Costa del Pacífico	500	25.067	-107.500
97	SIN-79	Dulcillo del Noroeste	La Chicura, Sinaloa	Sonorense	180	26.430	-108.630
98	SON-57	Dulcillo del Noroeste	La Chicura, Sonora	Sonorense	180	26.430	-108.630
99	CHIS-447	Dzit Bacal	El Zapotal, Chiapas	Altos de Chiapas	690	16.950	-92.717
100	QROO-20	Dzit Bacal	Isla Mujeres, Quintana Roo	Petén	10	21.250	-86.750
101	VER-96	Dzit Bacal	Amatlán, Veracruz	Oaxaca	700	18.850	-96.917
102	NAY-32	Elotero de Sinaloa	Milpas Viejas, Nayarit	Costa del Pacífico	100	22.400	-105.450
103	NAY-46	Elotero de Sinaloa	Las Peñitas, Nayarit	Costa del Pacífico	100	21.950	-105.217

No.	Población	Raza	Localidad	Provincia biogeográfica	Altitud	Latitud	Longitud
104	SIN-17	Elotero de Sinaloa	La Noria / Mazatlán, Sinaloa	Costa del Pacífico	200	23.500	-106.300
105	PUE-403	Elotes Cónicos	Santa María Coronango, Puebla	Faja Volcánica	2200	19.100	-98.300
106	PUE-510	Elotes Cónicos	Santa María Hidalgo, Puebla	Faja Volcánica	2200	18.967	-98.300
107	TLAX-251	Elotes Cónicos	Tenancingo, Tlaxcala	Faja Volcánica	2280	19.150	-98.200
108	GTO-1	Elotes Occidentales	S.P. de los Hernández, Guanajuato	Altiplano Sur (Zacatecano-Potosino)	1885	21.117	-101.683
109	JAL-54	Elotes Occidentales	Zapotitán, Jalisco	Faja Volcánica	1555	20.300	-103.500
110	JAL-633	Elotes Occidentales	Tepatitlán, Jalisco	Faja Volcánica	1800	20.817	-102.767
111	JAL-71	Elotes Occidentales	Tizapán, Jalisco	Faja Volcánica	1555	20.167	-103.067
112	NAY-29	Elotes Occidentales	Quimichis, Nayarit	Costa del Pacífico	46	22.350	-105.530
113	NAY-38	Elotes Occidentales	Tuxpan, Nayarit	Costa del Pacífico	46	21.930	-105.300
114	ZAC-180	Elotes Occidentales	Potrero de Gallegos, Zacatecas	Sierra Madre Occidental	2050	22.650	-103.700
115	ZAC-210	Elotes Occidentales	Jalpa, Zacatecas	Sierra Madre Occidental	1500	21.617	-102.950
116	CHIH-131	Gordo	Madera, Chihuahua	Sierra Madre Occidental	2095	29.217	-108.130
117	CHIH-160	Gordo	Guerrero, Chihuahua	Sierra Madre Occidental	2040	28.567	-107.483
118	CHIH-205	Gordo	Agua Blanca, Chihuahua	Sierra Madre Occidental	2510	26.750	-107.267
119	CHIH-214	Gordo	Guerrero, Chihuahua	Sierra Madre Occidental	2040	28.567	-107.483
120	CHIH-256	Gordo	La Junta, Chihuahua	Sierra Madre Occidental	1900	28.450	-107.350
121	NAY-24	Harinoso de Ocho	San Vicente, Nayarit	Costa del Pacífico	46	21.950	-105.300

No.	Población	Raza	Localidad	Provincia biogeográfica	Altitud	Latitud	Longitud
122	JAL-69	Jala	San Nicolás Tlaqq., Jalisco	Faja Volcánica	1616	20.650	-103.317
123	NAY-130	Jala	Jala, Nayarit	Costa del Pacífico	1095	21.100	-104.430
124	NAY-54	Jala	Jala, Nayarit	Costa del Pacífico	1040	21.100	-104.430
125	NAY-6	Jala	Jala, Nayarit	Costa del Pacífico	1098	21.083	-104.430
126	OAX-566	Mixeño	Totontepec, Oaxaca	Sierra Madre del Sur	1850	17.217	-96.050
127	OAX-567	Mixteco	Macuiltianguis, Oaxaca	Oaxaca	2100	17.567	-96.567
128	OAX-568	Mixteco	Yodocono, Oaxaca	Sierra Madre del Sur	2500	17.383	-97.367
129	OAX-569	Mixteco	Yodocono, Oaxaca	Sierra Madre del Sur	2500	17.383	-97.367
130	OAX-570	Mixteco	San José de Gracia, Oaxaca	Sierra Madre del Sur	2600	17.417	-97.483
131	CHIS-650	Motuzinteco	Motuzintla, Chiapas	Soconusco	1270	15.367	-92.250
132	CHIS-652	Motuzinteco	Motuzintla, Chiapas	Soconusco	1270	15.367	-92.250
133	CHIS-653	Motuzinteco	Motuzintla, Chiapas	Soconusco	1270	15.367	-92.250
134	JAL-127	Mountain Yellow	Tapalpa, Jalisco	Faja Volcánica	2060	19.950	-103.767
135	JAL-129	Mountain Yellow	Tapalpa, Jalisco	Faja Volcánica	2060	19.950	-103.767
136	JAL-131	Mountain Yellow	Tapalpa, Jalisco	Faja Volcánica	2060	19.950	-103.767
137	MICH-214	Mountain Yellow	Purépero, Michoacán	Faja Volcánica	2020	19.917	-102.000
138	OAX-G43	Mushito	Tanaco, Oaxaca	Faja Volcánica	2200	19.683	-102.117
139	MICH-328	Mushito	Tanaco, Oaxaca	Faja Volcánica	2200	19.683	-102.117
140	MICH-351	Mushito	Los Reyes, Michoacán	Faja Volcánica	2320	19.583	-102.467
141	MICH-371	Mushito	San Miguel Macasepo, Michoacán	Faja Volcánica	2100	19.500	-101.600
142	CAMP-48	Nal-Tel	Pueblo Nuevo, Campeche	Yucatán	50	19.783	-90.117
143	QROO-39	Nal-Tel	Tikosuco, Quintana Roo	Yucatán	10	20.250	-88.300
144	YUC-148	Nal-Tel	Oxkutzcab, Yucatán	Yucatán	30	20.300	-89.430
145	YUC-7	Nal-Tel	Dzitas, Yucatán	Yucatán	30	20.830	-88.517

No.	Población	Raza	Localidad	Provincia biogeográfica	Altitud	Latitud	Longitud
146	CHIS-196	Nal-Tel de Altura	Bejucal de Ocampo, Chiapas	Soconusco	2300	15.467	-92.150
147	OAX-298	Nal-Tel de Altura	Santa María Pápalo, Oaxaca	Oaxaca	1800	17.800	-96.730
148	OAX-301	Nal-Tel de Altura	Santa María Pápalo, Oaxaca	Oaxaca	2060	17.800	-96.730
149	OAX-310	Nal-Tel de Altura	Ayutla, Oaxaca	Sierra Madre del Sur	2060	17.000	-96.083
150	OAX-117	Negrito	Boca de León, Oaxaca	Sierra Madre del Sur	2100	17.430	-96.930
151	OAX-26	Negrito	Etla, Oaxaca	Sierra Madre del Sur	1640	17.200	-96.800
152	CHIS-230	Negro de Chimaltenango	Región Lacandona, Chiapas	Altos de Chiapas	400	16.900	-92.100
153	CHIS-698	Negro de Chimaltenango	Navenchau, Chiapas	Altos de Chiapas	2250	16.767	-92.717
154	CHIS-440	Olotillo	Totolapa, Chiapas	Altos de Chiapas	690	16.600	-92.683
<u>155</u>	<u>CHS-52</u>	<u>Olotillo</u>	<u>Terán, Chiapas</u>	<u>Costa del Pacífico</u>	<u>584</u>	<u>16.750</u>	<u>-93.167</u>
156	CHIS-53	Olotillo	Terán, Chiapas	Costa del Pacífico	579	16.750	-93.167
157	CHIS-56	Olotillo	Terán, Chiapas	Costa del Pacífico	560	16.750	-93.167
158	CHIS-562	Olotillo	Terán, Chiapas	Costa del Pacífico	570	16.750	-93.167
159	CHIS-81	Olotillo	El Zapotal, Chiapas	Costa del Pacífico	686	16.600	-92.717
160	SLP-108	Olotillo	Tampacán, San Luis Potosí	Golfo de México	305	21.417	-98.717
161	CHIS-684	Olotón	Chempil, Chiapas	Altos de Chiapas	2100	16.767	-92.450
162	CHIS-687	Olotón	Chilil, Chiapas	Altos de Chiapas	2200	16.730	-92.467
163	CHIS-695	Olotón	Romerillo, Chiapas	Altos de Chiapas	2390	16.800	-92.550
164	SON-105	Onaveño	Guadalupe Tayopa, Sonora	Sierra Madre Occidental	820	28.317	-108.967
165	SON-114	Onaveño	Bacanora, Sonora	Sonorense	410	28.967	-109.400
166	SON-155	Onaveño	Mazatán, Sonora	Sonorense	500	29.000	-110.117
167	SON-184	Onaveño	Tubutama, Sonora	Sonorense	520	30.883	-111.467

