



UNIVERSIDAD AUTÓNOMA CHAPINGO

DEPARTAMENTO DE FITOTECNIA

INSTITUTO DE HORTICULTURA

**RELACIÓN ENTRE LA DISTANCIA GENÉTICA Y
GEOGRÁFICA EN MAÍCES TUXPEÑO Y CÓNICO
NORTEÑO EN MÉXICO**

TESIS

QUE COMO REQUISITO PARCIAL
PARA OBTENER EL GRADO DE:

MAESTRO EN CIENCIAS EN BIOTECNOLOGÍA AGRÍCOLA

PRESENTA:

NOEMÍ GUADALUPE ORTEGA JIMÉNEZ

DIRECTOR DE TESIS: DR. JUAN PORFIRIO LEGARIA SOLANO



Chapingo, Estado de México, febrero de 2017



RELACIÓN ENTRE LA DISTANCIA GENÉTICA Y GEOGRÁFICA EN MAÍCES TUXPEÑO Y CÓNICO NORTEÑO EN MÉXICO

Tesis realizada por **NOEMÍ GUADALUPE ORTEGA JIMÉNEZ** bajo la dirección del comité asesor indicado, aprobada por el mismo y aceptada como requisito parcial para obtener el grado de:

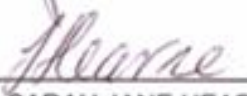
MAESTRO EN CIENCIAS EN BIOTECNOLOGÍA AGRÍCOLA

COMITÉ ASESOR

DIRECTOR:


DR. JUAN PORFIRIO LEGARIA SOLANO.

CO-DIRECTOR:


DRA. SARAH JANE HEARNE.

ASESOR:


DR. JAIME SAHAGUN CASTELLANOS.

Chapingo, Estado de México, febrero de 2017

DEDICATORIA

A mi madre: gracias a sus esfuerzos, paciencia, amor y sabios consejos me ha formado como una persona de bien y me ha enseñado que para alcanzar las metas hay que trabajar y esforzarse.

A mi abuelito: José Jiménez Medina, que se nos adelantó en el camino, que desde el cielo me cuida.

¡A todas aquellas personas y familiares que han hecho posible esta meta y han confiado en mí para realizarla!

AGRADECIMIENTOS

A mi comité asesor: la Dra. Sarah Hearne, no solamente por animarme e inspirarme a obtener este grado, también por su supervisión, consejo y apoyo en todos los aspectos a través de este periodo, al Dr. Legaría y al Dr. Sahagún, quien junto con la Dra. Sarah hicieron posible realizar esta maestría aprovechando mi trabajo en CIMMyT, mil gracias a los tres.

Al Dr. Kai Sonder y su grupo de investigadores del programa de GIS, en CIMMyT, por su apoyo técnico y disposición siempre.

Al Dr. Juan Burgueño, al Dr. Jorge Franco y sus colaboradores del programa de Bioestadística, en CIMMyT, por su apoyo y colaboración.

Gracias a CONACyT (Consejo Nacional de Ciencia y Tecnología) por apoyar a la juventud mexicana a continuar esforzándose por ser parte de la investigación, concediéndome financiamiento para cursar mis estudios de Maestría en Ciencias.

Mi gratitud también es para el CIMMyT, por permitirme realizar este proyecto e involucrarme en estudios de diversidad genética de maíz, que me ha dado la oportunidad de conocer a muchos investigadores de importancia científica y personal que siempre han tenido sus puertas abiertas, su tiempo y disposición para fomentar el análisis y aumento de conocimiento. En Chapingo, gracias nuevamente al Dr. Legaria y al Dr. Sahagún de quienes he recibido colaboración y apoyo, desde leer parte de mi escrito hasta ser mi conexión directa con la Universidad, respondiendo mis dudas, ayudándome a cumplir requisitos y documentos a tiempo. Muchas son las personas que me han ayudado y apoyado en los aspectos técnicos de la realización de esta Tesis, gracias a todos.

DATOS BIOGRÁFICOS



NOEMÍ GUADALUPE ORTEGA JIMÉNEZ

La autora de la presente tesis nació el 05 de junio de 1989, en el municipio de Ecatepec de Morelos, Estado de México, México. En el año 2011, obtuvo el título de Ingeniero en Biotecnología, con el proyecto “IDENTIFICACIÓN DE ALELOS DEL GEN *opaco-2*, EN CRUZAS DE MAICES DE CALIDAD PROTEINICA (QPM) MEDIANTE MARCADORES MOLECULARES”. En 2012 colaboró en el proyecto SEEDS OF DISCOVERY, para la genotipificación de material del banco de germoplasma de CIMMyT. Para 2015 ingresó a la Maestría en Ciencias en Biotecnología Agrícola en el Departamento de Fitotecnia de la Universidad Autónoma Chapingo. En febrero de 2017 obtuvo el grado de Maestro en Ciencias en Biotecnología Agrícola con la tesis titulada “RELACIÓN ENTRE LA DISTANCIA GENÉTICA Y GEOGRÁFICA EN MAÍCES TUXPEÑO Y CÓNICO NORTEÑO EN MÉXICO”

ÍNDICE DE CONTENIDO

1. INTRODUCCIÓN.....	1
1.1. Antecedentes.....	1
1.2. Objetivo general	2
1.2.1. Objetivos específicos	3
1.3. Hipótesis	3
1.3.1. Justificación	4
2. REVISIÓN DE LITERATURA	5
2.1. Importancia del maíz en el Mundo	5
2.2. Importancia del maíz en México	6
2.3. Distribución geográfica	7
2.4. Variabilidad del maíz en México y su clasificación racial	8
2.5. Las razas del maíz	12
2.6. Descripción de las razas	13
2.6.1. La raza Tuxpeño	14
2.6.2. La raza Cónico Norteño	16
2.7. Base genética y molecular de la relación entre la raza Tuxpeño y Cónico Norteño	18
2.8. Fuentes de variación genética (ADN)	20
2.9. Marcadores moleculares y sus contribuciones en estudios con plantas	21
2.10. Genotipificación por secuenciación GBS.....	23
2.11. Distancias Genéticas	29
2.12. Distancias Geográficas	30

2.13. Herramientas de análisis y visualización para datos genéticos y geográficos	31
3. MATERIALES Y MÉTODOS.....	33
3.1. Material vegetal	33
3.2. Colecta de material vegetal y extracción ADN.....	33
3.3. Análisis genético-molecular	34
3.4. Análisis geográfico.....	35
3.5. Análisis de correlaciones	36
4. RESULTADOS Y DISCUSIÓN	38
4.1. Análisis	38
4.1.1. Análisis genético-molecular.....	38
4.1.2. Análisis geográfico	44
4.2. Análisis comparativo entre distancias genéticas y geográficas.....	49
5. CONCLUSIONES	65
6. LITERATURA CITADA	66
7. APÉNDICES	77
Apéndice 1. Listado de las 830 accesiones de la raza Tuxpeño evaluadas en este estudio	77
Apéndice 2. Listado de las 1135 accesiones de la raza Cónico Norteño evaluadas en este estudio	78
Apéndice 3. Ejemplo de matriz de distancias calculadas mediante el método modificado de Roger (1983).	79
Apéndice 4. Ejemplo de matriz de salida del Análisis de Componentes Principales (PCA).	79

Apéndice 5. Ejemplo de matriz Cluster (análisis de asociación).....	80
Apéndice 6. Filtrado de la Información rescatada de la base de datos Germinate 3.....	80
Apéndice 7. Ejemplo de matriz de distancias puras.	81
Apéndice 8. Ejemplo de matriz de identificación de distancias calculadas reales.....	81
Apéndice 9. Ejemplo de matriz de distancias calculadas reales, determinando la longitud en metros y kilómetros, respectivamente.	82
Apéndice 10. Ejemplo de archivo excel para transformar en un archivo de extensión kml.....	82

LISTA DE CUADROS

Cuadro 1. Razas catalogadas en México. *Entre paréntesis está el número de colectas registradas en el catálogo LAMP (1991). **Los grupos están como se describe en Sánchez y Goodman (1992a).....	9
Cuadro 2. Descripción de complejos raciales de maíz.	13
Cuadro 3. Clasificación de los marcadores, ventajas y desventajas Schlötterer (2004): Aranguren-Méndez <i>et al.</i> (2005): Bedoya (2012).....	19
Cuadro 4. Ejemplo de información genotípica en formato numérico.	39

LISTA DE FIGURAS

Figura 1. Principales países productores de maíz (grano), 2015-2016, en millones de toneladas	6
Figura 2. Distribución geográfica de regiones donde se cultiva maíz, los puntos negros en el mapa son las zonas de mayor producción según lo describe Serratos, 1997.....	8
Figura 3. Análisis de componentes principales, basada en caracteres aloenzimáticos que contradicen la supuesta relación entre la raza Tuxpeño, Cónico y Chalqueño.....	19
Figura 4. Proceso para la generación de GBS-DArTseq, Hi-Seq (SAGA, 2016)	25
Figura 5. Dispersión de los genotipos correspondientes a las razas Tuxpeño y Cónico Norteño considerando los dos primeros componentes principales.....	40
Figura 6. Dendrograma de relaciones entre los 1965 genotipos evaluados, con una línea de corte a tres grupos (rojo) y cinco grupos (verde)	41
Figura 7. Representación de las distancias genéticas entre accesiones de las razas Tuxpeño y Cónico Norteño por el método MDS, mediante el software Curly Whirly	43
Figura 8. Página de la base de datos Germinate 3.....	44
Figura 9. Mapa de distribución geográfica de los sitios de colecta, generado en Google Earth.....	46
Figura 10. Mapa de la distribución geográfica de los sitios de colecta, generado en ArcMap	46
Figura 11. Distribución de las razas en la República Mexicana. A, Tuxpeño y B, Cónico Norteño, según lo descrito por Wellhausen y colaboradores en 1951.....	47

Figura 12. Mapa de distribución geográfica de los sitios de colecta de la raza Chalqueño, generado en ArcMap	48
Figura 13. Comparación de distancias genéticas entre las razas de maíz Tuxpeño y Cónico Norteño. A, como un conjunto. B, raza Tuxpeño ordenados de 0-830. C, raza Cónico Norteño ordenados de 0-1135	50
Figura 14. Imagen de distancias geográficas. A, como un conjunto de datos ordenados de 0 -1964. B, raza Tuxpeño ordenados de 0-830. C, raza Cónico Norteño ordenados de 0-1135.....	51
Figura 15. Histograma de distribución de frecuencias de distancias genéticas. A, datos ordenados de 0 -1964. B, distribución de frecuencias de distancias genéticas; raza Tuxpeño ordenados de 0-830. C, distribución de frecuencias de distancias genéticas; raza Cónico Norteño ordenados de 0-1135	52
Figura 16. Histograma de distribución de frecuencias de distancias geográficas. A, datos ordenados de 0 -1964. B, distribución de frecuencias de distancias geográficas, raza Tuxpeño ordenados de 0-830. C, distribución de frecuencias de distancias geográficas; raza Cónico Norteño ordenados de 0-1135	52
Figura 17. Gráfica de correlación entre las distancias genéticas (Y) y las distancias geográficas (X). A, del conjunto de datos. B, gráfica de correlación raza Tuxpeño. C, raza Cónico Norteño	53
Figura 18. Agrupamiento de accesiones de las razas Tuxpeño y Cónico Norteño. A, con base en el MDS de las distancia geográfica. B, con base en el MDS de las distancias genéticas. Rojo: Cónico Norteño. Negro: Tuxpeño.....	56
Figura 19. Agrupamiento de accesiones de las razas Tuxpeño y Cónico Norteño. A y B, gráficos con base en el MDS de las distancia geográfica. A. Raza Tuxpeño. B. Raza Cónico Norteño. C y D, gráficos con base en el MDS	

de las distancias genéticas. C. Raza Tuxpeño. D. Raza Cónico Norteño	56
Figura 20. Rotacion de coordenadas en el análisis procrustes, aplicado con la técnica MDS. A, del conjunto de datos. B, raza Tuxpeño. C, raza Cónico Norteño	59
Figura 21. Residuos del análisis procrustes aplicado con la técnica MDS. A, del conjunto de datos. B, residuales de la raza Tuxpeño. C, residuales de la raza Cónico Norteño	60
Figura 22. Histograma de residuales del análisis procrustes. A, conjunto de datos. B, raza Tuxpeño. C, raza Cónico Norteño	61
Figura 23. Gráfico de residuales en un mapa de la República Mexicana. A, del conjunto de datos. B, raza Tuxpeño. C, raza Cónico Norteño. Los puntos mas oscuros indican una mayor proximidad con el resto, por lo tanto los de color más claro muestran las relaciones menos cercanas entre la población	62
Figura 24. Distancias geográficas promedio para las accesiones de las razas de maíz. A, del conjunto de datos. B, raza Tuxpeño. C, raza Cónico Norteño. Los puntos mas oscuros indican una mayor proximidad con el resto, por lo tanto los de color mas claro muestran las relaciones mas cercanas entre la población.	63
Figura 25. Distribución genética promedio de las accesiones de las razas de maíz. A, del conjunto de datos. B, raza Tuxpeño. C, raza Cónico Norteño. Los puntos más oscuros indican una mayor proximidad con el resto, por lo tanto los de color más claro muestran las relaciones más cercanas entre la población	64

RELACIÓN ENTRE LA DISTANCIA GENÉTICA Y GEOGRÁFICA EN MAÍCES TUXPEÑO Y CÓNICO NORTEÑO EN MÉXICO

RESUMEN

La cuantificación e identificación de la variación genética de las accesiones es útil dentro de los bancos de germoplasma para usarse en los programas de mejora genética como una estrategia importante en el mantenimiento de los cultivares. La diversidad genética-molecular de accesiones de maíz que se encuentran en algunos de los principales ambientes de producción en México se investigó mediante genotipado por secuenciación (DArTseq) de alta calidad. Los marcadores DArTseq mostraron nuevas variaciones de polimorfismo de nucleótido único (SNP) en variedades locales que estuvieron asociadas a adaptación de ambientes estresantes por sequía y calor. Esto puede abrir nuevas vías para el pre-mejoramiento, al enriquecer el germoplasma élite con nuevos alelos para tolerancia al estrés abiótico. La raza de maíz tropical Tuxpeño es muy conocida porque ha contribuido en gran medida al desarrollo de los bancos de germoplasma de maíces tropicales y subtropicales. Con el fin de investigar la forma en que podría ser explotada esta variación genética existente en los bancos de germoplasma para una mejora futura de maíz, un conjunto de muestras fue caracterizado mediante el uso de marcadores SNP obtenidos por DArTseq. El análisis incluyó accesiones de las razas Tuxpeño y Cónico Norteño de la colección localizada en el banco de germoplasma del Centro Internacional de Mejoramiento de Maíz y Trigo (CIMMYT), evaluando un total de 1965 entradas con 30 plantas cada una. Se aplicó un enfoque de combinación entre los datos geográficos del pasaporte del banco de germoplasma y datos moleculares provenientes del genotipado de estas dos razas de maíz. El análisis individual de las razas mostró una mayor correlación entre distancias genéticas y geográficas para la Cónico Norteño que para Tuxpeño.