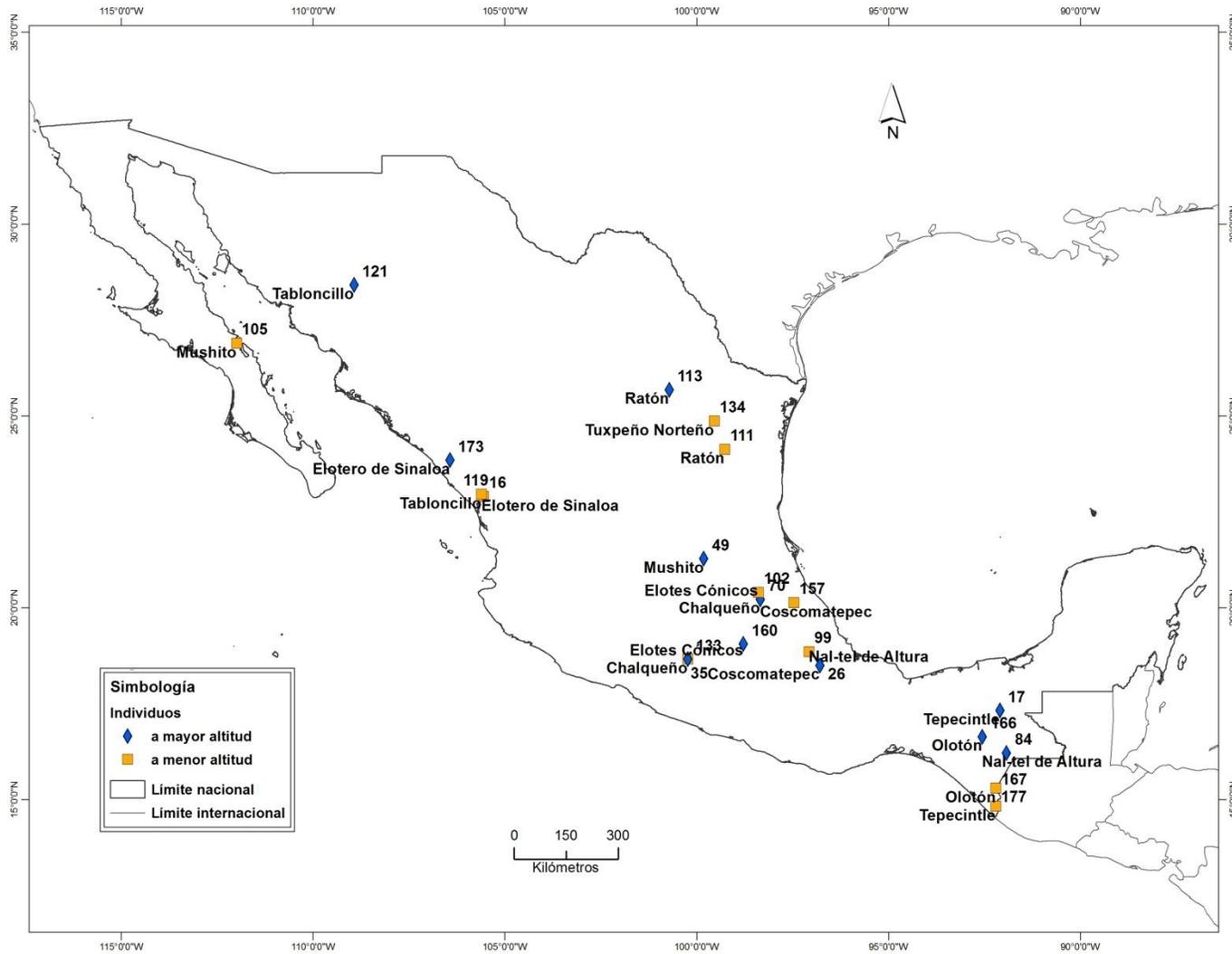
No.	Población	Raza	Localidad	Provincia biogeográfica	Altitud	Latitud	Longitud
168	CHIH-135	Palomero de Chihuahua	Madera, Chihuahua	Sierra Madre Occidental	2095	29.217	-108.130
169	CHIH148	Palomero de Chihuahua	Gómez Farías, Chihuahua	Sierra Madre Occidental	2140	29.350	-107.750
170	CHIH150	Palomero de Chihuahua	Gómez Farías, Chihuahua	Sierra Madre Occidental	2140	29.350	-107.750
171	CHIH-201	Palomero de Chihuahua	Gualaina, Chihuahua	Sierra Madre Occidental	2130	26.817	-107.067
172	JAL-142	Palomero de Jalisco	Juanacatlán, Jalisco	Faja Volcánica	2520	20.030	-103.683
173	JAL-154	Palomero de Jalisco	Juanacatlán, Jalisco	Faja Volcánica	2520	20.030	-103.683
174	JAL-159	Palomero de Jalisco	Juanacatlán, Jalisco	Faja Volcánica	2520	20.030	-103.683
175	MEX-211	Palomero Toluqueño	Toluca, Estado de México	Faja Volcánica	2620	19.283	-99.670
176	MEX-5	Palomero Toluqueño	México-Toluca km 57, Estado de México	Faja Volcánica	2652	19.283	-99.670
177	MEX-6	Palomero Toluqueño	Toluca (Mercado), Estado de México	Faja Volcánica	2652	19.283	-99.670
178	TLAX-311	Palomero Toluqueño	Calpulalpan, Tlaxcala	Faja Volcánica	2580	19.583	-98.567
179	GRO-3	Pepitilla	Iguala, Guerrero	Depresión del Balsas	747	18.350	-99.530
180	GRO-335	Pepitilla	Huitzuc, Guerrero	Depresión del Balsas	900	18.300	-99.350
181	MOR-102	Pepitilla	Chinameca, Morelos	Depresión del Balsas	1140	18.650	-98.983
182	MOR-17	Pepitilla	Paraimalco-Iguala, Morelos	Depresión del Balsas	747	18.350	-99.530
183	MOR-99	Pepitilla	Tlaltizapán, Morelos	Depresión del Balsas	1040	18.683	-99.117
184	COAH-25	Ratón	Santa Elena, Coahuila	Altiplano Norte (Chihuahuense)	1610	25.750	-102.967
185	NL-9	Ratón	Paso de las Cabras, Nuevo León	Tamaulipeca	250	25.583	-99.317

No.	Población	Raza	Localidad	Provincia biogeográfica	Altitud	Latitud	Longitud
186	TAMS-25	Ratón	Conrado Castillo, Tamaulipas	Tamaulipeca	290	24.567	-99.600
187	TAMS-66	Ratón	Comales, Tamaulipas	Tamaulipeca	50	26.130	-98.917
188	NAY-15	Reventador	Las Peñitas, Nayarit	Costa del Pacífico	46	22.300	-105.383
189	NAY-39	Reventador	Tuxpan, Nayarit	Costa del Pacífico	46	21.930	-105.300
190	SIN-55	Reventador	San Javier de Abajo, Sinaloa	Costa del Pacífico	750	25.950	-107.400
191	SIN-60	Reventador	San Javier de Abajo, Sinaloa	Costa del Pacífico	750	25.950	-107.400
192	JAL-133	Serrano de Jalisco	Tapalpa, Jalisco	Faja Volcánica	2060	19.950	-103.767
193	JAL-141	Serrano de Jalisco	Juanacatlán, Jalisco	Faja Volcánica	2520	20.030	-103.683
194	JAL-146	Serrano de Jalisco	Juanacatlán, Jalisco	Faja Volcánica	2520	20.030	-103.683
195	JAL-173	Serrano de Jalisco	Atemajac de Brisuela, Jalisco	Faja Volcánica	2565	20.130	-103.730
196	JAL-753	Serrano de Jalisco	Juanacatlán, Jalisco	Faja Volcánica	2520	20.030	-103.683
197	OAX-565	Serrano Mixe	Yacoche, Oaxaca	Sierra Madre del Sur	2350	17.130	-96.017
198	JAL-301	Tablilla de Ocho	Colotlán, Jalisco	Sierra Madre Occidental	1700	22.117	-103.267
199	JAL-306	Tablilla de Ocho	Colotlán, Jalisco	Sierra Madre Occidental	1700	22.117	-103.267
200	NAY-185	Tablilla de Ocho	Puente de Comatlán, Nayarit	Sierra Madre Occidental	1100	21.683	-104.067
201	NAY-189	Tablilla de Ocho	Puente de Comatlán, Nayarit	Sierra Madre Occidental	1100	21.683	-104.067
202	NAY-198	Tablilla de Ocho	Amatlán de Jora, Nayarit	Sierra Madre Occidental	1000	21.383	-104.130
203	ZAC-187	Tablilla de Ocho	Monte Escobedo, Zacatecas	Sierra Madre Occidental	2170	22.300	-103.550
204	JAL-100	Tabloncillo	La Rivera, Jalisco	Faja Volcánica	1890	20.367	-102.167

No.	Población	Raza	Localidad	Provincia biogeográfica	Altitud	Latitud	Longitud
205	JAL-102	Tabloncillo	Tecalitlán, Jalisco	Faja Volcánica	1280	19.467	-103.317
206	JAL-42	Tabloncillo	Autlán, Jalisco	Costa del Pacífico	1250	19.767	-104.367
207	JAL-43	Tabloncillo	Santa Ana, Jalisco	Faja Volcánica	1330	19.983	-104.267
208	JAL-44	Tabloncillo	Los Tocotes, Jalisco	Faja Volcánica	1300	19.950	-104.267
209	JAL-63	Tabloncillo	Autlán, Jalisco	Costa del Pacífico	1010	19.767	-104.367
210	JAL-87	Tabloncillo	Tuxpan, Jalisco	Faja Volcánica	1382	19.550	-103.383
211	NAY-12	Tabloncillo Perla	Las Peñitas, Nayarit	Costa del Pacífico	46	21.950	-105.217
212	NAY-16	Tabloncillo Perla	Chilapa, Nayarit	Costa del Pacífico	46	22.083	-105.250
213	NAY-41	Tabloncillo Perla	Palma Grande, Nayarit	Costa del Pacífico	100	21.930	-105.300
214	CHIS-159	Tehua	Santa Ana, Chiapas	Altos de Chiapas	747	16.500	-92.500
215	CHIS-204	Tehua	La Trinitaria, Chiapas	Altos de Chiapas	1600	16.130	-92.050
216	CHIS-229	Tehua	Mazapa, Chiapas	Soconusco	1000	15.383	-92.183
217	CHIS-234	Tehua	Col. Alvarado, Chiapas	Altos de Chiapas	-	16.500	-92.500
218	CHIS-29	Tehua	Finca Prusia, Chiapas	Soconusco	1000	15.750	-92.717
219	CHIS-596	Tehua	Rivera Zapote, Chiapas	Altos de Chiapas	820	16.967	-93.100
220	CHIS-225	Tepecintle	San Felipe-Escuintla, Chiapas	Costa del Pacífico	91	15.300	-92.630
221	CHIS-26	Tepecintle	Mapastepec, Chiapas	Costa del Pacífico	107	15.430	-92.900
222	CHIS-528	Tepecintle	Rosendo Salazar, Chiapas	Costa del Pacífico	730	16.717	-93.717
223	CHIS-76	Tepecintle	Chiapilla, Chiapas	Costa del Pacífico	579	16.500	-92.750
224	OAX-177	Tepecintle	Jicaltepec, Oaxaca	Costa del Pacífico	200	16.383	-98.017
225	OAX-9	Tuxpeño	Mata de Cana, Oaxaca	Golfo de México	100	18.117	-96.150
226	PUE-27	Tuxpeño	San José Acateno, Puebla	Golfo de México	152	20.130	-97.200
227	TAMS-125	Tuxpeño	Ursulo Galván, Tamaulipas	Golfo de México	150	22.283	-97.883
228	VER-128	Tuxpeño	Texistepec, Veracruz	Golfo de México	150	17.867	-94.783

No.	Población	Raza	Localidad	Provincia biogeográfica	Altitud	Latitud	Longitud
229	VER-39	Tuxpeño	Güiterrez Zamora, Veracruz	Golfo de México	107	20.483	-97.100
230	VER-44	Tuxpeño	El Palmar, Veracruz	Golfo de México	366	20.430	-97.000
231	CHIH-121	Tuxpeño Norteño	Janos, Chihuahua	Altiplano Norte (Chihuahuense)	1470	30.900	-108.167
232	CHIH-13	Tuxpeño Norteño	Saucillo, Chihuahua	Altiplano Norte (Chihuahuense)	1210	28.017	-105.267
233	COAH-21	Tuxpeño Norteño	Ojo de Apolo, Coahuila	Altiplano Norte (Chihuahuense)	1400	25.430	-102.183
234	COAH-36	Tuxpeño Norteño	Nadadores, Coahuila	Altiplano Norte (Chihuahuense)	900	27.067	-101.600
235	TAMS-29	Tuxpeño Norteño	Guadalupe, Tamaulipas	Tamaulipeca	290	24.567	-99.600
236	TAMS-3	Tuxpeño Norteño	Llera de Canales, Tamaulipas	Golfo de México	300	23.317	-99.017
237	JAL-756	Uruapeño	El Tepopote, Jalisco	Faja Volcánica	2060	19.830	-103.500
238	CHIS-112	Vandeño	Escuintla, Chiapas	Costa del Pacífico	107	15.317	-92.670
239	CHIS-114	Vandeño	Escuintla, Chiapas	Costa del Pacífico	108	15.300	-92.630
240	CHIS-25	Vandeño	Mapastepec, Chiapas	Costa del Pacífico	108	15.430	-92.900
241	CHIS-30	Vandeño	El Paval, Chiapas	Altos de Chiapas	614	16.500	-92.500
242	CHIS-30A	Vandeño	El Paval, Chiapas	Altos de Chiapas	600	16.500	-92.500
243	CHIS-31	Vandeño	Mapastepec, Chiapas	Costa del Pacífico	107	15.430	-92.900
244	CHIS-91	Vandeño	Col. Hidalgo, Chiapas	Altos de Chiapas	1800	16.250	-92.130
245	GRO-311	Vandeño	Tanganhuato, Guerrero	Depresión del Balsas	250	18.267	-100.600
246	GTO-191	Zamorano Amarillo	Ej. Albarradones, Guanajuato	Altiplano Sur (Zacatecano-Potosino)	1870	21.117	-101.683
247	JAL-513	Zamorano Amarillo	Viborero, Jalisco	Faja Volcánica	1930	20.670	-102.467
248	JAL-631	Zamorano Amarillo	Tepatitlán, Jalisco	Faja Volcánica	1800	20.817	-102.767

No.	Población	Raza	Localidad	Provincia biogeográfica	Altitud	Latitud	Longitud
249	MICH-5	Zamorano Amarillo	Zamora, Michoacán	Faja Volcánica	1560	19.983	-102.300
250	MICH-66	Zamorano Amarillo	El Tecolote, Michoacán	Faja Volcánica	1645	20.100	-101.950
251	CHIS-390	Zapalote Chico	Comalapa, Chiapas	Costa del Pacífico	700	15.650	-92.150
252	CHIS-662	Zapalote Chico	Col. Azteca, Chiapas	Costa del Pacífico	100	16.217	-93.950
253	OAX-48	Zapalote Chico	Reforma, Oaxaca	Costa del Pacífico	50	16.400	-94.230
254	OAX-50	Zapalote Chico	Niltepec, Oaxaca	Costa del Pacífico	107	16.567	-94.617
255	OAX-51	Zapalote Chico	Niltepec, Oaxaca	Costa del Pacífico	107	16.567	-94.617
256	OAX-52	Zapalote Chico	Niltepec, Oaxaca	Costa del Pacífico	100	16.567	-94.617
257	OAX-70	Zapalote Chico	Tehuantepec, Oaxaca	Costa del Pacífico	46	16.300	-95.230
258	CHIS-104	Zapalote Grande	Escuintla, Chiapas	Costa del Pacífico	107	15.317	-92.670
259	CHIS-224	Zapalote Grande	San Felipe-Escuintla, Chiapas	Costa del Pacífico	91	15.300	-92.630
260	CHIS-236	Zapalote Grande	Col. Lázaro Cárdenas, Chiapas	Costa del Pacífico	120	16.717	-93.800
261	CHIS-521	Zapalote Grande	Rosendo Salazar, Chiapas	Costa del Pacífico	730	16.717	-93.717



ANEXO III. Distribución geográfica de los pares de individuos de 11 razas con hipotenusa mayor a 100 km. Estos individuos fueron seleccionados para obtener el índice de diferenciación genética (F_{ST}). Los números corresponden a la columna Muestra del Anexo II. Diseño: Oswaldo Oliveros y Alejandra Breña Ochoa.

ANEXO IV. Localización de los 154 SNPs con índice F_{ST} igual o mayor a 0.15 en los 10 cromosomas de 169 individuos de 50 razas. Se indica el nombre del marcador, el cromosoma (C) del genoma de referencia B73, las coordenadas (pares de bases) en el mapa físico de B73, el tipo de polimorfismo, el gen correspondiente, la secuencia de referencia de la dbSNP (www.ncbi.nlm.nih.gov/snp/) y el valor de F_{ST} entre provincias biogeográficas. Aún no se conoce la posición sinónima o no sinónima de algunos marcadores en exones de genes, por lo que se colocó exón como tipo de polimorfismo.