Palabras clave: *Zea mays*; estrés hídrico; genotipificación por secuenciación.

RELATIONSHIP BETWEEN GENETIC AND GEOGRAPHIC DISTANCES IN TUXPEÑO Y CÓNICO NORTEÑO MAIZE IN MÉXICO

ABSTRACT

Quantification and identification of genetic variation of accessions are useful within germplasm banks for use in genetic breeding programs as an important strategy in the maintenance of cultivars. The genetic-molecular diversity of maize accessions that are found in some of the major production environments in Mexico was investigated by genotyping by sequencing (DArTseq) of high quality. DArTseq marks showed new variations of single nucleotide polymorphism (SNP) in local varieties that were associated with adaptation to stressful drought and heat environments. This may open new pathways for pre-breeding, enriching elite germplasm with alleles for abiotic stress tolerance. The tropical maize Tuxpeño race is well known because it has contributed greatly to the development of tropical and subtropical maize germplasm banks. In order to investigate how this genetic variation existing in germplasm banks could be exploited for future maize improvement, a set of samples was characterized by use of SNP markers obtained from DArTseq. The analysis included accessions of the Tuxpeño and Cónico Norteño races from the collection located in the International Maize and Wheat Improvement Center (CIMMYT) Germplasm Bank, evaluating a total of 1965 entries with 30 plants each. An approach combining the geographic data of the passport of the germplasm bank and the molecular data from the genotyping of these two maize races was applied. The individual analysis of the races showed a higher correlation between genetic and geographic distances for Cónico Norteño than for Tuxpeño.

Keywords: *Zea mays*, water stress, genotyping by sequencing

1. INTRODUCCIÓN

1.1. Antecedentes

El maíz (*Zea mays L.*) es el cereal más utilizado bajo diversos contextos. Este cereal forma parte de la alimentación diaria de los mexicanos, constituye un factor de suma importancia en la ganadería y se usa para la fabricación de múltiples productos industriales. Es el cultivo de mayor impacto en el país, desde el punto de vista alimentario, económico, político y social. México es centro de origen y diversidad de este cereal. Por otra parte, posee una importante fuente de variación genética útil para su mejoramiento.

No obstante, la gran inversión del gobierno mexicano en el campo y años de influencia de los programas de mejoramiento formal en la introducción masiva de variedades mejoradas de maíz, existen hoy día agricultores que continúan cultivando variedades ancestrales de maíz criollo de manera aislada y han conservado sus características de orden morfológico, ecológico, genético y cultural, de tal modo que es posible diferenciarlas.

En el banco de germoplasma de maíz del CIMMyT existen cerca de 28,000 muestras de semilla de maíz y entre ellas se encuentran algunas variedades criollas pertenecientes a diferentes razas. Mucho de este germoplasma ya no se siembra en los terrenos de los agricultores, pero ha servido a investigadores como base para la mejora de variedades que se utilizan en la actualidad. El análisis de la diversidad genética en correlación al área geográfica en que se localizan las razas de maíz en México, en especial Tuxpeño y Cónico Norteño, es de gran importancia tanto para definir sus atributos, así como para su empleo en el mejoramiento genético, debido a que dichas razas se encuentran ampliamente distribuidas en el país, mostrando sobresaliente calidad agronómica en planta y en resistencia a condiciones climáticas adversas como la sequía y el calor.

Esta investigación fue parte del proyecto 'Seeds of Discovery' de CIMMyT, cuyo propósito es el desarrollo de variedades de maíz de alto rendimiento que aborden los desafíos futuros del cambio climático y el incremento de la población mundial.

La información que genere este estudio puede ser útil para seleccionar accesiones del banco de germoplasma y aprovecharlas por su potencial en programas de pre-mejoramiento y explotación de alelos de genes candidatos para resistencia a sequía y al estrés por calor.

El mejoramiento genético convencional ha contribuido en buena parte a la reducción de la variabilidad genética de los cultivos de maíz. Además, a medida que se consiguen mayores perfiles de rendimiento mediante sofisticados proyectos de mejoramiento se han observado ascendentes niveles de vulnerabilidad genética ante la diversificación del clima y susceptibilidad a enfermedades, lo cual amenaza a la seguridad alimentaria.

En este contexto, las variedades nativas pueden constituir una fuente importante de genes útiles para ampliar la variabilidad genética explotable para desarrollar mejores variedades. Actualmente existe gran cantidad de germoplasma disponible en los bancos, sin embargo, esta es escasamente explotada, por esta razón se busca evidenciar el potencial y posible aprovechamiento de estos recursos genéticos para garantizar la seguridad alimentaria ante los efectos de la sequía y de las temperaturas altas.

En este contexto, se plantearon los objetivos siguientes para el presente proyecto de investigación:

1.2. Objetivo general

Analizar la relación entre la distancia genética y geográfica en maíces de las razas Tuxpeño y Cónico Norteño, con datos obtenidos mediante tecnología de próxima generación (NGS), genotipificación por secuenciación (DArTseq) y datos geográficos procedentes del pasaporte, almacenados en el banco de Germoplasma del CIMMyT en México.

1.2.1. Objetivos específicos

Realizar matrices de distancia genética por el método modificado de Roger con base en las secuencias obtenidas mediante DArTseq, calculando matrices de disimilitud (MRD), análisis de componentes principales (PCA), escalamiento multidimensional (MDS) y análisis de asociación (clúster).

Identificar datos geográficos (latitud, longitud) de cada accesión a evaluar, mediante la base de datos Germinate 3, así como herramientas de visualización tales como Google earth y ArcGIS, para calcular distancias puras entre accesiones y transformar a distancias físicas.

Evaluar la relación existente entre distancias genéticas y geográficas de las accesiones de las razas Tuxpeño y Cónico Norteño usando un análisis de Mantel y un análisis Procrustes.

1.3. Hipótesis

Estadísticamente existe una relación positiva entre las distancias genética y geográfica, según esta, entre más corta sea la distancia geográfica entre dos razas, menor será la distancia genética, en contraste entre mayor es la distancia geográfica mayor será la distancia genética.

1.3.1. Justificación

La importancia que tiene el estudio de las razas Tuxpeño y Cónico Norteño radica en las características endémicas de estas razas, asociadas a la tolerancia que han mostrado a cambios climáticos como la carencia de agua o sequía durante los últimos 50 años, por ello con este trabajo se pretende analizar formas de explorar esta característica genética. Al conocer el germoplasma disponible en los bancos, escasamente explotado, se busca resaltar su potencial y aprovechamiento, que es de suma importancia en el mejoramiento actual de las plantas para garantizar la seguridad alimentaria ante los efectos del cambio climático particularmente aquellos que provocan sequías y temperaturas altas.

El análisis de relaciones entre poblaciones de plantas implica evaluar distancias genéticas y distancias geográficas. Estas se han utilizado muy frecuentemente debido a que han demostrado ser las que mejor representan las relaciones poblacionales, a partir de datos generados por computadora, esto con la finalidad de comprobar las teorías acerca de que a mayor distancia geográfica la distancia genética será afectada proporcionalmente. Los resultados de esta investigación podrán ser utilizados como una guía para la selección de accesiones de estas dos importantes razas para el mejoramiento de maíz y además servirán para diseñar programas de conservación ya sea *in situ* o *ex situ*. Los resultados de este trabajo también darán idea de si se da o no prioridad a la conservación *in situ* y/o se decide por la recopilación de nuevos materiales para la conservación *ex situ* en zonas donde la raza o la diversidad genética se ven amenazadas. De esta manera es considerada una estrategia óptima para la elección adecuada de lugares de colecta, así como para analizar el número adecuado y cercanía entre colectas.

2. REVISIÓN DE LITERATURA

2.1. Importancia del maíz en el mundo

El maíz es una de la plantas cultivadas que posee un nivel elevado de domesticación, perfeccionado mediante la selección masal conducida por miles de generaciones, que originó una especie totalmente dependiente del hombre, debido a que la transformación eliminó características de sobrevivencia en la naturaleza, pero a su vez generó una gran variedad de maíces (más de 300 razas y miles de variedades) adaptadas a los ambientes más diversos y a las preferencias de sus cultivadores (Paterniani, 2000).

Actualmente, el maíz es el segundo cultivo a nivel mundial debido a su producción, después del trigo, y es considerado el primer cereal en rendimiento de grano por hectárea y de gran importancia económica como alimento humano, alimento para ganado o fuente de un gran número de productos industriales. El Departamento de Agricultura de Estados Unidos (USDA, 2016) estima que la Producción Mundial de Maíz del 2015 fue de 968.86 millones de toneladas y para 2016/2017 será de 1011.07 millones de toneladas. Los 1011.07 millones de toneladas estimados este año podrían significar un aumento de 42.21 millones de toneladas o un 4.36 % en la producción de maíz alrededor del mundo.

La diversidad de ambientes en los que se cultiva es mayor que para cualquier otro cultivo, hoy día se siembra hasta los 58° de latitud norte en Canadá, en Rusia y hasta los 40° de latitud sur en Argentina y Chile. La mayor parte del maíz se siembra a altitudes medias, pero también por debajo del nivel del mar en las planicies del Caspio y hasta los 3 800 msnm en la cordillera de los Andes (Paliwal, 2011).

2.2. Importancia del maíz en México

En México, el maíz forma parte de la alimentación diaria, suministra el 50 % de las calorías consumidas por la población mexicana. Es el cultivo de mayor presencia en el país, constituyendo un insumo para la ganadería y para la obtención de numerosos productos industriales, por lo que, desde el punto de vista alimentario, económico, político y social, es el cultivo agrícola más importante en el país, ya que es el quinto productor (Figura 1) de maíz a nivel mundial con 25,000,000 toneladas anuales reportadas en 2015 (SIAP, 2016).

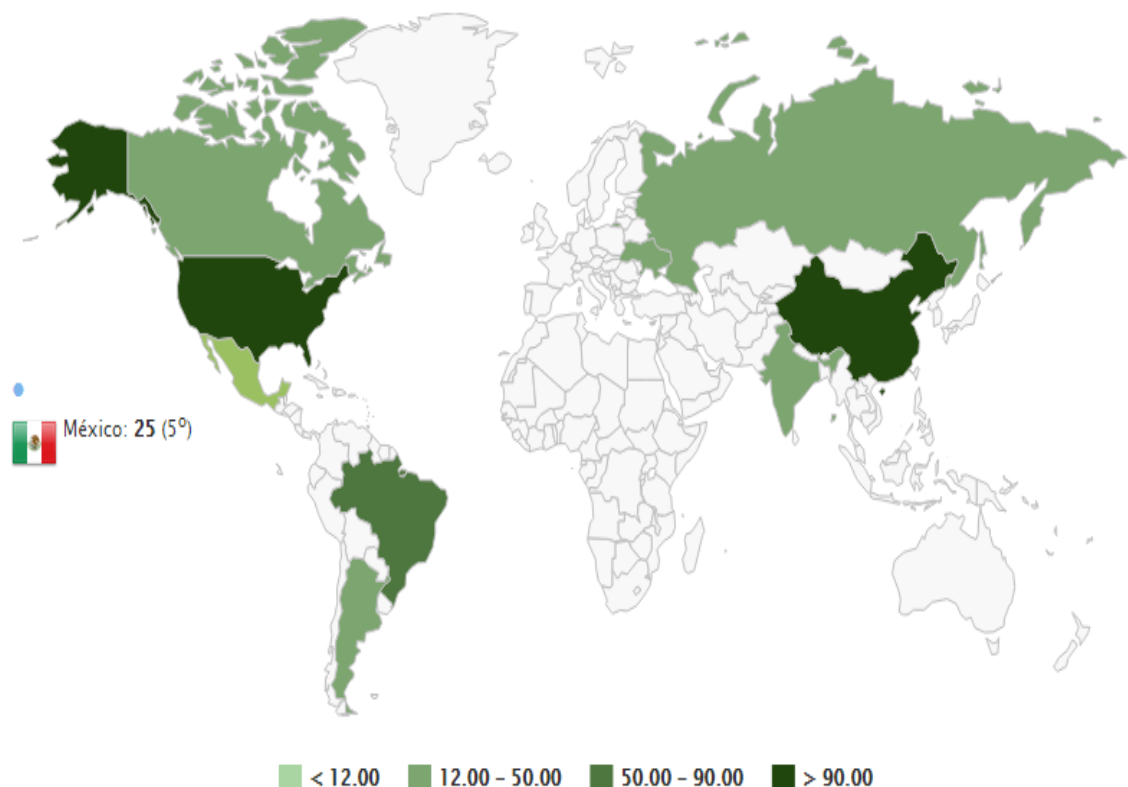


Figura 1. Principales países productores de maíz (grano), 2015-2016, en millones de toneladas.

México es el centro de origen y diversidad del maíz. Se reconoce que los productores mexicanos conservan la diversidad del cultivo por razones sociales,

económicas y culturales y las razas locales muestran un comportamiento agronómico superior al de las mejoradas (Bellon, 2002).

El maíz en México se siembra en un amplio rango de climas y altitudes, desde el nivel del mar hasta los 3,400 msnm, cultivándose también en zonas tórridas con escasa precipitación, en regiones templadas, en las faldas de las montañas, en ambientes muy cálidos y húmedos, en suelo escaso, en laderas pronunciadas o en valles fértiles, en diferentes épocas del año y bajo múltiples sistemas de manejo y desarrollo tecnológico (Hernández, 1985b).