Marcador	C	Coordenada (pb)	Tipo	Gen	dbSNP (#ss)	F_{ST}
PZE-101000450	1	999598	exón	GRMZM2G073979	ss196422192	0.166
PZE-101015149	1	8623017	no sinónimo	GRMZM2G171324	ss196422378	0.153
PZE-101049044	1	33799525	intrón	GRMZM2G110913	ss196423275	0.174
PZE-101066291	1	49393138	intrón	GRMZM2G092442	ss196423830	0.165
PZA00468.8	1	50096948	intrón	GRMZM2G394212	ss196414500	0.152
PZE-101116836	1	140961510	no sinónimo	GRMZM2G157127	ss196427215	0.157
PZE-101117480	1	142337816	no génico	-	ss196427274	0.150
PZE-100001428	1	161457991	sinónimo	GRMZM2G028036	ss196422009	0.179
PZE-101140546	1	181559144	no génico	-	ss196428960	0.157
SYN25877	1	182076775	sinónimo	GRMZM2G023906	ss196513466	0.155
SYN1740	1	183843865	sinónimo	GRMZM2G113899	ss196506473	0.168
SYN1741	1	183848635	intrón	GRMZM2G113911	ss196506485	0.164
SYN1739	1	183849259	sinónimo	GRMZM2G113911	ss196506465	0.171
PZE-101142851	1	183850352	sinónimo	GRMZM2G113911	ss196429068	0.171
ZM012177-0317	1	183850558	sinónimo	GRMZM2G113911	ss196530724	0.158
PZE-101160748	1	203331217	sinónimo	GRMZM2G139683	ss196429694	0.154
PZE-101190298	1	236090113	-	GRMZM2G162874	ss196431085	0.162
SYN13216	1	241222849	sinónimo	GRMZM2G081554	ss196503242	0.156
SYNGENTA3100	1	241223114	sinónimo	GRMZM2G081554	ss196530228	0.157
SYN8671	1	283601031	intrón	GRMZM2G065374	ss196528208	0.159
PUT-163a-60401045-3037	1	285061831	exón	GRMZM2G361220	ss196420306	0.158

Marcador	C	Coordenada (pb)	Tipo	Gen	dbSNP (#ss)	F_{ST}
PZA03699.1	2	7525572	intrón	GRMZM2G164277	ss196414975	0.232
SYN28948	2	37868398	no génico	-	ss196515971	0.157
SYN23120	2	54966030	sinónimo	GRMZM2G079323	ss196511219	0.180
PZE-102085119	2	74045919	no génico	-	ss196436086	0.212
PZE-102092853	2	103583409	no génico	-	ss196436939	0.157
PZE-102092975	2	103821194	no génico	-	ss196436957	0.188
PZE-102092976	2	103821271	no génico	-	ss196436959	0.196
PZE-102109861	2	141561790	no génico	-	ss196438621	0.169
PZE-102119420	2	161608716	sinónimo	GRMZM2G158575	ss196439463	0.168
PZE-102137516	2	186383746	no sinónimo	GRMZM2G359746	ss196440374	0.169
PZE-102156631	2	203773309	exón	GRMZM2G010941	ss196440999	0.205
SYN39126	2	203900206	no génico	-	ss196524119	0.168
PZE-102169738	2	213204599	no génico	-	ss196441321	0.179
SYN28305	2	225530997	sinónimo	GRMZM2G458077	ss196515396	0.150
PZE-103045135	3	45705092	intrón	GRMZM2G096548	ss196443793	0.166
PZE-103045167	3	45791696	no génico	-	ss196443797	0.153
PZE-103045276	3	46130382	sinónimo	GRMZM2G130366	ss196443805	0.160
SYN4078	3	46176192	exón	GRMZM2G130348	ss196524466	0.155
SYN4083	3	46178588	intrón	GRMZM2G130348	ss196524470	0.155
PZE-103064196	3	97373259	no génico	-	ss196445719	0.185
PZE-103064150	3	97582711	no génico	-	ss196445706	0.156
PZE-103064029	3	97752254	no génico	-	ss196445696	0.159
PZE-103063980	3	97904716	no génico	-	ss196445692	0.150
PZE-103075246	3	124666487	no génico	-	ss196446690	0.162
PZE-103075282	3	124730507	intrón	GRMZM2G018372	ss196446692	0.162
PZB00451.1	3	124732596	-	GRMZM2G018527	ss196415419	0.151
PZE-103078709	3	130049328	sinónimo	GRMZM2G040750	ss196446944	0.187

Marcador	C	Coordenada (pb)	Tipo	Gen	dbSNP (#ss)	F_{ST}
PZE-103080851	3	134023298	intrón	GRMZM2G021471	ss196447133	0.151
PZE-103088303	3	146525161	intrón	GRMZM2G103342	ss196447663	0.199
PZE-103096467	3	156312085	no génico	-	ss196448216	0.156
SYN10954	3	186828855	sinónimo	GRMZM2G024806	ss196501280	0.156
SYN34635	3	209726331	intrón	GRMZM2G080499	ss196520496	0.191
SYN1962	3	211982835	intrón	GRMZM2G152686	ss196508381	0.199
PZE-103172227	3	219335861	exón	GRMZM2G079089	ss196450517	0.152
PUT-163a-74235050-3609	3	124732209-2347	-	GRMZM2G018372	ss196420791	0.152
SYN26447	4	17236520	exón	GRMZM2G080168	ss196513883	0.191
PZE-104111271	4	98321469	no génico	-	ss196458369	0.214
PZE-104111268	4	98325326	no génico	-	ss196458365	0.185
PZE-104057394	4	108408787	no génico	-	ss196455148	0.168
PZB02053.2	4	118066179	sinónimo	GRMZM2G096171	ss196415786	0.162
PZE-104093384	4	169960496	exón	GRMZM2G082222	ss196457656	0.151
PZE-104096422	4	172833586	exón	GRMZM2G162052	ss196457766	0.155
SYN2309	4	181057827	exón	GRMZM2G029323	ss196511183	0.155
PZE-104122141	4	199167311	intrón	GRMZM2G083309	ss196458920	0.151
PZE-104123862	4	201110345	intrón	GRMZM2G079436	ss196459019	0.165
SYN16139	4	240463135	sinónimo	GRMZM2G163227	ss196505424	0.185
PZE-105029276	5	15315326	intrón	GRMZM2G100412	ss196461091	0.150
PZE-105046526	5	35190971	no génico	-	ss196461859	0.163
PZE-105051255	5	44719418	intrón	GRMZM2G145236	ss196462371	0.152
PZE-105051311	5	44987971	intrón	GRMZM2G159890	ss196462375	0.154
PZE-105051344	5	45089554	no génico	-	ss196462377	0.164
PZE-105051409	5	45333476	intrón	GRMZM2G067789	ss196462387	0.166
PZE-105051472	5	45440557	intrón	GRMZM2G083076	ss196462397	0.156
PZE-105051581	5	45905841	no génico	-	ss196462407	0.151
PZE-105060093	5	59045273	intrón	GRMZM2G439198	ss196462935	0.151

Marcador	C	Coordenada (pb)	Tipo	Gen	dbSNP (#ss)	F_{ST}
PZE-105060180	5	59143184	sinónimo	GRMZM2G135570	ss196462937	0.153
PZE-105063649	5	63326340	sinónimo	GRMZM2G088737	ss196463110	0.187
PZE-105063667	5	63375651	intrón	GRMZM2G389462	ss196463112	0.224
PZE-105065789	5	67505063	no génico	-	ss196463237	0.168
PZE-105065790	5	67505086	no génico	-	ss196463239	0.151
PZE-105072737	5	77750221	sinónimo	GRMZM2G010920	ss196463644	0.160
SYN16842	5	84799961	exón	GRMZM2G427468	ss196506031	0.168
PZE-105085845	5	107798973	no génico	-	ss196464779	0.159
PZE-105086589	5	110179006	no génico	-	ss196464841	0.160
SYN702	5	112041787	no génico	-	ss196526853	0.193
SYN21719	5	204008438	exón	GRMZM2G018126	ss196510105	0.160
SYN32227	5	204316878	no sinónimo	GRMZM2G012213	ss196518651	0.156
PZE-106025869	6	61728691	no génico	-	ss196471131	0.168
PZE-106025873	6	61740963	no génico	-	ss196471133	0.199
PZE-106025971	6	61857661	no génico	-	ss196471143	0.197
PZE-106032882	6	76167968	no génico	-	ss196471713	0.158
PZE-106032919	6	76236367	intrón	GRMZM2G051338	ss196471719	0.162
PZE-106123501	6	164558994	no génico	-	ss196475032	0.181
SYN25439	7	4161522	exón	GRMZM2G175930	ss196513099	0.154
SYN32201	7	8467890	intrón	GRMZM2G312910	ss196518633	0.166
PZE-107039304	7	75985070	no génico	-	ss196477463	0.167
PZE-107040496	7	79193083	no génico	-	ss196477579	0.155
PZE-107051697	7	101576864	intrón	GRMZM2G016660	ss196478650	0.164
PZE-107051699	7	101585302	intrón	GRMZM2G016660	ss196478652	0.161
PZE-107051773	7	101742073	sinónimo	GRMZM2G049364	ss196478656	0.154
PZE-107055203	7	106994194	intrón	GRMZM2G067752	ss196478929	0.184
PZA03728.1	7	142783202	sinónimo	GRMZM2G048170	ss196416672	0.159