2.3. Distribución geográfica

El maíz se cultiva en zonas extensas del mundo desde Estados Unidos, China, Brasil, México, Argentina, India, Indonesia, Francia, Sudáfrica, Ucrania, Canadá, Italia, Hungría, Rumania, Nigeria, Filipinas, Rusia, Egipto, Serbia y Alemania.

En América Latina (Figura 2), regiones como el centro de México, Honduras, Nicaragua, Venezuela, Ecuador, Perú, Bolivia, Brasil y Argentina son las zonas de mayor producción de maíz en la actualidad. También, en zonas de Estados Unidos, Guatemala, Cuba, Colombia, Venezuela, Ecuador, Perú, Chile, Bolivia, Argentina, Paraguay, Uruguay, pero principalmente en México se ha colectado la mayor cantidad de germoplasma nativo (Serratos, 1997).

Estudios varios sugieren que las variedades de maíz más antiguas sobrevivientes son de las tierras altas de México, que se han extendido a lo largo de dos grandes rutas, cotejado con registros arqueológicos que indican la diversificación en las tierras altas de México extendiéndose a las tierras bajas de Centroamérica y a Sudamérica.

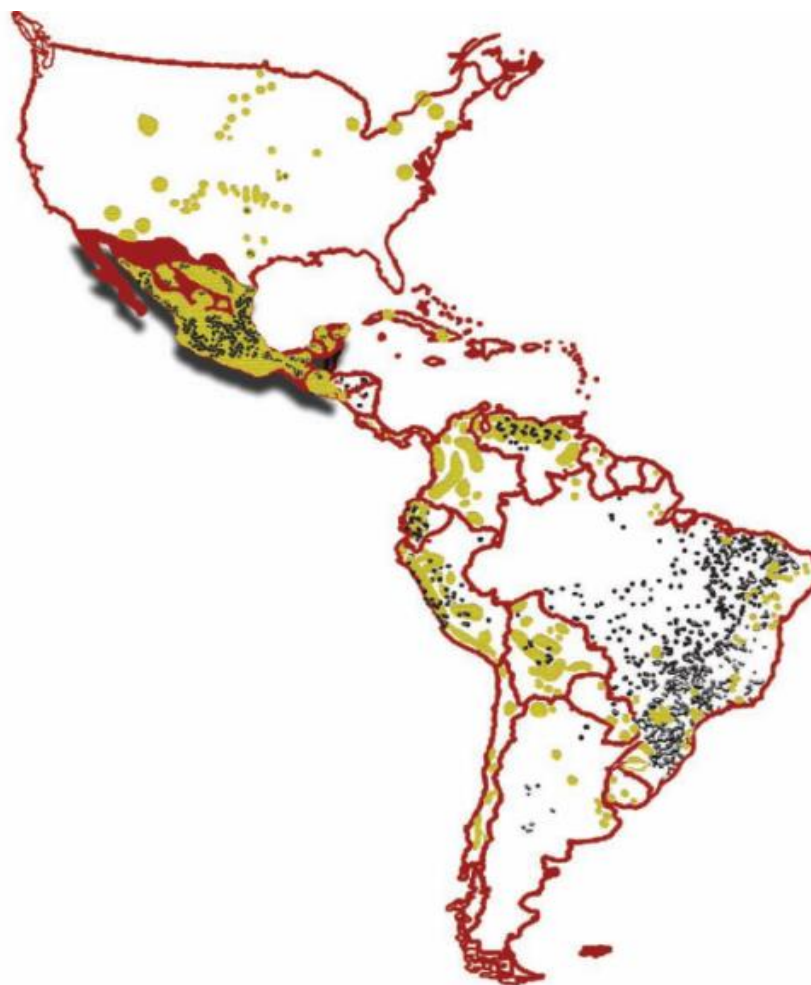


Figura 2. Distribución geográfica de regiones donde se cultiva maíz, los puntos negros en el mapa son las zonas de mayor producción según lo describe Serratos, 1997.

2.4. Variabilidad del maíz en México y su clasificación racial

En diferentes regiones de México, el cultivo de maíz por los pequeños agricultores es altamente dinámico debido a diversos factores como el manejo humano y el ambiente. Entre los primeros se destaca el intercambio de semillas (Badstue *et al.*, 2006) así como la selección que hacen los agricultores de caracteres fisiológicos, de planta y mazorca (Soleri y Cleveland, 2001; Pressoir y Berthaud, 2004b). En los segundos esta la adaptación natural y gradientes altitudinales (Muñoz *et al.*, 2002; Ruíz-Corral *et al.*, 2008; Mercer *et al.*, 2008).

Cárdenas y Taba (1995) catalogan las razas de maíz conforme a sus caracteres fenotípicos, morfológicos y de localidad sobresalientes como Celaya, Cónico, Chalqueño, Maíz dulce, Tabloncillo, Harinoso de ocho, Olotillo, entre otros, posteriormente en 1992, Sánchez y Goodman los agrupan por estados (Cuadro 1), marcando el número de colectas registradas por cada estado según LAMP en 1991.

Cuadro 1. Razas de maíz catalogadas en México según Cárdenas F (Taba, 1995a).

Estado*	Razas catalogadas de maíz
Aguascalientes (59)	Celaya, Cónico, Cónico Norteño, Chalqueño, Elotes Cónicos
Baja California Sur (11)	Tuxpeño, Tabloncillo Perla
Campeche (182)	Dzit-Bacal, Nal-Tel, Clavillo
Chihuahua (348)	Tuxpeño, Celaya, Cónico, Cónico Norteño, Chalqueño, Tabloncillo, Reventador, Tabloncillo Perla, Bolita, Maíz Dulce, Harinoso de Ocho, Palomero, San Juan, Dulcillo del Noroeste, Tuxpeño Norteño, Azul, Lady Finger, Blandito, Cristalino de Chihuahua, Gordo, Tehua, Apachito, Maizon
Chiapas (795)	Tuxpeño, Celaya, Cónico, Elotes Occidentales, Olotillo, Tabloncillo Perla, Dzit-Bacal, Vandeño, Nal-Tel, Tepecintle, Oloton, Zapalote Chico, Zapalote Grande, Clavillo, Comiteco
Coahuila (124)	Tuxpeño, Celaya, Cónico Norteño, Elotes Occidentales, Tuxpeño Norteño, Tehua
Colima (29)	Tuxpeño, Tabloncillo, Reventador, Tabloncillo Perla, Vandeño, Jala
Durango (270)	Tuxpeño, Celaya, Cónico, Cónico Norteño, Chalqueño, Elotes Occidentales, Tabloncillo, Reventador, Tabloncillo Perla, Bolita, Pepitilla, San Juan, Dulcillo del Noroeste, Bofo, Blandito de Sonora, Blandito, Cristalino de Chihuahua, Gordo, Tablilla, Tunicata

Cuadro 1. Continuación

Guerrero (383)	Tuxpeño, Elotes Cónicos, Elotes Occidentales, Olotillo, Tabloncillo, Reventador, Vandeño, Nal-Tel, Pepitilla, Mushito, Tepecintle, Ancho, Conejo
Guanajuato (370)	Tuxpeño, Celaya, Cónico, Cónico Norteño, Chalqueño, Elotes Cónicos, Elotes Occidentales, Reventador, Maíz Dulce, Mushito, Fasciado
Hidalgo (236)	Tuxpeño, Celaya, Cónico, Cónico Norteño, Chalqueño, Elotes Cónicos, Elotes Occidentales, Olotillo, Bolita, Dzit-Bacal, Mushito, Cacahuacintle, Arrocillo Amarillo, Olotón, Arrocillo
Jalisco (683)	Tuxpeño, Celaya, Cónico, Cónico Norteño, Chalqueño, Elotes Cónicos, Elotes Occidentales, Tabloncillo, Reventador, Tabloncillo Perla, Bolita, Vandeño, Pepitilla, Maíz Dulce, Harinoso de Ocho, San Juan, Azul, Jala, Zamora, Complejo Serrano de Jalisco
Estado de México (724)	Tuxpeño, Celaya, Cónico, Cónico Norteño, Chalqueño, Elotes Cónicos, Bolita, Pepitilla, Cacahuacintle, Palomero, Arrocillo Amarillo, Ancho, Azul
Michoacán (528)	Tuxpeño, Celaya, Cónico, Cónico Norteño, Chalqueño, Elotes Cónicos, Elotes Occidentales, Olotillo, Reventador, Dzit-Bacal, Vandeño, Pepitilla, Maíz Dulce, Mushito, Cacahuacintle, Palomero, Conejo, Zamora
Morelos (165)	Tuxpeño, Chalqueño, Olotillo, Tabloncillo, Vandeño, Pepitilla, Tuxpeño Norteño, Ancho
Nayarit (336)	Tuxpeño, Celaya, Cónico, Cónico Norteño, Chalqueño, Elotes Occidentales, Olotillo, Tabloncillo, Reventador, Tabloncillo Perla, Vandeño, Maíz Dulce, Harinoso de Ocho, Bofo, Jala, Tablilla de Ocho
Nuevo León (118)	Tuxpeño, Cónico Norteño, Tabloncillo, Tablilla de Ocho
Oaxaca (562)	Tuxpeño, Celaya, Cónico, Cónico Norteño, Chalqueño, Elotes Cónicos, Olotillo, Bolita, Vandeño, Nal-Tel, Mushito, Tepecintle, Olotón, Conejo, Zapalote Chico, Zapalote Grade

Cuadro 1. Continuación

Puebla (943)	Tuxpeño, Celaya, Cónico, Cónico Norteño, Chalqueño, Elotes Cónicos, Elotes Occidentales, Olotillo, Bolita, Pepitilla, Mushito, Cacahuacintle, Palomero, Arrocillo Amarillo, Arrocillo
Quintana Roo (132)	Tuxpeño, Olotillo, Dzit-Bacal, Nal-Tel, Tepecintle
Querétaro (115)	Tuxpeño, Celaya, Cónico, Cónico Norteño, Chalqueño, Elotes Cónicos, Bofo, Onaveño, Fasciado
Sinaloa (187)	Tuxpeño, Tabloncillo, Reventador, Tabloncillo Perla, Maíz Dulce, Harinoso de Ocho, San Juan, Dulcillo del Noroeste, Blandito de Sonora, Lady Finger, Onaveño, Chapalote, Harinoso
San Luis Potosí (206)	Tuxpeño, Celaya, Cónico, Cónico Norteño, Chalqueño, Elotes Cónicos, Elotes Occidentales, Olotillo, Tabloncillo, Dzit-Bacal, Harinoso de Ocho
Sonora (183)	Tuxpeño, Tabloncillo, Reventador, Tabloncillo Perla, Nal-Tel, Harinoso de Ocho, San Juan, Dulcillo del Noroeste, Blandito de Sonora, Lady Finger, Onaveño, Chapalote
Tabasco (35)	Tuxpeño, Olotillo, Vandeño, Nal-Tel, Zapalote Grande
Tamaulipas (148)	Tuxpeño, Dzit-Bacal, Carmen
Tlaxcala (332)	Cónico, Chalqueño, Elotes Cónicos, Cacahuacintle, Palomero, Arrocillo Amarillo, Arrocillo
Veracruz (741)	Tuxpeño, Celaya, Cónico, Cónico Norteño, Chalqueño, Elotes Cónicos, Elotes Occidentales, Olotillo, Bolita, Dzit-Bacal, Nal-Tel, Pepitilla, Mushito, Cacahuacintle, Palomero, Tepecintle, Arrocillo Amarillo, Olotón, Coscomatepec
Yucatán (249)	Tuxpeño, Olotillo, Dzit-Bacal, Nal-Tel, Tepecintle, Zapalote Chico, Xmeneja
Zacatecas (263)	Celaya, Cónico, Cónico Norteño, Chalqueño, Elotes Cónicos, Elotes Occidentales, Tabloncillo, Bolita, Maíz Dulce, San Juan, Dulcillo del Noroeste, Bofo, Tablilla

*Entre paréntesis está el número de colectas registradas en el catálogo LAMP (1991). **Los grupos están como se describe en Sánchez y Goodman (1992a).

2.5. Las razas del maíz

El término raza se ha utilizado en el maíz y en las plantas cultivadas para agrupar individuos o poblaciones que comparten características en común, de orden morfológico, ecológico, genético y de historia de cultivo, que permiten diferenciarlas como grupo (Hernández y Alanís, 1970). Las razas se agrupan a su vez en grupos o complejos raciales, los cuales se asocian a una distribución geográfica y climática más o menos definida y a una historia evolutiva común (Goodman y McK. Bird, 1977).

Los indígenas mexicanos fueron quienes hicieron evolucionar al maíz, sembraron las razas derivadas, es decir, las razas nativas y razas criollas. Con la formación de las razas obtenidas y con los cruzamientos interraciales se dio origen a las razas modernas, de las cuales se han obtenido los híbridos actuales de muy altos rendimientos (Márquez-Sánchez, 2008).

El banco de germoplasma de CIMMyT cuenta con 28,000 accesiones de semilla de maíz, en las que se incluyen las razas criollas; la mayoría de este maíz ya no se siembra de manera recurrente entre los agricultores, pero ha servido a investigadores como base para el mejoramiento de germoplasma que es útil en la actualidad (CIMMyT, 2016).

La evaluación de la diversidad en maíces criollos es importante para el planteamiento de estrategias de conservación, caracterización y uso del germoplasma en el mejoramiento genético, dado su enorme potencial como fuente de características nuevas, exóticas y favorables, debido a su escaso uso en las tareas mencionadas es importante retomar la riqueza genética que se puede aportar con estos maíces (CIMMyT, 2014).

2.6. Descripción de las razas

Se ha adoptado la categoría de raza para la clasificación de los maíces. El primer estudio para el caso de México es el reportado por Wellhausen y colaboradores en 1951, en el que se describen 25 razas, tres sub-razas y siete tipos pobremente definidos. En el Cuadro 2, se muestra una clasificación más reciente con base en la tabla descriptiva de razas de maíz en México, detallada por Perales y colaboradores en 2009, en donde indican que los maíces mexicanos se agrupan en siete complejos raciales (Cónico, Sierra de Chihuahua, Ocho hileras, Maíces dentados tropicales, Maduración temprana a tardía, Chapalote y Razas no asociadas a un grupo específico), que incluyen 51 razas.

Cuadro 2. Descripción de complejos raciales de maíz.