Marcador	C	Coordenada (pb)	Tipo	Gen	dbSNP (#ss)	F_{ST}
SYN34204	7	151242431	intrón	GRMZM2G025242	ss196520142	0.176
PZE-107095739	7	151409024	sinónimo	GRMZM2G029082	ss196480686	0.169
PZE-107118935	7	165344935	intrón	GRMZM2G080567	ss196481150	0.169
PZE-108021327	8	20514430	no génico	-	ss196482277	0.181
PZE-108027532	8	28744724	no génico	-	ss196482539	0.206
PZE-108027534	8	28746956	no génico	-	ss196482541	0.196
PZE-108031468	8	35862710	no génico	-	ss196482886	0.185
PZE-108034774	8	44401700	sinónimo	GRMZM2G117401	ss196483201	0.182
PZE-108034832	8	44551786	exón	GRMZM2G012601	ss196483215	0.178
PZA00417.2	8	45262229	intrón	GRMZM2G142870	ss196416844	0.178
PZE-108038903	8	45262229	intrón	GRMZM2G142870	ss196483771	0.178
PZE-108041594	8	67989802	no génico	-	ss196484020	0.188
PZE-108042026	8	68976451	intrón	GRMZM2G070013	ss196484067	0.150
PZE-108042154	8	69130177	no sinónimo	GRMZM2G125294	ss196484085	0.150
PZE-108042248	8	69412154	exón	GRMZM2G098129	ss196484089	0.164
PZE-108042363	8	69754674	no génico	-	ss196484093	0.150
PZE-108042667	8	70677401	no génico	-	ss196484139	0.170
PZE-108046489	8	77169519	sinónimo	GRMZM2G416201	ss196484463	0.172
PZE-108068741	8	120768244	exón	GRMZM2G380561	ss196486289	0.157
PZE-108073195	8	127626059	exón	GRMZM2G135366	ss196486608	0.155
PZE-108080073	8	136138158	exón	GRMZM2G100229	ss196486968	0.159
PZE-108091524	8	148675564	intrón	GRMZM2G130586	ss196487490	0.168
PZE-109014052	9	14155059	intrón	GRMZM2G047138	ss196489141	0.179
PZE-109030126	9	33803014	intrón	GRMZM2G065355	ss196489837	0.184
PZE-109030176	9	33890865	exón	GRMZM2G144173	ss196489843	0.203
PZE-109030178	9	33986832	no génico	-	ss196489845	0.188
PZE-109038290	9	56595583	intrón	GRMZM2G171031	ss196490631	0.169

Marcador	C	Coordenada (pb)	Tipo	Gen	dbSNP (#ss)	F_{ST}
PZE-109038687	9	57600573	no génico	-	ss196490684	0.164
PZE-109039048	9	58496990	no génico	-	ss196490733	0.151
PZE-109043223	9	72010825	no génico	-	ss196491167	0.170
PZE-109047651	9	82284479	exón	GRMZM2G009335	ss196491625	0.179
PZE-109049639	9	86396507	sinónimo	GRMZM2G407790	ss196491783	0.195
PZE-109063042	9	105253087	intrón	GRMZM5G863602	ss196492608	0.186
PZE-109075044	9	121602110	sinónimo	GRMZM2G004222	ss196493187	0.152
SYN16617	10	3165902	intrón	GRMZM2G356817	ss196505807	0.160
SYN12503	10	10290662	exón	GRMZM2G093895	ss196502593	0.159
PZE-110017754	10	19851929	sinónimo	GRMZM2G431157	ss196495279	0.179
PZE-110021154	10	28237501	no génico	-	ss196495632	0.182
PZE-110022148	10	30863669	no génico	-	ss196495746	0.191
PZE-110022324	10	31311797	intrón	GRMZM2G115424	ss196495772	0.198
PZE-110022966	10	33615965	intrón	GRMZM2G171518	ss196495881	0.153
PZE-110027837	10	37810598	no génico	-	ss196496414	0.164
PZE-110032375	10	50920595	no génico	-	ss196496854	0.160
PZE-110032461	10	51436407	no génico	-	ss196496866	0.155
PZE-110052325	10	98944143	no génico	-	ss196498592	0.167
PZE-110053452	10	100682765	no génico	-	ss196498681	0.168
PZE-110053456	10	100683137	no génico	-	ss196498683	0.160
SYN16891	10	102292971	sinónimo	GRMZM2G035899	ss196506080	0.154
PUT-163a-78116163-4360	10	111843474	sinónimo	GRMZM2G317406	ss196421386	0.194
PZE-110066819	10	123229863	no sinónimo	GRMZM2G107306	ss196499543	0.169
PZE-110072863	10	129354223	no génico	-	ss196499759	0.185

ANEXO V

Anexo V.A. Estadísticos de diversidad genética por provincia biogeográfica de 34 isoenzimas y de 96 microsatélites (sombreado). Los genotipos representan a 259 o 201 poblaciones y 59 o 58 razas, respectivamente. Los totales se calcularon a partir de las medias de cada provincia. El error estándar está entre paréntesis.

Provincia biogeográfica	N¹ (R²)	RA³	AP⁴	H_O⁵	H_E⁶
Altiplano Norte (Chihuahuense)	60 (2)	1.264 (0.039)	0.032 (0.006)	0.199 (0.033)	0.260 (0.038)
	5 (2)	1.780 (0.020)	0.388 (0.033)	0.566 (0.026)	0.779 (0.019)
Altiplano Sur (Zacatecano-Potosino)	180 (6)	1.292 (0.041)	0.064 (0.015)	0.215 (0.033)	0.290 (0.040)
	9 (6)	1.809 (0.013)	0.403 (0.031)	0.564 (0.022)	0.808 (0.013)
Altos de Chiapas	204 (7)	1.275 (0.042)	0.048 (0.011)	0.194 (0.032)	0.272 (0.040)
	8 (5)	1.778 (0.017)	0.384 (0.031)	0.509 (0.025)	0.777 (0.017)
Depresión del Balsas	96 (3)	1.271 (0.042)	0.048 (0.009)	0.207 (0.037)	0.265 (0.042)
	4 (2)	1.777 (0.020)	0.385 (0.034)	0.589 (0.032)	0.767 (0.021)
Golfo de México	120 (3)	1.268 (0.037)	0.056 (0.012)	0.194 (0.029)	0.265 (0.036)
	5 (3)	1.761 (0.018)	0.380 (0.034)	0.526 (0.026)	0.759 (0.017)
Faja Volcánica	840 (18)	1.286 (0.041)	0.048 (0.011)	0.213 (0.033)	0.285 (0.040)
	49 (16)	1.816 (0.014)	0.404 (0.028)	0.557 (0.017)	0.813 (0.014)
Costa del Pacífico	804 (19)	1.272 (0.038)	0.043 (0.006)	0.191 (0.030)	0.268 (0.038)
	49 (19)	1.823 (0.014)	0.399 (0.028)	0.550 (0.015)	0.818 (0.014)
Petén	12 (1)	1.243 (0.041)	0.027 (0.009)	0.218 (0.044)	0.235 (0.041)
	2 (1)	1.747 (0.025)	0.374 (0.040)	0.589 (0.039)	0.740 (0.026)
Sierra de Oaxaca	48 (3)	1.291 (0.039)	0.051 (0.009)	0.236 (0.037)	0.287 (0.038)
	3 (2)	1.784 (0.017)	0.394 (0.040)	0.616 (0.027)	0.777 (0.019)
Sierra Madre Occidental	516 (10)	1.266 (0.040)	0.050 (0.011)	0.198 (0.032)	0.262 (0.039)
	31 (10)	1.814 (0.014)	0.411 (0.076)	0.558 (0.017)	0.809 (0.014)
Sierra Madre Oriental	96 (3)	1.297 (0.039)	0.054 (0.015)	0.220 (0.033)	0.293 (0.038)
	7 (3)	1.808 (0.013)	0.374 (0.026)	0.629 (0.025)	0.808 (0.013)
Sierra Madre del Sur	276 (9)	1.300 (0.042)	0.061 (0.012)	0.225 (0.034)	0.298 (0.041)
	14 (8)	1.811 (0.013)	0.382 (0.029)	0.566 (0.019)	0.810 (0.013)
Soconusco	96 (4)	1.260 (0.041)	0.056 (0.015)	0.178 (0.034)	0.253 (0.040)
	7 (4)	1.766 (0.019)	0.350 (0.031)	0.483 (0.021)	0.766 (0.018)

Provincia biogeográfica	N¹ (R²)	RA³	AP⁴	H_O⁵	H_E⁶
Sonorenses	144 (4)	1.265 (0.041)	0.043 (0.011)	0.196 (0.033)	0.264 (0.040)
	12 (4)	1.809 (0.014)	0.423 (0.027)	0.527 (0.019)	0.804 (0.014)
Tamaulipeca	48 (2)	1.259 (0.039)	0.038 (0.010)	0.211 (0.035)	0.260 (0.038)
	3 (1)	1.721 (0.022)	0.368 (0.034)	0.571 (0.032)	0.719 (0.022)
Yucatán	60 (1)	1.182 (0.041)	0.050 (0.023)	0.134 (0.029)	0.188 (0.040)
	5 (1)	1.764 (0.018)	0.428 (0.040)	0.457 (0.029)	0.762 (0.018)
Total	3600	1.268 (0.007)	0.048 (0.002)	0.202 (0.008)	0.265 (0.010)
	213	1.786 (0.007)	0.390 (0.005)	0.554 (0.006)	0.782 (0.004)

¹N: número de individuos.

²R: razas por provincia.

³RA: riqueza alélica.

⁴AP: alelos privados.

⁵H_O: heterocigosis observada.

⁶H_E: heterocigosis esperada.

Anexo V.B. AMOVA y StAMOVA por provincia biogeográfica de 34 isoenzimas que representan a 259 poblaciones de 59 razas.

Fuente	df¹	SS²	MS³	p⁴	Φ_{ST}
AMOVA					
Provincia Biogeográfica	15	2004.265	133.618	0.001	0.055
Error	3584	36704.731	10.241		
Total	3599	38708.996			
StAMOVA					
Altitud	1	265.390	265.390	0.004	
Latitud	1	3.943	3.943	0.482	
Longitud	1	-1.621	-1.621	0.510	
Altitud-Latitud	1	577.838	577.838	0.001	
Altitud-Longitud	1	37.848	37.848	0.351	
Latitud-Longitud	1	200.644	200.644	0.031	
Provincia Biogeográfica	15	1226.069	81.738	0.001	0.033
Error	3578	36398.885	10.173		
Total	3599	38708.996	10.756		

¹df: grados de libertad

²SS: suma de cuadrados

³MS: cuadrados medios

⁴p: significancia.

Anexo V.C. AMOVA y StAMOVA por provincia biogeográfica de 96 microsatélites que representan a 201 poblaciones de 58 razas.

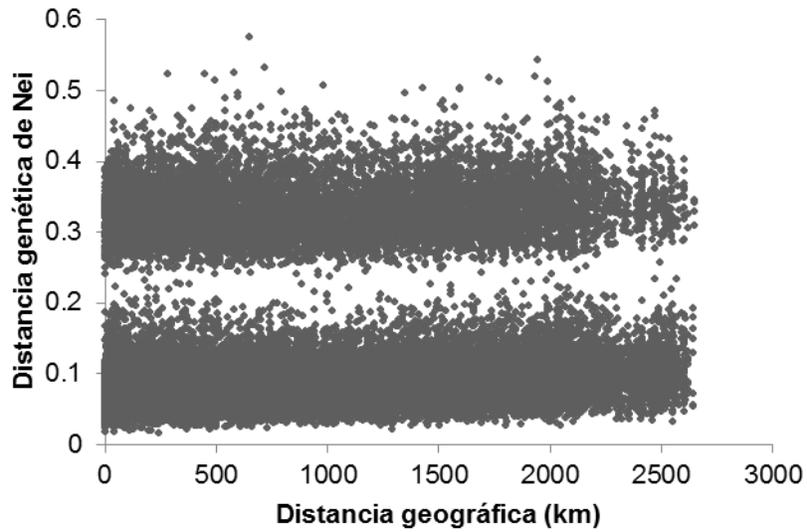
Fuente	df¹	SS²	MS³	p⁴	Φ_{ST}
AMOVA					
Provincia Biogeográfica	15	1876.027	125.069	0.001	0.023
Error	197	19192.786	97.425		
Total	212	21068.813			
StAMOVA					
Altitud	1	89.331	89.331	0.159	
Latitud	1	4.720	4.720	0.420	
Longitud	1	-1.272	-1.272	0.444	
Altitud-Latitud	1	182.245	182.245	0.064	
Altitud-Longitud	1	15.543	15.543	0.407	
Latitud-Longitud	1	168.233	168.233	0.098	
Provincia Biogeográfica	15	1533.278	102.218	0.001	0.002
Error	191	19076.735	99.878		
Total	212	21068.813	99.381		

¹df: grados de libertad

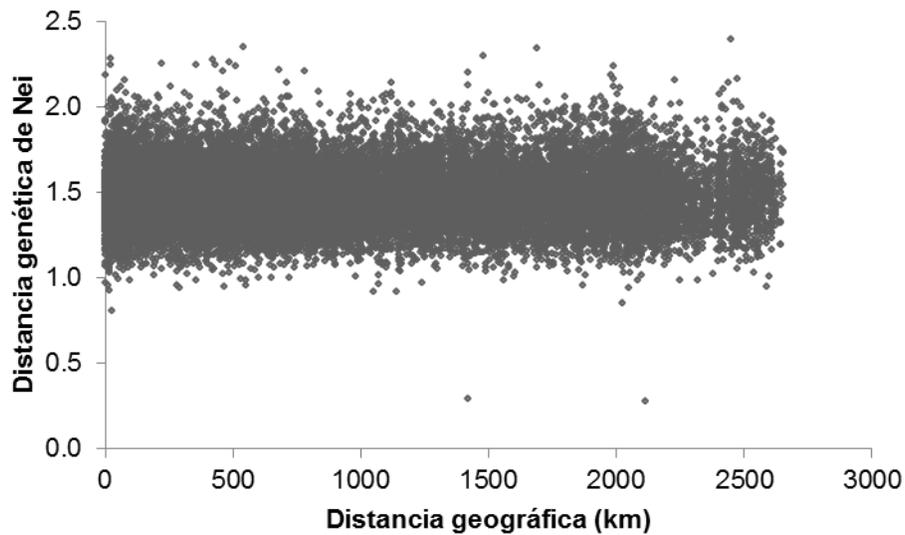
²SS: suma de cuadrados

³MS: cuadrados medios

⁴p: significancia.

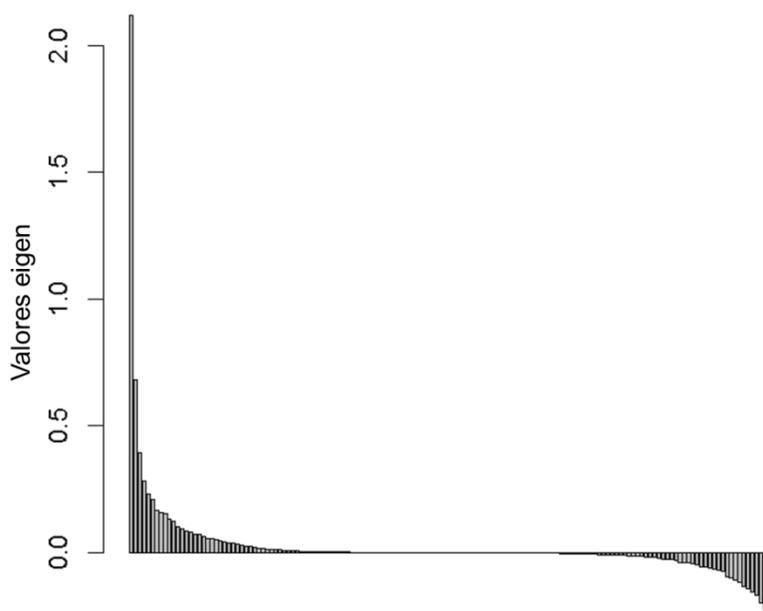


Anexo V.D. Correlación ($r_M=0.039$; $p<0.05$) entre la distancia geográfica (en kilómetros) y la distancia genética de Nei de 34 isoenzimas que representan a 259 poblaciones de 59 razas.

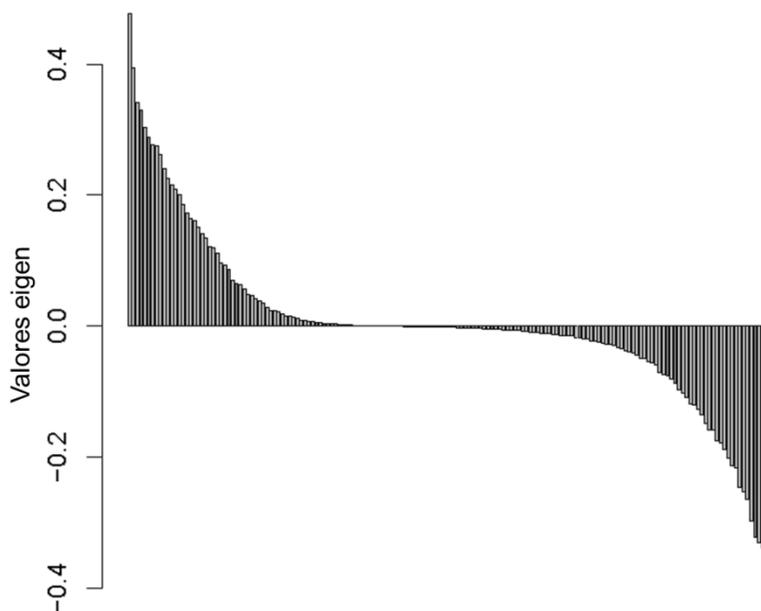


Anexo V.E. Correlación ($r_M=0.002$; $p=0.442$) entre la distancia geográfica (en kilómetros) y la distancia genética de Nei de 96 microsatélites que representan a 213 individuos de 58 razas.