Complejos raciales	Razas
Cónico	Chalqueño, Cónico, Elotes Cónicos, Cónico Norteño, Arrocillo, Cacahuacintle, Palomero Toluqueño, Pepitilla, Mushito.
Sierra de Chihuahua	Palomero de Chihuahua, Gordo, Azul, Cristalino de Chihuahua, Apachito.
Ocho Hileras	Harinoso de Ocho, Elotes Occidentales, Tabloncillo, Tabloncillo Perla, Bofo, Jala, Onaveño, Blando de Sonora, Reventador, Tablilla de Ocho, Ancho, Bolita, Zamorano Amarillo.
Maíces dentados tropicales	Tuxpeño, Vandeño, Tuxpeño Norteño, Celaya, Olotón, Olotillo, Dzit-Bacal, Comiteco, Tehua, Coscomatepec.
Maduración temprana a tardía	Conejo, Nal-Tel, Ratón, Tepecintle, Zapalote Chico.
Chapalote	Chapalote, Dulcillo del Noroeste, Elotero de Sinaloa.
Razas no asociadas a un grupo específico	Mixteco, Mountain Yellow, Negro de Chimaltenango, Complejo Serrano de Jalisco, Serrano Mixe, Serrano.

2.6.1. La raza Tuxpeño

La raza Tuxpeño perteneciente al complejo de maíces dentados tropicales, se caracteriza por sus mazorcas grandes, cilíndricas, de grano dentado, predominando la tonalidad blanca, pero puede presentar diversas tonalidades, tiene un alto número de hileras y granos por hilera, presenta también buena calidad agronómica en planta y resistencia a enfermedades, lo que la hace una de las razas más productivas de México (Wellhausen *et al.*, 1951).

Domina grandes áreas principalmente las partes bajas tropicales bajo temporal, así como en las subtropicales bajo riego; se puede ampliar más su siembra utilizando para ello grandes áreas subutilizadas como potreros, predomina la siembra de primavera-verano, y en áreas con disposición de riego permiten su cultivo en el ciclo otoño-invierno (Wellhausen *et al.*, 1951).

Kato (1984) encontró que el territorio costero del Golfo de México fue una de las rutas que siguió prácticamente el origen del maíz y actualmente la raza que predomina en esos lugares es el Tuxpeño; en el occidente de México y el territorio sur de Mesoamérica (Oaxaca, Chiapas y Guatemala) es donde actualmente se encuentra la mayor variación racial (27 razas mexicanas más las guatemaltecas) y es ahí donde convergieron germoplasmas de los complejos Zapalote, Tuxpeño y de Tierras altas de Guatemala. En la costa del Golfo de México predomina la raza Tuxpeño, debido a que solo uno de los germoplasmas originales tuvo influencia, el complejo Tuxpeño.

Estudios posteriores como el de Ortega-Paczka (1985) mostraron la existencia de razas que no habían sido descritas; con el objetivo de verificar las colecciones resguardadas en el Banco de germoplasma del entonces Instituto Nacional de Investigaciones Agrícolas, hoy Instituto Nacional de Investigaciones Forestales y Agropecuarias logró describir cinco razas nuevas: Ratón, Tuxpeño Norteño, Onaveño, Cristalino de Chihuahua y Palomero de Chihuahua.

Con la finalidad de recopilar datos de pasaporte en entradas del Banco en 1986-87, comenzaron intensos esfuerzos para evaluar características generales tales

como la morfología y la estructura de la planta. Cerca de 40 ensayos se llevaron a cabo en grupos raciales establecidos recurriendo a la información del pasaporte recién compilado. Se resumieron los resultados para el complejo raza Tuxpeño y se les utilizó para desarrollar subconjuntos centrales de ese complejo. Una comparación posterior, a través del proyecto de base de datos mundial de pasaporte, en accesiones de la raza Tuxpeño en los bancos del CIMMyT y el INIFAP mostró que la institución mexicana cuenta con 570 accesiones que no se incluyeron en la evaluación de CIMMyT.

El complejo Tuxpeño en el área Occidental y norte del Estado México contiene introgresiones de germoplasma de maíz del complejo Mesa Central. En las zonas del sur de México, el germoplasma de Tuxpeño es influenciado por el complejo Zapalote (Kato, 2005).

Las colecciones del INIFAP contienen accesiones de Tuxpeño traídas de Chiapas, Tabasco y Quintana Roo (CIMMyT, 1994). En 2011, Taba consideró que existía un enorme potencial para hacer un amplio uso de esta raza en el futuro y en el mejoramiento. Wen por su parte en 2012 con el fin de indagar la forma de explotar la raza de maíz tropical Tuxpeño en la mejora futura de maíz evaluó el uso de marcadores de polimorfismo de nucleótido único (SNP) en todo el genoma.

Actualmente es considerada una raza importante a nivel nacional y es la más utilizada para el mejoramiento, aunque ha aumentado su área de distribución, no está claro cuál es su centro de distribución y de diversidad. Predominan formas de grano blanco, a excepción de la península de Yucatán en donde también son muy importantes las formas de color amarillo y azul-negra. Recientemente se ha colectado una gran variación de color (blanco, varias tonalidades de amarillo, variegados, jaspeados y rojos) en la región de las Huastecas (Hidalgo, Veracruz, San Luis Potosí) y en Tuxtla, Veracruz. (CONABIO, 2016).

2.6.2. La raza Cónico Norteño

A partir de la conquista se han desarrollado las razas Modernas incipientes que son: Bolita, Chalqueño, Celaya y Cónico Norteño (Wellhausen *et al.*, 1951). Entre 1974 y 1976, se evaluaron bajo temporal 1353 colectas de maíz de la región Centro-Norte de México (C-N) y de Valles Altos (Zapata y Luna, 1989; Ortega *et al.*, 1991); de ahí se derivaron compuestos con colectas sobresalientes, como el compuesto de colecciones blancas (CCB), formado con 81 colectas; el Con-Bol conformado de colectas de las razas Cónico Norteño y Bolita, seleccionadas por aptitud combinatoria general.

De acuerdo con Stretta y Mosiño en 1963, y Contreras en 1990) las zonas semiáridas y áridas del C-N comprenden gran parte de los estados de Zacatecas, Durango, San Luis Potosí, Guanajuato y Aguascalientes, y pequeña parte de Jalisco. Las estadísticas de la SAGARPA indican que constantemente el cultivo de maíz ocupa el segundo lugar en importancia agrícola.

Entre 1991 y 1994 cubrió 30 y 35 % del área cultivada, con 870 mil ha. De esa superficie la mayor parte (87 %) es de temporal y de esa una pequeña proporción se cultiva bajo los sistemas de producción de bajío, abanicos aluviales o “de húmedo”.

De la superficie total de temporal sembrada con maíz en la región C-N, en promedio 27 % no produce grano, principalmente por efecto de la sequía (Luna, 1993) y con rendimientos entre 0.30 y 0.45 t·ha⁻¹ en Aguascalientes, San Luis Potosí, Guanajuato y Jalisco, mientras que en Zacatecas y Durango entre 0.67 y 0.77 t·ha⁻¹.

Aún con bajos rendimientos, el cultivo de maíz de temporal es muy importante en la región C-N, pues representa un cultivo de subsistencia para más de 320 mil familias que obtienen forraje para sus animales o para vender y lo más importante, tienen la esperanza de cosechar algo de grano (Montañez y Warman, 1985).

La región semiárida-árida se localiza entre los 21° y 26° de latitud N, 99° y 106° de longitud O entre los 1700 a 2200 msnm. Donde predominan los climas BS1kw(w)(e) y BS1hw(w)(e) en área de productividad frecuente, con precipitación que fluctúa entre 400 y 500 mm, temperatura media entre 15 y 21° C, entre 110 a 120 días de ciclo de cultivo y BS0 en zona de baja productividad con 250 y 400 mm de precipitación, 1 a 20 ° C de temperatura y 80 a 100 días por ciclo (Luna y Zapata, 1988; Luna, 1993).

En casi toda la región 65 % de la lluvia se registra antes de la antésis, adicional a esto aproximadamente, 51 % de los suelos tienen menos de 50 cm de profundidad, 46 % entre 50 y 100 cm y su contenido de materia orgánica es inferior a 1 %; por esto, su capacidad de retención de humedad es muy baja, no obstante que predominan las texturas cercanas a la franca (Luna, 1996).

La raza Cónico Norteño pertenece al complejo Cónico, correspondiente con la descripción de Perales (Perales y Golicher, 2011), se ha utilizado en el perfeccionamiento de razas resistentes a sequía, y generalmente para el mejoramiento genético en la región templada semiárida, en combinación con poblaciones precoces de la raza Bolita. Apropia para llanos al oriente de la Sierra Madre Occidental, norte de Guanajuato y hasta el sur de Chihuahua (zonas templadas semiáridas BSh), altura arriba de 1600 a 2500 msnm, con formas precoces para siembra de temporal escaso o crítico (Avendaño *et al.*, 2005).

También presume una adaptación a la zona templada semiárida del norte del país, caracterizada por plantas de porte bajo, mazorcas cónicas y textura de grano semi- dentado, de color blanco predominantemente, pero ocasionalmente se encuentran diferentes tonalidades del amarillo al morado (Rincón *et al.*, 2010). Raza de interés para el CIMMyT por los altos niveles de tolerancia a sequía observados en muchas de sus accesiones provenientes de su banco de germoplasma.

2.7. Base genética y molecular de la relación entre la raza Tuxpeño y Cónico Norteño

Evaluaciones realizadas en el Colegio de Posgraduados, campus Montecillos, incluyeron complejos mexicanos como Celaya y Vandeño, así como las razas del altiplano Cónico, Cónico Norteño, y Chalqueño.

Basados en análisis fenotípicos de caracteres morfológicos Wellhausen et al. (1952) propusieron que la raza Tuxpeño al hibridarse con la raza Cónico dio lugar a la raza Chalqueño. Hipótesis que para John F. Doebley no parece razonable analizar incluso sobre una base morfológica, simplemente porque Chalqueño es morfológicamente similar a Cónico, y ambas distintas a Tuxpeño. Por lo demás, Cónico y Chalqueño son razas adaptadas a altitudes elevadas, mientras que Tuxpeño es una raza adaptada a bajas altitudes.

Dadas estas consideraciones, no fue extraordinario encontrar que información aloenzimática demostrara que Cónico y Chalqueño se encuentran estrechamente relacionados y que esta última raza no es intermedia entre Tuxpeño y Cónico (Figura 3).

Wellhausen *et al.* (1952) presentaron una serie de hipótesis muy explícitas concernientes a las relaciones evolutivas existentes entre las razas de maíz de México. Estos autores basaron sus hipótesis en análisis fenéticos de características morfológicas. Ellos concluyeron que la mayoría de las razas de maíz mexicanas surgieron a partir de la hibridación entre otras razas. Por ejemplo, dichos investigadores consideraron a la raza Celaya como un híbrido derivado de las razas Tuxpeño y Tabloncillo. Si este punto de vista fuese correcto, uno podría esperar que la raza Celaya fuese intermedia en su composición aloenzimática entre Tuxpeño y Tabloncillo. No obstante, la hipótesis del origen híbrido de la raza indicada no resulta confirmada.

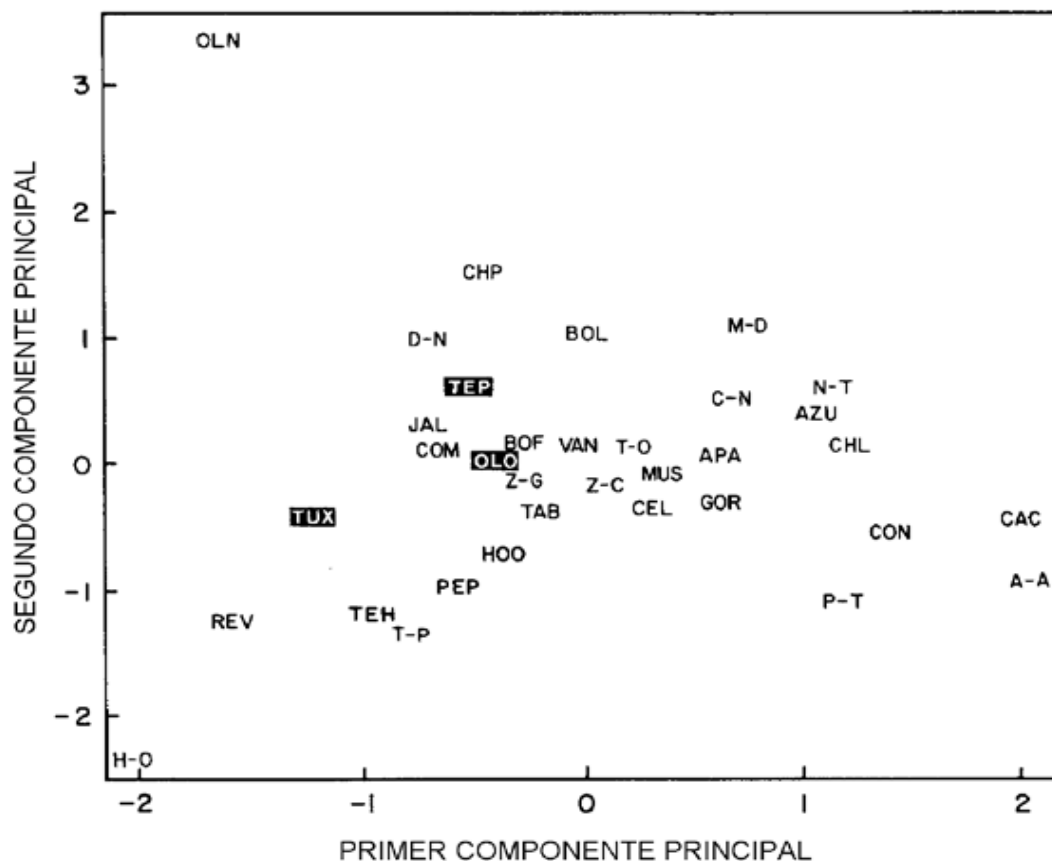


Figura 3. Análisis de componentes principales, basada en caracteres aloenzimáticos que contradicen la supuesta relación entre la raza Tuxpeño, Cónico y Chalqueño.

A partir del análisis de su composición aloenzimática de modo evidente, Tuxpeño está adaptado a las bajas altitudes (0 a 500 msnm). Por lo anterior, parece improbable que uno pudiese hibridar formas de maíz adaptadas a altitudes bajas o intermedias y consecuentemente producir una forma adaptada a altitudes elevadas.

Respecto a la raza Tuxpeño, Wellhausen *et al.* (1952,1953) indicaron que “existe poca razón para dudar que Tuxpeño es básicamente el producto de la hibridación entre Olotillo y Tepecintle”. Mientras que el análisis de Doebley () revela que Tuxpeño no es intermedio en su constitución aloenzimática, por lo que la hipótesis sobre su origen híbrido es rechazada.

Ruíz en 2013 clasificó las razas de maíz mexicanas mediante taxonomía numérica empleando variables descriptivas correspondientes a condiciones de humedad de los sitios de colecta de maíces nativos, dejando ver que existen grupos de razas con relación a adaptación y déficit hídrico durante el ciclo de temporal. El grupo dos integrado por las razas Chapalote, Dulcillo del Noroeste, Tuxpeño Norteño, Cónico Norteño, Tablilla de Ocho y Gordo se identificó como el de mayor adaptación a un temporal más seco, incluso con condiciones de semi-aridez (índice de humedad estacional (IHe) entre 0.21 y 0.5). Mientras que el grupo cuatro con adaptación a temporal semiárido se encuentra integrado por las razas Chapalote, Dulcillo del Noroeste, Gordo, Tablilla de Ocho, Cónico Norteño y Tuxpeño Norteño, que se ajustan a los ambientes con el mayor déficit de humedad, precipitación del ciclo de 409 a 496 mm, (IHe menor que 0.5).

Según Linhart y Grant (1996) todas las razas de este grupo podrían aportar genes de resistencia o tolerancia a sequía. Señalando que en México existen recursos genéticos de maíz adaptados a condiciones de sequía y que pueden contribuir a la generación de variedades adaptables a condiciones hídricas menos favorables por la presencia del cambio climático.

2.8. Fuentes de variación genética (ADN)

Un aspecto fundamental de la biodiversidad es la variación genética de las poblaciones y las especies. esta variación puede ser clasificada en cuatro tipos en función de sus consecuencias en la adecuación: variación de alelos letales, deletéreos o neutros, y la asociada con alelos que aumentan la adecuación de sus portadores. Asimismo, se pueden dividir los tipos de variación genética en aquellos que estudian rasgos cualitativos (determinados por un locus) y los que son cuantitativos (Piñero *et al.*, 2008).

A continuación, se enlistan algunos de los posibles eventos fuentes de variación en las especies:

Cambio de base: Se producen al cambiar en una posición un par de bases por otro, como es el caso del polimorfismo de secuencia simple (SNP). El efecto de este cambio de base puede ser muy pequeño, en donde el resultado es sobre el fenotipo: “neutro”, pero también se pueden dar cambios muy abundantes, especialmente en poblaciones naturales y algunos cultivos.

Inserciones/Deleciones (InDel): Son las anomalías estructurales cromosómicas que consisten en la pérdida o ganancia de un fragmento de ADN, como ejemplo de estas variaciones tenemos a los transposones, que se encuentran con frecuencia cambiando patrones de expresión para genes próximos.

Rearreglos del genoma, ploidía, metilación. Estos son inestables y reversibles, la mayoría “sensibles” a cambios desde el punto de vista evolutivo.

Estos eventos fuentes de variación en las especies, pueden ocurrir por cuestiones naturales o inducidas y estas variaciones pueden ser útiles para el mejoramiento asistido.

2.9. Marcadores moleculares y sus contribuciones en estudios con plantas

Los marcadores moleculares o genéticos son cualquier carácter heredable que muestra variabilidad o polimorfismo en los individuos. En estudios de plantas, los marcadores moleculares tienen muchas utilidades, entre las que destacan: obtención de la huella genética, son ventajosos para el soporte al registro de variedades, estudio de diversidad genética, relaciones filogenéticas, determinación de niveles de estabilidad genética en procesos *in vitro*, soporte para selección asistida por marcadores (SAM), caracterización de bancos de germoplasma, registro de parentales, mapeo genético, mapeo por asociación, así como en selección genómica (Collard *et al.*, 2005).

Los marcadores moleculares utilizados en plantas incluyen: proteínas; isoenzimas, marcadores de ADN; polimorfismo de longitud amplificado (AFLP), SSR (secuencia corta repetida) y el polimorfismo de nucleótido simple (SNP). Estos marcadores están presentes en forma dominante (presencia / ausencia en

un individuo o población) o co-dominantes (se pueden identificar estados heterocigotos).

Para la elección de los tipos de marcadores moleculares a utilizar, es necesario tener en cuenta que el marcador cumpla con herencia co-dominante, interacción nula con el medio o con otros genes, elevado polimorfismo, expresión en cualquier tejido o estado de desarrollo, rapidez, simplicidad en la detección y bajo costo (Phillips, 1995).

Los marcadores se pueden clasificar como dominantes entre los que destacan los AFLP's, o co-dominantes como los SSR's y SNP's que incorporan el polimorfismo más pequeño posible de evaluar y los más abundantes del genoma (Cuadro 3).

Cuadro 3. Clasificación de los marcadores, ventajas y desventajas (Schlötterer, 2004; Aranguren-Méndez *et al.*, 2005; Bedoya, 2012).

Marcador	Ventajas	Desventajas
Secuencias Repetidas Cortas Microsatelites - SSR	Altamente informativos Fácil de aislar Interpretación sencilla de los resultados Reproducibilidad alta	Alta tasa de mutación Comportamiento complejo de mutación Comparación entre laboratorios requiere preparación
Polimorfismo de base única SNPs	Baja tasa de mutación Abundantes en el genoma Facilidad de caracterización Nuevos enfoques analíticos se están desarrollando Repetitividad entre laboratorios Bases de datos existentes y accesibles	Costosos de aislar y caracterizar Bajo contenido de información para un único SNP
Secuenciación de ADN	Resultados altamente reproducibles El contenido de información es máximo Fácil comparación entre estudios	Significativamente más costoso que las otras técnicas Presenta exigencias técnicas

Una amplia proporción de la variación genética valiosa permanece oculta en las repisas de los bancos de germoplasma, como lo es el caso de las razas de maíces criollas, mientras que los retos para la producción de cultivos se vuelven más grandes y el fortalecimiento de la seguridad alimentaria está en constante polémica.

2.10. Genotipificación por secuenciación GBS, DArTseq

En el contexto de que un marcador molecular es una secuencia de ADN (expresado o no) capaz de detectar polimorfismo y que posee un comportamiento mendeliano se han evaluado distintos métodos para la identificación de estas secuencias como: hibridación, PCR, PCR seguido de fragmentación, fragmentación seguida de PCR (Azofeifa-Delgado, 2006).

Para su detección se han empleado técnicas tales como: Marcación por fluorescencia, Marcación por nitrato de plata, Marcación por bromuro de etidio, Marcación por ^{32}P o ^{33}P , Secuenciación Sanger y Secuenciación NGS.

Con la aparición de métodos rápidos de secuenciación de ADN se ha acelerado en gran medida la investigación y descubrimientos biológicos (Sansaloni, 2015).

La secuenciación de ADN es un proceso que determina el orden preciso de nucleótidos dentro de una molécula de ADN, este proceso incluye cualquier método o tecnología que es utilizado para determinar el orden de cuatro bases: Adenina (A), Guanina (G), Citosina (C) y Timina (T).

La Secuenciación Sanger se basa en el empleo de dideoxinucleótidos que carecen del grupo hidroxilo del carbono 3', de manera que cuando uno de estos nucleótidos se incorpora a una cadena de DNA en crecimiento, esta cadena no puede continuar elongándose, esto es así ya que la DNA polimerasa necesita un grupo terminal 3' OH para añadir el siguiente nucleótido y el dideoxinucleótido incorporado carece de este grupo hidroxilo.

La Secuenciación de última generación, NGS por sus siglas en inglés, es un proceso que emplea micro y nanotecnologías para reducir el tamaño de los componentes de la muestra, reduciendo costos de reactivos, y permitiendo reacciones de secuenciación masivamente paralelas.

Los principales tipos de NGS son Sequencing by synthesis: Illumina: HiSeq 2000 / HiSeq 2500, Lief Technologies: Ion Proton, Roche 454 pyrosequencing; Sequencing by ligation: SOLiD; Single-molecule sequencing: HeliScope, Real Time Sequencing.

Entre las principales características que posee la NGS son: altamente multiplexados, lo que permite la secuenciación simultánea y el análisis de millones de muestras. Su llegada ha sido reciente y comenzó a estar disponible en el mercado alrededor de 2005, métodos de secuenciación diferentes se han desarrollado, los cuales están siendo continuamente mejorados a un ritmo asombroso.

El genotipado por secuenciación abreviado como GBS, DArTseq, por sus siglas en inglés, es un sistema altamente multiplexado (Figura 4) que genera cientos de miles de marcadores de tipo SNPs (Polimorfismo de nucleótido simple) y PAVs (Variación de Presencia/Ausencia), para construir bibliotecas de perfil reducido a partir de representaciones genómicas reducidas y valiéndose de plataformas de secuenciación de última generación (NGS) de Illumina desarrollada en Cornell University por Rob Elshire.

Se genera un gran número de SNPs para su uso en los análisis genéticos. Los componentes clave de este sistema son: manejo de la muestra reducida, menos pasos de PCR y purificación, así como menores costos. Se utilizan enzimas de restricción para reducir la complejidad del genoma y evitar la fracción repetitiva del genoma. El resultado es un genotipo con decenas a cientos de miles de marcadores SNP, listo para analizar (Elshire *et al.*, 2011).

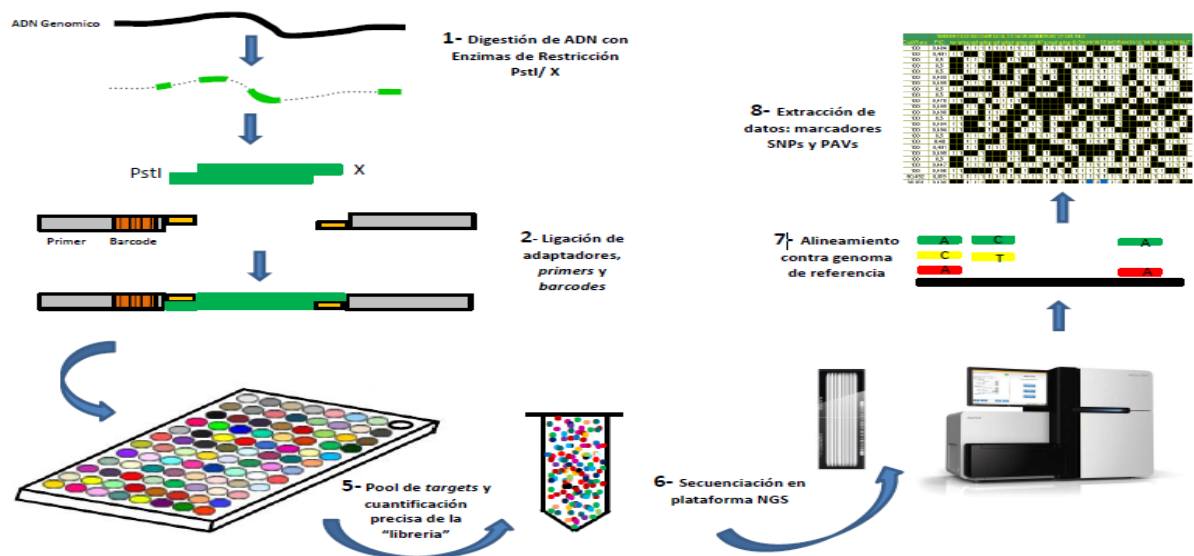


Figura 4. Proceso para la generación de GBS-DArTseq, Hi-Seq (SAGA, 2016).

Reducción de complejidad genómica (digestión/ligación): Esto se realiza mediante la digestión del ADN genómico con enzimas de restricción tipo II (sensibles a la metilación) para recortar el genoma y la ligación de adaptadores para identificación de las muestras y generar la biblioteca de fragmentos que será secuenciada. El uso de este tipo de enzimas hace que el método sea rápido, sencillo, altamente específico, reproducible, y la sensibilidad a la metilación permite alcanzar regiones del genoma difíciles de estudiar con otros métodos (Elshire *et al.*, 2011).

Habitualmente es utilizada una enzima de corte frecuente y una de corte específico. La enzima de corte frecuente más utilizada es PstI (C_TGCA^G), la cual se combina con otra(s) enzimas de corte específico dependiendo de la especie y el objetivo del estudio (Ej. NspI para estudios de diversidad y HpaII para selección genómica). La selección de la enzima es importante para lograr una mejor reducción de complejidad. En el caso del maíz, una de las enzimas más adecuada es la ApeKI, un tipo de endonucleasa de restricción de corte frecuente que reconoce una secuencia degenerada de 5 pares de bases (pb) (GCWGC, donde W puede ser una A o T) y es parcialmente sensible a la

metilación y de poco reconocimiento para la mayoría de retrotransposones en maíz (Elshire *et al.*, 2011). Junto con las enzimas de restricción debe agregarse una ligasa, la cual se encarga de unir un par de adaptadores los cuales permitirán el manejo de las muestras a lo largo del proceso (identificación) y amplificación (común) (Elshire *et al.*, 2011).

Dos tipos de adaptadores se utilizan en este protocolo (Elshire *et al.*, 2011): los identificadores o “barcode” (secuencias de 4–9 pb), que tienen la función de identificar las muestras y diferenciarlas al momento de realizar la mezcla de todas las muestras y debe contener la secuencia afín a la enzima de corte frecuente. Esto es necesario cuando se tienen diferentes especies o genotipos, en un mismo carril de la celda de secuenciación. El segundo tipo de adaptador llamado “común” contiene una secuencia afín al sitio de corte de la enzima secundaria (específica para la especie). Los dos tipos de adaptadores tienen secuencias complementarias para los iniciadores que se utilizan en la amplificación selectiva. La ligación de los adaptadores es facilitada por una ligasa en la misma reacción de digestión del ADN. Por último, los adaptadores permitirán el manejo de las muestras a lo largo del proceso de identificación y amplificación (Elshire *et al.*, 2011).