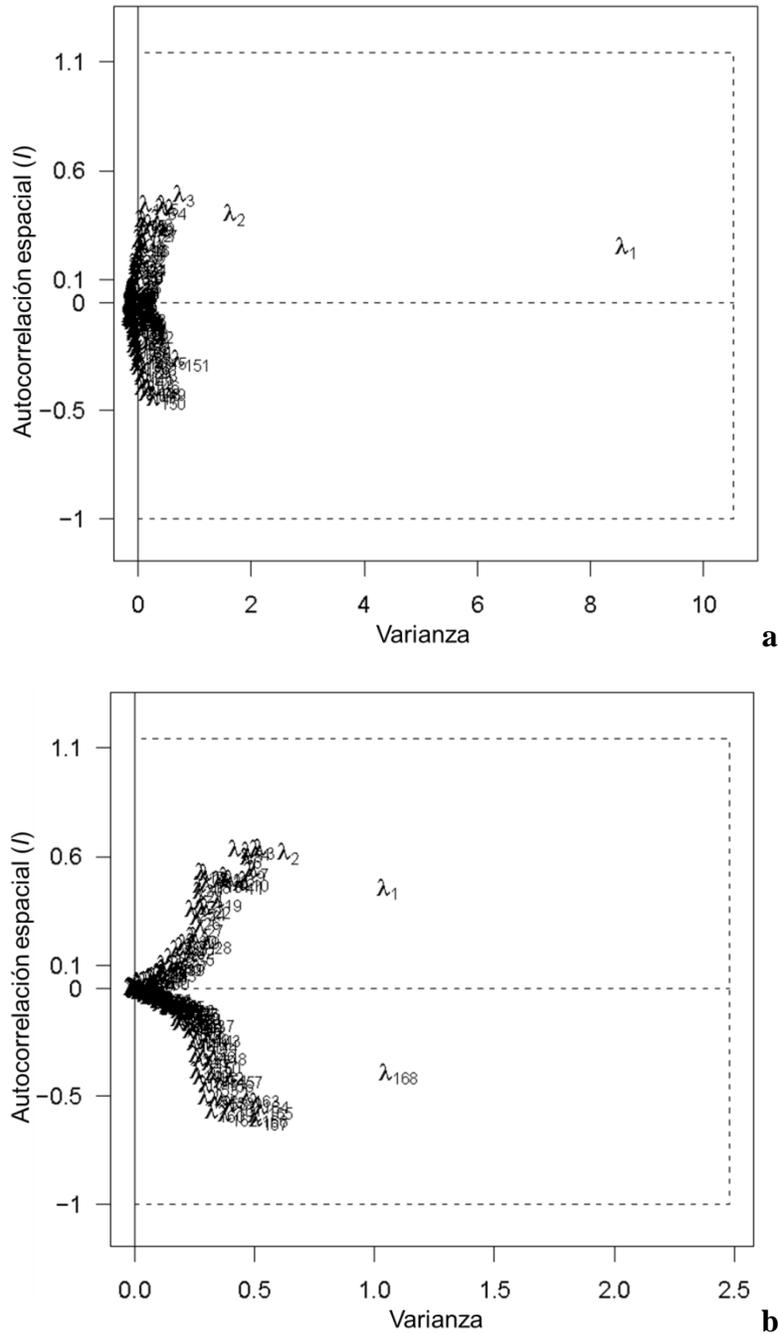
ANEXO VI



Anexo VI.A. Gráfica de barras de los valores eigen del sPCA de 153 SNPs con índice $F_{ST} \geq 0.15$ de 169 individuos de 50 razas. Los valores eigen positivos representan estructuras regionales, mientras que los negativos, estructuras locales.



Anexo VI.B. Gráfica de barras de los valores eigen del sPCA de 200 SNPs con índice $F_{ST} \leq 0.02$ de 169 individuos de 50 razas. Los valores eigen positivos representan estructuras regionales, mientras que los negativos, estructuras locales.



Anexo VI.C. Gráficas de los valores eigen en función de los componentes de varianza y de autocorrelación espacial (I de Moran) de cada uno. **a.** sPCA de 153 SNPs con índice $F_{ST} \geq 0.15$. **b.** sPCA de 200 SNPs con índice $F_{ST} \leq 0.02$. La línea punteada vertical a la derecha indica la máxima varianza obtenida con una combinación lineal de alelos en un PCA ordinario. Las líneas punteadas horizontales en los extremos indican el intervalo de variación del estadístico I de Moran de acuerdo a la matriz de pesos.

MATERIAL SUPLEMENTARIO

Análisis de las isoenzimas con un modelo bayesiano de asignación

La determinación de las agrupaciones genéticas presentes en el maíz a partir de las isoenzimas se llevó a cabo con Structure v.2.3 (Pritchard *et al.* 2000). El modelo bayesiano de asignación de este programa infiere el número de poblaciones y la frecuencia de cada alelo en las distintas agrupaciones. Los individuos son asignados a un número predeterminado de agrupaciones (K), minimizando el desequilibrio de Hardy-Weinberg y de ligamiento dentro de cada población. El algoritmo Monte Carlo de cadenas de Markov (MCMC) permite estimar las frecuencias alélicas de cada K y la proporción del genoma individual derivado de cada agrupación (q_K). La población de procedencia de los genotipos no fue usada en el análisis. El número de K varió entre 1-6 (10 repeticiones en cada una) y para cada una se generaron 1,000,000 de MCMC para estimaciones, después de haber eliminado las primeras 100,000 (burn-in). Los parámetros usados están preestablecidos por el programa con excepción de Alphapropsd, la desviación estándar de la distribución que permite establecer la proporción de mezcla, el cual se elevó de 0.05 a 0.5 para aumentar el intervalo de alfa. Debido a que este análisis es sensible a los datos faltantes, las isoenzimas con más del 50% de datos faltantes fueron eliminadas (*ACO 1, DIA 1 y 2, HEX 2*).

El número de agrupaciones se determinó al identificar la ΔK mayor (Evanno *et al.* 2005). Sin embargo, la ΔK mayor ocurrió cuando K es igual a dos, mientras que la máxima verosimilitud ($L(K)$), cuando K es igual a seis (Figura S1). Esta discrepancia entre la ΔK y la $L(K)$ también se observó en el caso de los microsatélites (Vigouroux *et al.* 2008). La ΔK se enfoca en la tasa de cambio de segundo orden de la función de verosimilitud con respecto a K y en la mayoría de los casos ha mostrado indicar el número de agrupaciones biológicas más probable (Evanno *et al.* 2005). En cambio, la $L(K)$ no necesariamente representa agrupaciones biológicas (Pritchard *et al.* 2000; Evanno *et al.* 2005). Además, el mayor nivel de mezcla (α) ocurre cuando K es igual a dos. Debido a esto, se representó gráficamente el porcentaje de pertenencia individual a cada uno de los dos grupos genéticos probables (Figura S2).

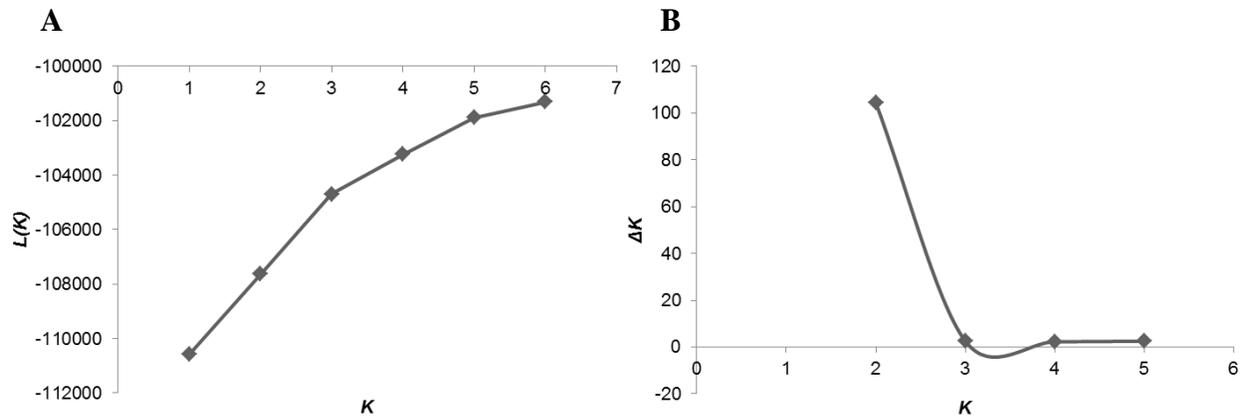


Figura S1. Gráficas del logaritmo de verosimilitud ($L(K)$; A) y de ΔK (B) del análisis de asignación bayesiana para 30 isoenzimas de 3600 individuos de 59 razas.

La proporción de pertenencia de los genotipos individuales de las isoenzimas a una de las dos agrupaciones sustentadas por la ΔK parece mostrar un patrón altitudinal. Al ordenar los individuos de menor a mayor altitud, el coeficiente de pertenencia (Q) tiende a aumentar para una agrupación genética y disminuir para la otra a lo largo de este gradiente (Figura S2). Por consiguiente, la altitud y probablemente los gradientes ambientales que se asocian a ésta podrían determinar la distribución espacial de las frecuencias alélicas en el maíz.