Amplificación selectiva: La amplificación de estos fragmentos (targets, en inglés) ocurrirá gracias al uso de cebadores (primers, en inglés) complementarios a los adaptadores utilizados en la digestión/ligación. Sólo los fragmentos que tengan ambos adaptadores serán amplificados. Además, estos cebadores también contienen secuencias adicionales en su extremo 5' para hacerlos compatibles con fragmentos únicos de secuenciación (Elshire *et al.*, 2011).

Revisión de la calidad de fragmentos amplificados: Para revisar la calidad de los fragmentos amplificados se realiza un gel de agarosa, en el cual se busca obtener un pequeño barrido de ADN. Los barridos deben ser uniformes en intensidad, tamaño y migración en el gel si son de la misma especie y si se utiliza el mismo método de reducción de complejidad. Si se observan bandas definidas y dímeros

de cebadores (de aproximadamente 200 pb de longitud) pondrá en duda la calidad obtenida (Elshire *et al.*, 2011).

Mezclado, purificación y cuantificación: Se mezclan todos los fragmentos o targets en un solo tubo para simplificar su manejo y finalmente se purifican con un kit comercial. La mezcla es posible gracias al uso de los adaptadores “barcode”. Posteriormente se verifica por segunda ocasión en un gel de agarosa esperando observar barridos uniformes. Una vez combinados y purificados, se realiza la cuantificación para medir la cantidad de ADN necesario para la secuenciación, mediante equipo especializado como nanodrop, lectores de placas o espectrofotómetros (Elshire *et al.*, 2011).

Secuenciación en plataforma Illumina: Previo a la secuenciación en el HiSeq 2500 debe prepararse la celda de secuenciación (flowcell, en inglés), el cual es un portaobjetos con un canal dividido en 8 carriles, los cuales contienen cebadores específicos a los que se unen los fragmentos de ADN durante la reacción de amplificación de puente (bridge-PCR). Estos fragmentos formarán cúmulos concentrados llamados “clústeres” con una misma secuencia, los cuales son el material de lectura para el secuenciador (Elshire *et al.*, 2011).

Generación de clusters en cBot Illumina: El cBot es el equipo especializado en preparar la celda de secuenciación que contiene las muestras que serán secuenciadas. Previo a este paso las muestras de ADN purificado deben pasar por tres pasos: (a) Dilución: la concentración de ADN determina si se forman muchos o pocos clústeres, por lo cual se diluyen las muestras para ajustarlas a una misma concentración que forma una cantidad intermedia de clústeres; (b) Desnaturalización, consiste en la separación de las cadenas de ADN, formando ADN de cadena sencilla (ssDNA, en inglés). Luego de desnaturalizar el ADN es necesario volver a diluir las muestras utilizando un amortiguador de hibridación. Finalmente se produce la (c) Hibridación o amplificación de puente (bridge-PCR) (Elshire *et al.*, 2011).

Secuenciación por síntesis: Luego de la generación de los clusters en el cBot, la celda de secuenciación esta lista para ser introducida en el secuenciador HiSeq

2500 de Illumina, utilizando una serie de reactivos provistos por el mismo fabricante. Este equipo tiene la capacidad de correr entre 2 a 8 placas de 96 muestras por celda de secuenciación. El HiSeq 2500 se basa en tecnología de secuenciación por síntesis (SBS, por sus siglas en inglés) de Illumina, mediante la reacción química denominada Terminación Reversible Cíclica (CTR, por sus siglas en inglés), la cual incorpora nucleótidos marcados fluorescentemente y luego toma fotografías de los clusters que posteriormente son procesadas para determinar la secuencia de los fragmentos (Elshire *et al.*, 2011).

Obtención y análisis de los datos: Verificación de la calidad de datos: La calidad técnica de los datos obtenidos se revisa utilizando el software Illumina Sequencer Analysis Viewer (SAV). Este software permite ver las métricas de calidad generadas por el Analizador en Tiempo Real del secuenciador (RTA, por sus siglas en inglés) (Elshire *et al.*, 2011).

Selección de datos e identificación de SNPs: Para la identificación de los marcadores SNPs se utiliza un Sistema de Gestión de Información de Laboratorio (LIMS: Laboratory Information Management System, en inglés). Este sistema permite conservar los marcadores de buena calidad y desechar las lecturas que no cumplen con los estándares mínimos. Posteriormente, segrega las secuencias con base a los “barcode” para que sea posible el análisis de cada muestra por separado. Una vez producidos los marcadores, un investigador capacitado debe revisar manualmente los resultados para seleccionar únicamente los mejores marcadores con base en su alineamiento a un genoma de referencia, así como la cantidad de repeticiones del mismo marcador que se obtuvo a lo largo del experimento (Elshire *et al.*, 2011).

Se han llevado a cabo con éxito GBS en muchos organismos, incluyendo: el maíz, el sorgo, la yuca, el sauce, el arroz, las manzanas, la vaca, el zorro y muchos más. Actualmente se desarrolla un método para análisis de compuestos de 30 individuos de cada población y frecuencias SNP (Hearne *et al.*, in prep).

2.11. Distancias Genéticas

Pueden ser divididas en dos grupos. El primero con las distancias geométricas (D_A), que reflejan la similitud entre poblaciones con respecto a las frecuencias alélicas que éstas presentan en determinados sistemas genéticos. A este grupo corresponden las distancias euclidianas y no euclidianas, las cuales establecen simplemente una visión geométrica de las distancias entre puntos en una hipersfera cuyo número de dimensiones es igual al de alelos considerados en la comparación (Paredes, 2014).

De igual manera la D_A es utilizada frecuentemente con microsatélites debido a que ha demostrado ser la que mejor representa las relaciones poblacionales, a partir de datos generados por computadora (Demarchi, 2009):

$$D_A = 1 - \sum_u \sqrt{X_u Y_u} \quad \begin{array}{l} X_u = \text{frecuencia alélica del alelo } u \text{ en la población } X \\ Y_u = \text{frecuencia alélica del alelo } u \text{ en la población } Y \end{array}$$

A este mismo grupo corresponde la distancia de Roger, que es útil para un determinado locus con k alelos, si x_i e y_i son las frecuencias del alelo i en las poblaciones X e Y , respectivamente, la Distancia de Roger se define como:

$$D = \sqrt{0.5 \sum_{i=1}^k (x_i - y_i)^2}$$

donde la sumatoria es para todos los alelos. Cuando se consideran los datos de varios *loci*, la estima de la distancia genética promedio se calcula como la media aritmética de los valores calculados para cada uno de los *loci*. La distancia genética de Roger puede tomar valores entre cero y uno. La similitud de Roger sería $S=1-D$ (Demarchi, 2009).

El segundo grupo comprende a aquellas que se basan en modelos evolutivos, donde la variación es atribuida principalmente a la acción de la mutación y la deriva, asumiendo que los marcadores utilizados no están sujetos a la selección natural (Demarchi, 2009).

La gran mayoría de estas medidas de distancia, se ajustan al modelo de alelos infinitos (Infinite Alleles Model-IAM). Según este modelo cada mutación crea aleatoriamente un alelo nuevo “y”, por tanto, cada estado alélico está relacionado por igual a cualquier otro.

En el caso particular de los microsatélites, las distancias se calculan sin tomar en cuenta el número de repeticiones en que difieren dos estados alélicos. Esto implica que el tamaño no importa. La distancia basada en el modelo IAM más utilizada es la distancia estándar DA, que también ha demostrado ser adecuada para el análisis de microsatélites (Demarchi, 2009).

2.12. Distancias Geográficas

La Tierra es un cuerpo tridimensional con aspecto próximo a la esfera, achatada por los polos y ensanchada por el Ecuador, parecido a la figura geométrica elipsoide. Evidentemente no es posible la representación de la Tierra a tamaño real, por lo que se plantea una relación entre una distancia sobre el mapa y su proporción sobre la superficie terrestre: la escala. Los métodos de proyección cartográfica y las dimensiones, conducen al concepto de escala, estos dos factores hacen posible la relación entre el mapa y la realidad (Instituto Geográfico Nacional, 2016).

La representación de la superficie terrestre sobre una superficie plana, sin que haya deformaciones, es geoméricamente imposible. En cartografía, este problema se resuelve mediante las proyecciones, así, una proyección cartográfica es una correspondencia biunívoca entre los puntos de la superficie terrestre y sus transformados en el plano llamado plano de proyección (Instituto Geográfico Nacional, 2016).

Básicamente la localización geográfica de un punto se puede realizar detallando uno de estos dos parámetros: coordenadas geográficas en forma longitud-latitud y coordenadas (x, y) UTM, Universal Transversal Mercator (Fernández, 2001).

La distancia geográfica es el espacio que existe entre un punto geográfico o lugar y otro. Para estudiar estas distancias se utiliza el Sistema de Información Geográfica (SIG en español o GIS en inglés) que es un conjunto de herramientas que integra y relaciona diversos componentes (usuarios, hardware, software, procesos) para permitir la organización, almacenamiento, manipulación, análisis y modelización de cantidades de datos vinculados a una referencia espacial (Instituto Geográfico Nacional, 2016).

En relación con la red geográfica que forman los paralelos y meridianos se definen las coordenadas geográficas que permiten ubicar con precisión la ubicación de un punto cualquiera de la superficie terrestre. Estas dos coordenadas se miden como la distancia desde el punto en cuestión hasta las líneas de base del sistema y reciben el nombre de: Latitud: su línea de base es el Ecuador, Longitud: su línea de base es el Meridiano de Greenwich (Fernández, 2001).

Estas coordenadas se expresan en grados sexagesimales, Para los paralelos, sabiendo que la circunferencia que corresponde al Ecuador mide 40.076 km, 1º equivale a 113,3 km. Para los meridianos, sabiendo que junto con sus correspondientes antimeridianos se forman circunferencias de 40.007 km de longitud, 1º equivale a 111,11 km. (Instituto Geográfico Nacional 2016).

2.13. Herramientas de análisis y visualización para datos genéticos y geográficos

Debido a los avances tecnológicos en cuanto a caracterización y obtención de información tanto genética como geográfica, la cantidad de datos que se pueden generar mediante estas nuevas tecnologías, ha ido en aumento haciendo en ocasiones un tanto complejo su análisis y visualización, por ello el área bioinformática en relación con otras áreas de interés ha desarrollado sistemas y aplicaciones para facilitar la manipulación e interpretación de la excesiva cantidad de información generada en la actualidad.

En este contexto podemos mencionar algunos de los programas utilizados actualmente en análisis geográficos y genéticos.

Germinate 3, es una aplicación basada en Java y desarrollada por el Instituto James Hutton. Esta plataforma almacena información referente a los conjuntos de datos experimentales tales como datos de caracterización, conjuntos fenotípicos, mapas moleculares, frecuencias alélicas, climas, búsqueda geográfica, conjuntos de datos analizados, entre otros (James Hutton Institute, 2016).

Curly Whirly, es una herramienta útil en representaciones visuales y técnicas de interacción, aprovechan sus ojos, sobresaliendo en tareas tales como: detección de bordes, reconocimiento de patrones, discernimiento de las diferencias en color, proporcionan capacidad para ver, explorar y comprender grandes cantidades de información a la vez (James Hutton Institute, 2016).

ArcMAP, es el componente principal de Esri's ArcGIS, conjunto de programas de procesamiento geo-espaciales, y se utiliza principalmente para visualizar, editar, crear y analizar datos geo-espaciales. ArcMap permite al usuario explorar datos dentro de un conjunto de datos, simbolizar funciones en consecuencia y crear mapas.

Dentro de las técnicas multivariantes podemos citar al Escalamiento Multidimensional (Multidimensional Scaling, MDS). Es una técnica multivariante de interdependencia que trata de representar en un espacio geométrico de pocas dimensiones las proximidades existentes entre un conjunto de objetos o de estímulos. Esta técnica, aunque tiene sus raíces a principios del siglo XX, hoy día sigue siendo infrutilizada en muchas áreas (Casas, 2012).

3. MATERIALES Y MÉTODOS

3.1. Material vegetal

Se analizaron 830 y 1135 accesiones de las razas de maíces criollos Tuxpeño y Cónico Norteño, respectivamente (Apéndice 1). Dichas accesiones fueron seleccionadas de un panel principal, debido a que en ensayos previos han mostrado tolerancia a condiciones climáticas adversas como el estrés hídrico.

Los materiales fueron proporcionados por el banco de germoplasma del Centro Internacional de Mejoramiento de Maíz y Trigo (CIMMyT) y se sembraron en charolas de unicel, con 128 cavidades en invernadero dentro de la Estación Experimental Batán-CIMMyT de clima tropical, ubicado en el municipio de Texcoco, que se encuentra en las coordenadas GPS: Longitud (dec): -98.856389, Latitud (dec): 19.526944, a una altura de 2300 metros sobre el nivel del mar.

El germoplasma proporcionado por el banco de CIMMyT, se trabajó por el grupo Seeds of Discovery, del proyecto MasAgro Biodiversidad, que tiene como propósito la utilización de recursos genéticos inapreciables mediante tecnologías de última generación, así como el desarrollo de capacidades, con la finalidad de ofrecer un paquete de herramientas a los investigadores que contribuyen al fitomejoramiento. Para alcanzar estas metas se está implementando lo que llaman “Atlas Molecular” donde se ingresará la información obtenida de diversos materiales evaluados e información de interés que ofrece germoplasma con capacidades para afrontar y tolerar cambios climáticos y fortalecer la seguridad alimentaria. Mediante estrategias que buscan el fortalecimiento de capacidades en cursos y talleres a representantes de las redes de mejoramiento de maíz y trigo.

3.2. Colecta de material vegetal y extracción ADN

Se usaron 30 hojas de plantas elegidas al azar de cada accesión para hacer una muestra representativa compuesta de ADN, para ello se cosecharon hojas frescas y sanas tomando tejido del área sin nervadura central, zonas necrosadas

o lesiones, de plantas de alrededor de 30 días de edad. Las muestras de tejido foliar se tomaron del invernadero #2 de la estación experimental “El Batán”. El tejido vegetal se congeló a -80 °C para su posterior secado mediante frío, a una temperatura ≤ 50 °C y alto vacío, ≤ 0.020 milibares (mBar), utilizando un liofilizador (Labconco® FreeZone®18, Serie 77550, USA). Posterior a este periodo, las muestras se colocaron en orden para cortar fragmentos de 2 mm de diámetro por hoja y se procedió a pulverizar las muestras usando un disruptor TissueLyser (Quiagen, Germany).

La extracción de ADN se realizó en el Laboratorio de Biotecnología, en el complejo Multi-Tecnológico Biociencias del CIMMyT, utilizando el método de purificación de ADN genómico en plantas a base de CTAB (bromuro de cetiltrimetilamonio) (CIMMyT, 2006), que tiene como referencia el protocolo estándar CTAB (Doyle, 1987).

La calidad del ADN obtenido se verificó mediante un gel de agarosa (skeaam) al 1 %, así mismo para la cuantificación fue requerido un nano-espectrofotómetro NanoDrop® 1400 (Thermo Scientific, Wilmington, DE).

El ADN obtenido se envió para servicio de genotipo, el cual se realizó a través de Next Generation Sequency (NGS), técnica usada originalmente mediante DArTSeq, (Diversity Arrays Technology) en la Universidad de Canberra (Australia) y aplicada por los servicios de SAGA (Servicio de Análisis Genético para la Agricultura) en las instalaciones del Centro Nacional de Recursos Genéticos (México). Las muestras de ADN fueron digeridas con la enzima de restricción ApeKI y analizadas en el secuenciador HiSeq2500 Illumina.

3.3. Análisis genético - molecular

El análisis de datos obtenidos mediante Genotipificación por Secuenciación se dividió en tres categorías principales:

Análisis primario, que comprende las etapas de filtro de calidad de ‘barcodes’ (P=30), filtro de calidad de secuencias completas (P=10), división de ‘targets’

según sus 'barcodes', identificación de 'tags' por muestras, generación de reporte de calidad, y visualización en LIMS/base de datos.

Análisis Secundario, que implica la selección de muestras de buena calidad para análisis (LIMS), DArTSoft Upload de datos en Beskidy (Software de análisis), alineación de 'tags' sobre genoma de referencia, identificación de SNPs, e identificación de polimorfismos.

Visualización de datos, que comprende la verificación de calidad de datos, verificación de reproducibilidad entre réplicas, selección de marcadores PAVs y SNPs polimórficos de alta calidad, así como la preparación de reportes.

Para este análisis se partió de dos primeros archivos, los datos PAV y los datos SNP se unieron para formar un solo archivo el cual se sometió a filtros para obtener un archivo en formato Excel el cual fue la referencia para el uso de R, con estas matrices se construyó la matriz de distancias con el método modificado de Roger (MRD), partiendo de los datos de disimilitud se realizó un clúster y PCA (análisis de componentes principales), de igual manera se realizó un análisis de MDS (Multidimensional scaling), así como un análisis básico conteniendo el resumen de la diversidad existente en este conjunto de muestras.

Para la visualización de los resultados de la diversidad genética y las distancias genéticas se usó el programa libre Curly Whirly proporcionado por el Instituto James Hutton, que es útil para la observación y análisis de datos multidimensionales en un espacio 3D, utilizando soporte de rotación, zoom, filtro jerárquico y una gama de interacciones con puntos.

3.4. Análisis geográfico

La distancia geográfica se determinó con base a la latitud y altitud de la región de colecta original de las accesiones, en primera instancia se realizó una matriz de distancias puras utilizando la fórmula:

$(90 - \text{LatitudA}) * \cos(\text{Radian}(90 - \text{LatitudB})) + \sin(\text{Radian}(90 - \text{LatitudA})) * \sin(\text{Radian}(90 - \text{LatitudB})) (\text{LongitudA} - \text{LongitudB})$

Se realizó un mapa de coordenadas geográficas previo utilizando herramientas de Google earth, partiendo de un archivo xlm, posteriormente se determinaron las coordenadas utilizando un conjunto de herramientas SIG (Sistema de Información Geográfica). Algunos datos geográficos se tomaron de las bases de datos del programa Germinate 3, del Centro Internacional de Mejoramiento de Maíz y Trigo, Internacional (CIMMYT, Int).

Germinate 3 es un sistema genérico de gestión de datos de planta diseñado para contener datos de pasaporte (i.e. de colecta) y una gama adicional de diferentes tipos de datos, incluyendo datos fenotípicos y marcadores moleculares, estas herramientas fueron desarrolladas por Paul Shaw y Sebastian Raubach.

3.5. Análisis de correlaciones

Los estudios de asociación entre datos obtenidos en un área de interés genético y provenientes de otras áreas permiten detectar interacciones relevantes de los genotipos. El estudio de la asociación de información fenotípica y genotípica se realiza frecuentemente con procedimientos univariados (Brachi *et al.*, 2010; Cockram *et al.*, 2010), considerando cada característica de interés separadamente. Sin embargo, la reaparición de aproximaciones holísticas y multivariadas en los últimos años ha generado un cambio de perspectiva en la biología (Carrari *et al.*, 2006), probablemente facilitado por importantes avances en el desarrollo de tecnologías e instrumental para la caracterización multidimensional de los distintos niveles de organización. Así, las ciencias “ómicas” como la fenómica y la genómica, tienen en común la búsqueda del estudio integral de los seres vivos, para lo cual las asociaciones de distintos tipos de datos son cruciales (Houle *et al.*, 2010). Las técnicas estadísticas del análisis multivariado permiten abordar los estudios de asociación multidimensional entre diferentes tipos de datos (Jombart *et al.*, 2009; Balzarini *et al.*, 2011; Teich *et al.*,

2011). sin embargo, aun cuando las técnicas multivariadas clásicas permiten operar sobre variables correlacionadas, suponen independencia (no correlación) entre las unidades de análisis. Una cuestión fundamental que debe considerarse en cualquier estudio de asociación es la existencia de estructuras de correlación en los datos. Cuando se lleva a cabo un análisis de asociación sin considerar los efectos de dependencia de los datos, se aumenta el riesgo de detectar asociaciones ilegítimas tanto como el de no detectar asociaciones (Keitt *et al.*, 2002). En la naturaleza es frecuente que la variabilidad genética presente estructuras espaciales, particularmente en especies de plantas (Vekemans y Hardy, 2004).

Uno de los procedimientos multivariados diseñados para cuantificar asociaciones no restringidas espacialmente es el análisis denominado Procrustes Generalizado, métodos que permite abordar la problemática de los datos analizando tablas múltiples (individuos x variables x condiciones). (Gower, 1975). A pesar de no ser técnicas de desarrollo reciente, su aplicación en el contexto de datos biológicos masivos sí lo es (Bramardi *et al.*, 2005).

No obstante, estos procedimientos de análisis de asociaciones multivariadas se basan en modelos lineales generales que suponen datos independientes, por lo que podrían producir resultados errados bajo auto correlación espacial. Otra aproximación al análisis de datos correlacionados espacialmente que se encuentra cada vez más difundida, se basa en el uso de modelos lineales mixtos (MLM) (Littell *et al.*, 2006). Los MLM han permitido integrar conceptos relativos a la estructura de correlaciones de cualquier tipo de datos normales en un marco analítico flexible y unificado. Si bien podrían aplicarse a variables sintéticas de distribución continua, como puede ser una componente principal o una variable canónica, su aplicación en el estudio de asociaciones de tipo multivariado es prácticamente nula (Teich, 2015).

4. RESULTADOS Y DISCUSIÓN

4.1. Análisis

4.1.1. Análisis genético-molecular

Las muestras fueron enviadas al Servicio de Análisis Genético para la Agricultura [SAGA], para ser analizadas con la tecnología de DArTseq con base en lo desarrollado por Elshire *et al.* (2011) y descrito en el inciso 2.8. Las muestras de ADN se digirieron con la enzima de restricción ApeKI y se analizaron en el secuenciador HiSeq 2500 Illumina.

De este proceso de secuenciación, se obtuvieron dos tipos de marcadores, los de presencia / ausencia (PAV) y los de frecuencias (SNP).

En lo referente al control de calidad, los marcadores con frecuencia del alelo menor (MAF, por sus siglas en inglés) de 0.05 fueron descartados, debido a que producen resultados inestables (Corvin *et al.*, 2010; Weng *et al.*, 2011). Con este primer filtrado se eliminaron también aquellos marcadores monomórficos, cuyo genotipo en todos los mestizos es el mismo (MAF = 0) y por consiguiente no explican la variabilidad en cuanto a nivel de severidad. Un segundo filtrado se realizó para descartar aquellos marcadores con más de 20 % de datos faltantes.

Se identificaron un total de 616,967 SNPs. Después del filtrado se obtuvo un total de 55,540 SNPs de alta calidad. Las bases de datos genotípicas originales están disponibles en <https://www.dropbox.com/s/fi8tplghvz65s8/MarkerSummary.csv?dl=0> (Hearne *et al.*, 2016). El Cuadro 4 es un ejemplo de los datos obtenidos de genotipificación DArTseq, en formato numérico.

Cuadro 4. Ejemplo de información genotípica en formato numérico.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	SNP	X518429	X518437	X688412	X688436	X688448	X688460	X43856	X520252	X518161	X518195	X518211	X518219	X518277	X518281
2	101949330 F 0--14:C>G	NA	NA	NA	0	0	0	NA	NA	NA	0	NA	NA	NA	NA
3	101949331 F 0--17:C>T	0	0	0	0	0	0	NA	0	0	0	0	0	0	0
4	101949332 F 0--20:G>A	0	0	0	0	0	0	NA	0	0	0	0	0	0	0
5	9713381 F 0--5:T>C	NA	NA	NA	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
6	9713380 F 0--9:T>C	NA	NA	NA	0.988235	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
7	9713380 F 0--30:A>G	NA	NA	NA	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
8	7060668 F 0--28:A>G	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	7059890 F 0--15:T>C	0	0	0	0	0	0	0	0	0	0	0	0	1	0.94
10	7059890 F 0--17:C>T	0.02439	0	0	0	0	1	0	0	0	0	0	0	0	0
11	9701239 F 0--5:T>G	NA	NA	NA	NA	NA	0	NA	NA	NA	NA	NA	NA	NA	NA
12	9701239 F 0--8:G>A	0	0	0	0	0	0.006579	0	0	0	0	0	0	0	0
13	5584491 F 0--11:A>T	NA	NA	0	0	NA	0	0	NA	0	0	0	0	0	0
14	5584491 F 0--17:G>T	NA	NA	0	0	NA	0	0	NA	0	0	0	0	0	0
15	101949373 F 0--6:C>A	NA	0	0	0	0	0	0	0	0	0	0	0	0	0
16	101949374 F 0--8:T>G	NA	0	0	0	0	0	0	0	0	0	0	0	0	0
17	101949373 F 0--12:A>G	NA	0	0	0	0	0	0	0	0	0	0	0	0	0
18	101949373 F 0--17:A>G	NA	0	0	0	0	0	0	0	0	0	0	0	0	0
19	101949373 F 0--18:G>A	NA	0	0.022222	0	0	0	0	0	0	0	0	0	0	0
20	101949373 F 0--22:A>C	NA	0	0	0	0	0	0	0	0	0	0	0	0	0
21	101949380 F 0--6:C>G	NA	NA	NA	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
22	101949386 F 0--25:G>A	0.777778	0.875	NA	0	0	0	0.421053	0.5	0.875	0.666667	1	0.6	0.971831	0.971831
23	101949387 F 0--50:G>C	0.777778	0.875	NA	0	0	0	0.421053	0.5	0.875	0.666667	0.6	0	0.042254	0.042254
24	9676003 F 0--13:C>G	NA	NA	NA	NA	NA	NA	NA	0	0	NA	NA	NA	NA	NA
25	101949395 F 0--30:C>A	NA	NA	NA	NA	NA	NA	NA	0	0	NA	NA	NA	NA	NA

Para los cálculos de distancias genéticas se utilizó el software libre estadístico R versión 3.3.1., previo un filtrado de datos en el cual se consideró la profundidad, que en la secuenciación del ADN se refiere al número de veces que un nucleótido se lee durante el proceso, así como el número de heterocigotos y el número de datos perdidos.

Para el análisis molecular se construyó la matriz de distancias de las 1965 accesiones mediante el método modificado de Roger (MRD). El Apéndice 3 muestra un ejemplo del contenido total de la matriz de 3,859,290 relaciones de distancias calculadas. En este se indica el genotipo y la interacción de cada genotipo con los demás, además de tratarse de una matriz de distancias euclidianas.

Partiendo de la matriz de distancias (MRD) se realizó un Análisis de Componentes Principales (ACP), y una muestra de este se representa en el Apéndice 4. Este tipo de análisis puede reducir efectivamente la estructura de una matriz de interacción Genotipo x Ambiente de G puntos (genotipos) en E dimensiones (ambiente) a un sub-espacio de pocas dimensiones.

En la Figura 5 puede apreciarse la dispersión de los genotipos considerando los primeros dos componentes principales. La finalidad de estos componentes es la de simplificar la estructura de los datos, sin obedecer a un modelo fijo, para poder explicar, en pocos componentes, la mayoría de la información que contienen las variables originales. En esta imagen puede verse que las dos razas de maíz evaluadas se encuentran claramente diferenciadas, aunque también se observan algunos genotipos intermedios que bien valdría la pena retomarlos y reevaluarlos dado que podrían ser producto de la hibridación entre las dos razas.