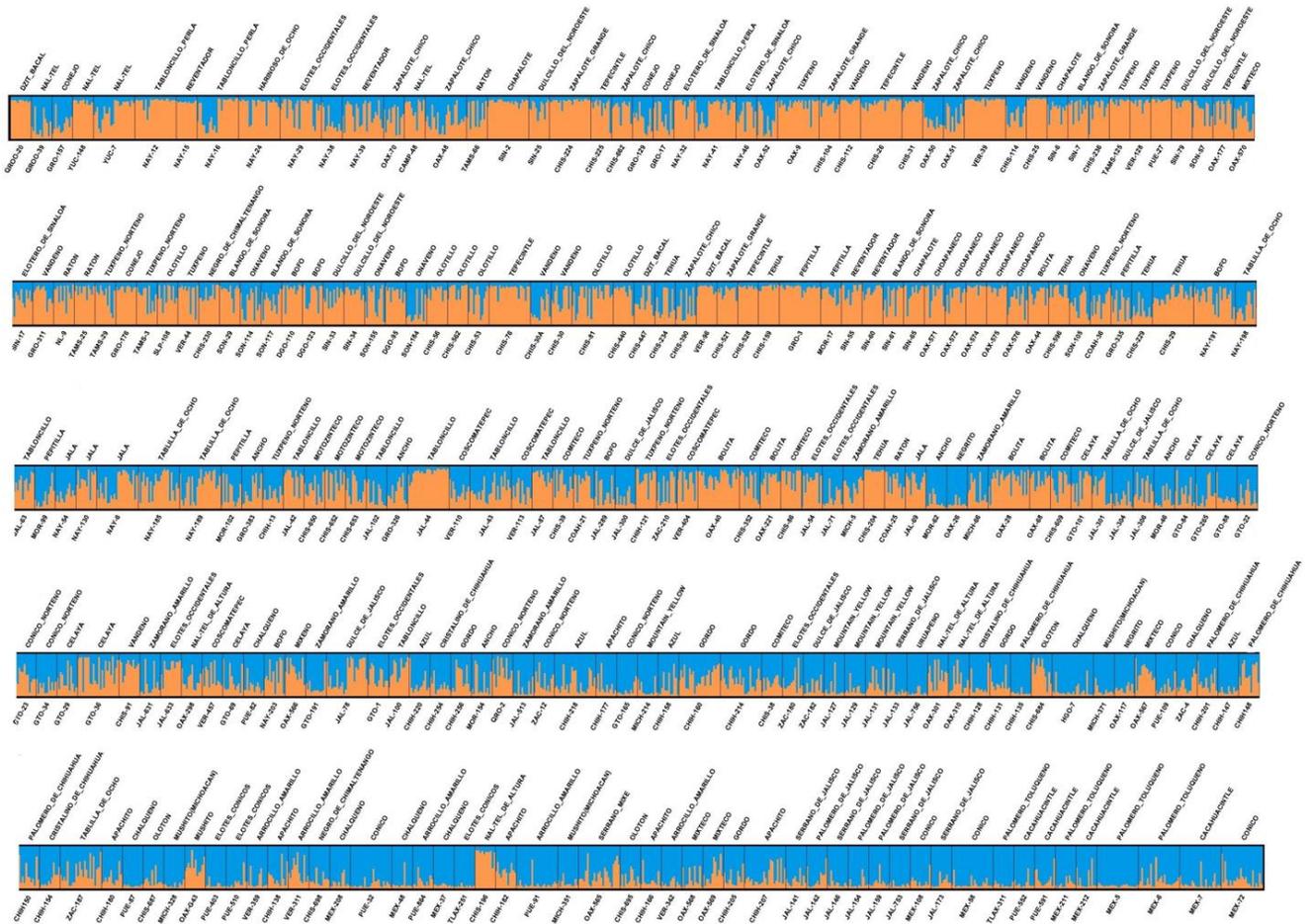


Figura S2. Gráfica del coeficiente de pertenencia (Q) de cada individuo a las dos agrupaciones genéticas ($K=2$) sustentadas por la ΔK en el análisis de asignación bayesiana de 30 isoenzimas de 59 razas. Las 259 poblaciones, cada una con 12 ó 24 individuos, se agrupan de menor a mayor altitud. Las razas se nombran arriba de la gráfica y las poblaciones, abajo.

Distribución del índice F_{ST} en isoenzimas y microsatélites

La distribución del índice F_{ST} entre razas de maíz se comparó con la de teocintles anuales (*Z. m. ssp. mexicana*, *Z. m. ssp. parviglumis*) y perennes (*Zea diploperennis*, *Zea perennis*) para establecer el grado de diferenciación genética de las isoenzimas y los microsatélites en este cultivo. Los genotipos individuales de todos los teocintles se obtuvieron del portal de Panzea (www.panzea.org). Los datos de isoenzimas incluyen a 20 marcadores de 504 individuos de 25 poblaciones (Doebley *et al.* 1984), mientras que los de microsatélites incluyen a 93 marcadores de 213 individuos (Fukunaga *et al.* 2005). Los marcadores son los mismos, aunque menos que en el maíz, y los genotipos se han reportado en la literatura.

La diferenciación genética de las 34 isoenzimas analizadas del maíz es similar a la de los teocintles anuales y perennes (Figura S3). El índice F_{ST} en el cultivo varía entre 0.05 y 0.15. En cambio, los valores de este estadístico en las 20 isoenzimas analizadas de los teocintles varían entre 0.10 y 0.15, para los anuales, y entre 0.05 y 0.20, para los perennes. La similitud en la distribución de la diferenciación genética entre el cultivo y sus parientes silvestres podría deberse a que el número de alelos por locus en estos marcadores no es alto (Anexo V.A). Sin embargo, también sugiere que los cambios en las frecuencias alélicas de las isoenzimas analizadas podrían no estar influidos por la selección natural (positiva).

La diferenciación genética de los 96 microsatélites analizados del maíz es más alta que la de los 93 marcadores de los teocintles anuales y perennes (Figura S4). La mayoría de los valores de F_{ST} se encuentran entre 0.20 y 0.30, en el cultivo, entre 0.05 y 0.15, en los teocintles anuales, y entre 0.12 y 0.20, en los teocintles perennes. La mayor diferenciación genética en el maíz podría deberse al alto número de alelos por locus (Anexo V.A) y podría sugerir que estos marcadores son adecuados para analizar el efecto de los barridos selectivos en la variación neutral de este cultivo. Es posible que algunos de estos marcadores se localicen cerca de loci en el genoma influidos por la selección natural (positiva). No obstante, la estructura espacial encontrada con los microsatélites muestra un patrón de aislamiento por distancia, ya que la mayoría de estos marcadores probablemente se localizan en loci neutrales.

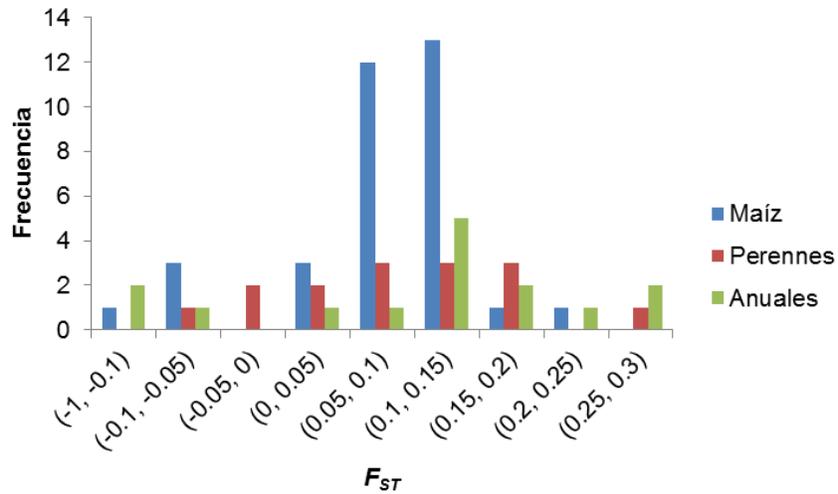


Figura S3. Distribución del índice F_{ST} de isoenzimas obtenido entre razas de maíz (*Z. m. ssp. mays*), entre teocintles anuales (*Z. m. ssp. mexicana* y *Z. m. ssp. parviglumis*) y entre teocintles perennes (*Z. diploperennis* y *Z. perennis*).

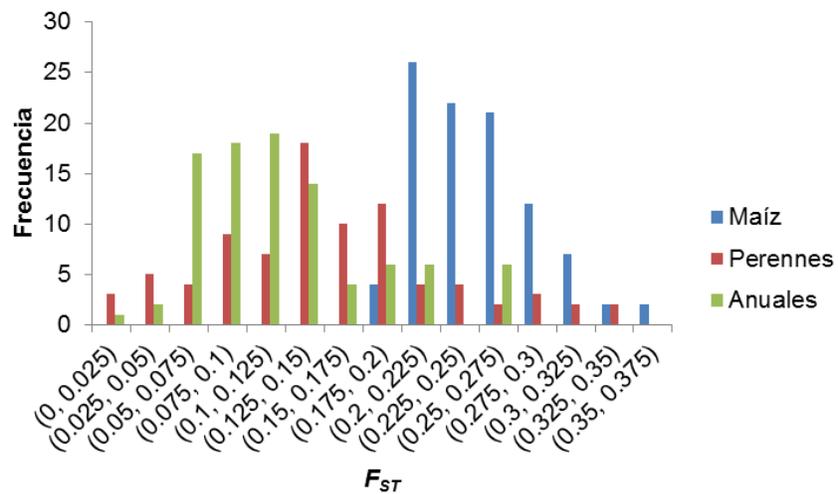


Figura S4. Distribución del índice de F_{ST} de microsatélites obtenido entre razas de maíz (*Z. m. ssp. mays*), entre teocintles anuales (*Z. m. ssp. mexicana* y *Z. m. ssp. parviglumis*) y entre teocintles perennes (*Z. perennis* y *Z. diploperennis*).