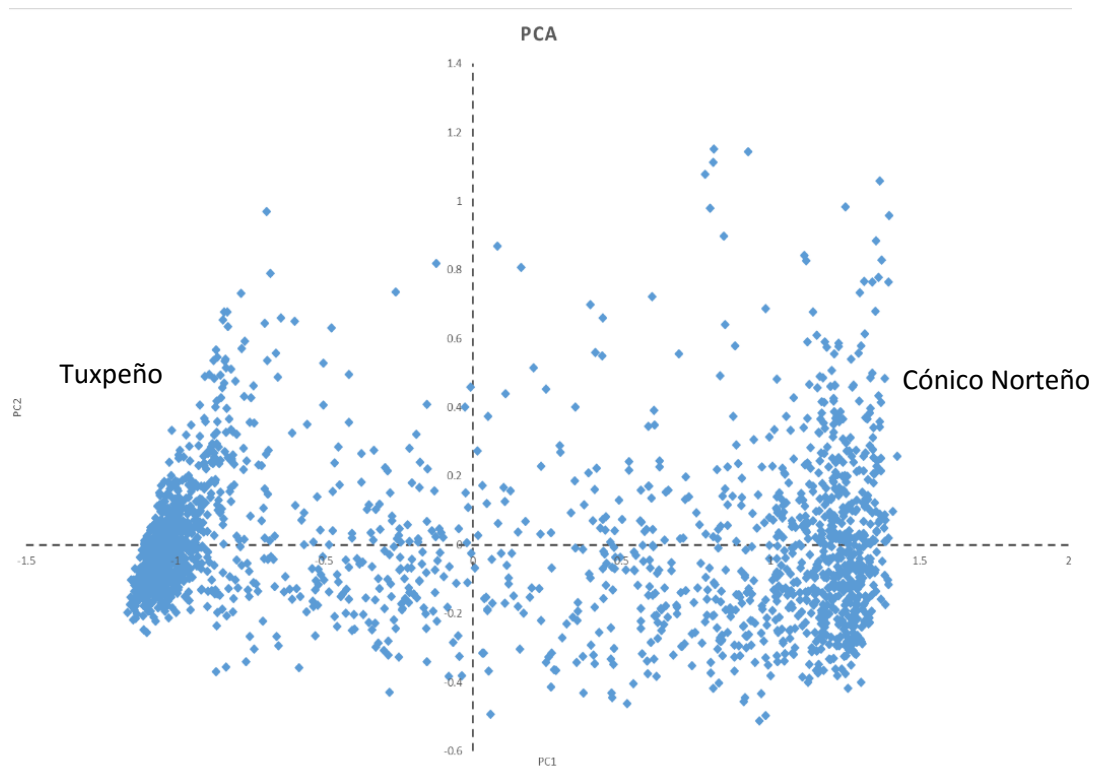


Figura 5. Dispersión de los genotipos correspondientes a las razas Tuxpeño y Cónico Norteño considerando los dos primeros componentes principales.

Cuevas (2003) menciona que teniendo en cuenta que el punto de partida del ACP es la matriz de correlación observada y que esta matriz anticipa la estructura y la varianza de los primeros pocos componentes (PC 's), vale la pena explorar esta matriz de correlación para buscar posibles grupos de variables con correlación alta o baja conforme Jolliffe (2002).

Un ACP frecuentemente revela relaciones insospechadas con anterioridad, razón por la cual se emplea como un paso metodológico intermedio en investigaciones a largo plazo, en la que, métodos tales como la regresión múltiple o los análisis de agrupamiento pudiesen ser pertinentes en etapas posteriores.

Por lo anterior se construyó una matriz de asociación o clúster, de la cual una muestra se observa en el Apéndice 5, para posteriormente construir un diagrama de asociación o dendrograma como se muestra en la Figura 6.

Los métodos de agrupamiento utilizan un algoritmo secuencial al agrupar secuencias individuales o grupos de secuencias, usualmente referidos como unidades taxonómicas operacionales (OTUs), que son más similares entre ellos, esto es, para los que la distancia genética es menor. Cuando se agrupan dos OTUs, se tratan como una nueva unidad OTU de manera recursiva para inferir la estructura, así las 1965 accesiones procedentes de las dos razas en estudio se agruparon a su raza respectiva, pero de igual manera algunas accesiones se sub-agruparon en distintos conjuntos por región.

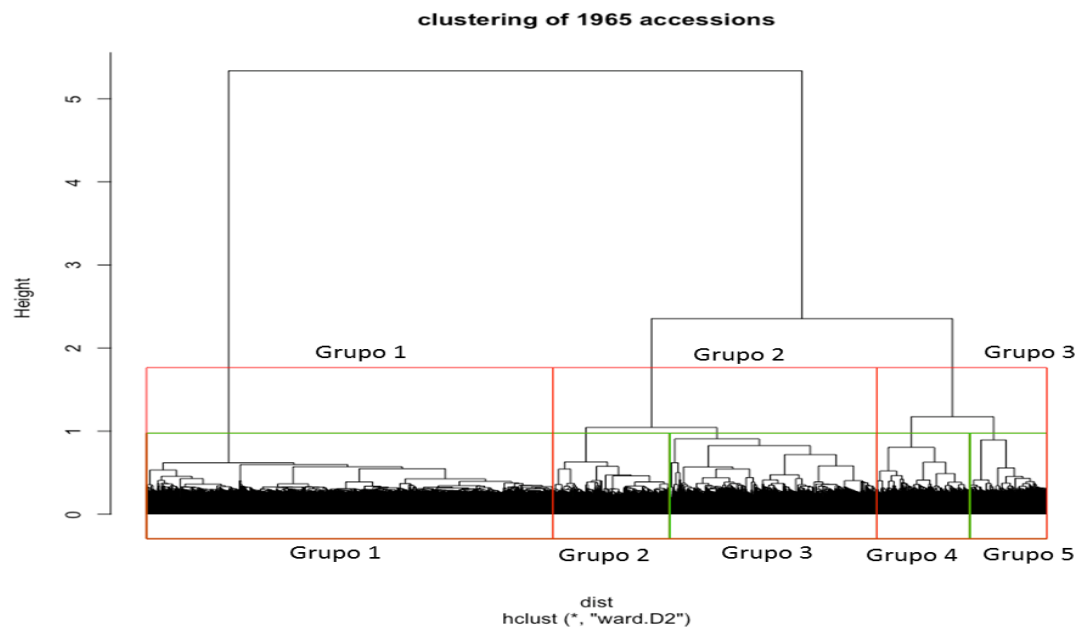


Figura 6. Dendrograma de relaciones entre los 1965 genotipos evaluados, con una línea de corte a tres grupos (rojo) y cinco grupos (verde).

El análisis de agrupamiento partiendo de la matriz MRD, que se muestra en la Figura 6, con una línea de corte en 3 grupos, encerrados en rojo, indicó que el primer grupo se compone de 887 genotipos de los cuales 882 pertenecen a la raza Cónico Norteño y cinco a Tuxpeño, el segundo grupo compuesto por 707 individuos incluyó 671 Tuxpeños y 36 Cónicos Norteños, mientras que el tercer grupo se formó con 371 individuos y contiene 217 Cónicos Norteños y 154 Tuxpeños.

Cuando se usó el criterio en un corte a 5 grupos como se observa en verde, el primer grupo en donde abundan los Cónicos Norteños permanece igual que con el corte anterior con 882 Cónicos Norteños y cinco Tuxpeños, el segundo grupo con 255 de los cuales 253 pertenecen a Tuxpeño y dos a Cónico Norteño, el tercer grupo abarcó 452 accesiones de las cuales 418 fueron de Tuxpeño y 34 de Cónico Norteño, el cuarto grupo contiene 203 subdivididos en 146 Tuxpeños y 57 Cónicos Norteños, en tanto el quinto grupo se formó con 168 y contiene 160 Cónicos Norteños y 8 Tuxpeños, siendo de esta manera los grupos 1 y 5 en los cuales se encuentra la mayor cantidad de individuos pertenecientes a la raza Cónico Norteño.

Se utilizaron una serie de herramientas tecnológicas desarrolladas en el Instituto Hutton con el objetivo de obtener una representación de datos biológicos complejos (Figura 7).

Las técnicas de análisis multivariado como el Análisis de Componentes Principales (ACP) y el Escalamiento Multidimensional (MDS) se utilizan a menudo con datos genéticos de poblaciones para producir "Mapas estadísticos" de individuos o poblaciones muestreados (Menozzi *et al.*, 1978; Zhivotovsky *et al.*, 2003; Patterson *et al.*, 2006). Por ello se realizó un ACP en primera instancia, pero debido a que este tipo de técnicas conservan principalmente la varianza y no la distancia que es el objeto de interés en este estudio se optó por realizar un MDS (Figura 7).

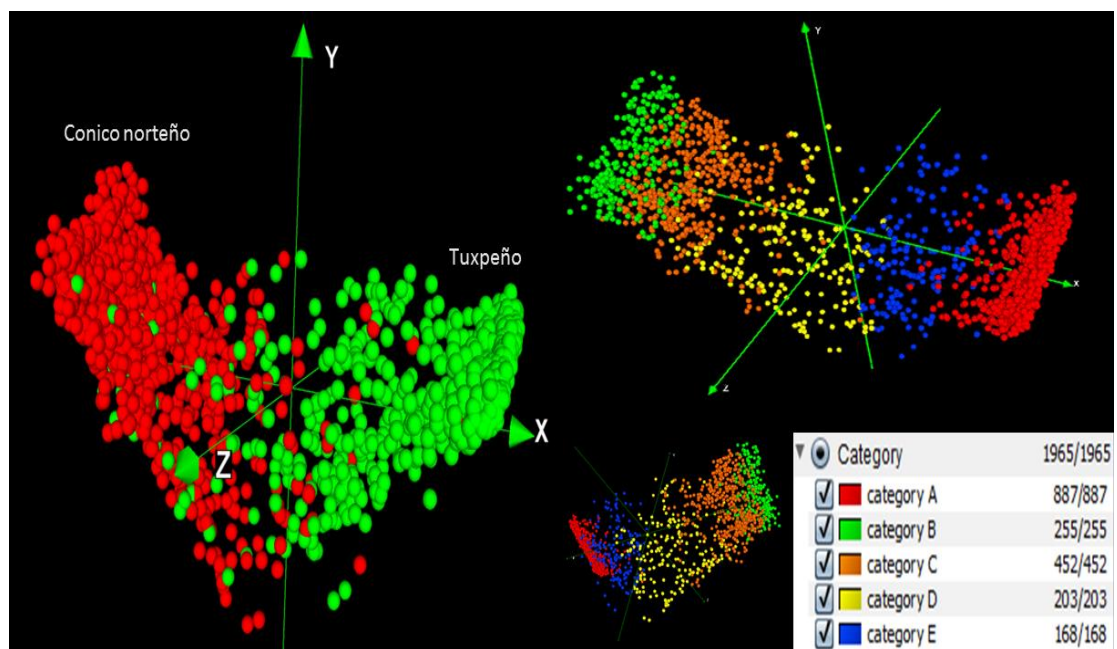


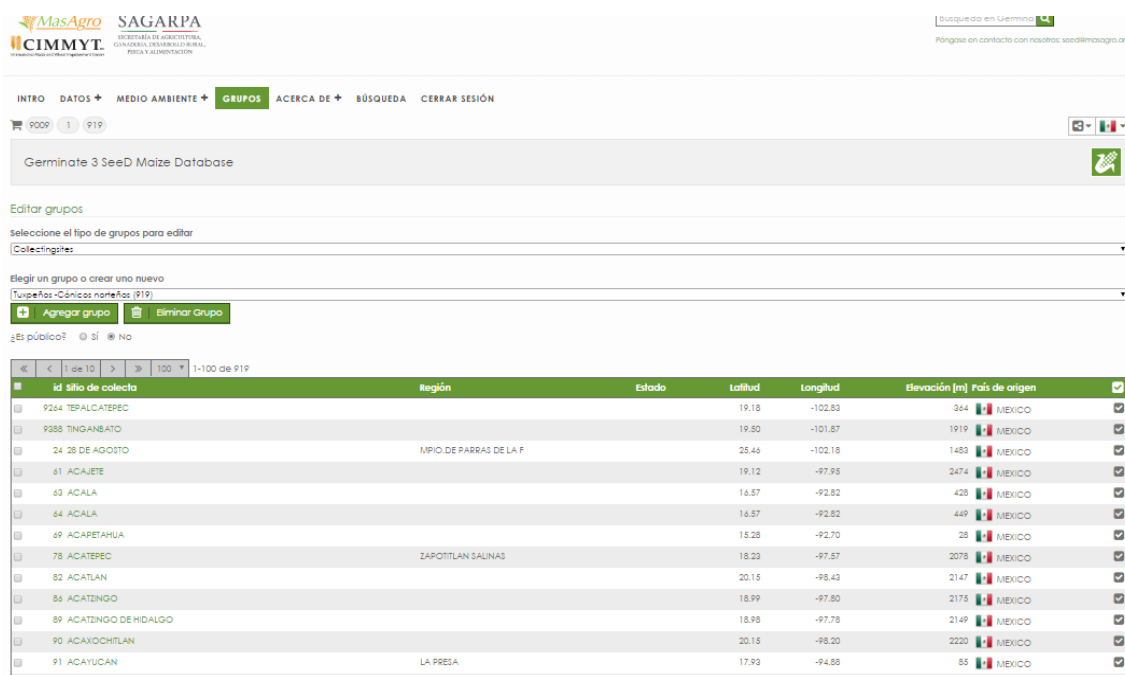
Figura 7. Representación de las distancias genéticas entre accesiones de las razas Tuxpeño y Cónico Norteño por el método MDS, mediante el software Curly Whirly.

La Figura 7 muestra una separación muy clara entre las accesiones de las dos razas, aunque existen algunas que se traslapan y que podrían considerarse híbridos, entre los que se podrían considerar como híbridos existe toda una gama donde algunos de ellos como los naranjas podrían portar mayor cantidad del genoma de Tuxpeño, mientras que en azul se encuentran los que poseen mayor similitud con Cónico Norteño, siendo los amarillos los más intermedios entre estas dos razas

Todo esto puede deberse a diversos contextos como la movilidad de las semillas por caminos y carreteras, en un componente tiempo.

4.1.2. Análisis geográfico

En la parte correspondiente a los datos geográficos, con base en la clave GID (Identificación de Germoplasma), se realizó una búsqueda mediante la plataforma Germinate 3 (Figura 8), en la cual se logró obtener datos tales como sitios de colectas, longitud, latitud, nombre de la colecta y país, entre otros.



#	id sitio de colecta	Región	Estado	Latitud	Longitud	elección [m]	País de origen	
9264	TERALCATEPEC			19.18	-102.83	384	MEXICO	
9285	TINGANBATO			19.30	-101.87	1919	MEXICO	
24	28 DE AGOSTO		MPIO. DE PARRAS DE LA F	25.46	-102.18	1483	MEXICO	
61	ACAJETE			19.12	-97.95	2474	MEXICO	
63	ACALA			16.57	-92.82	428	MEXICO	
64	ACALA			16.57	-92.82	449	MEXICO	
69	ACAPETAHUA			15.28	-92.70	28	MEXICO	
78	ACATEPEC		ZAPOTITLAN SALINAS	18.23	-97.57	2078	MEXICO	
82	ACATLAN			20.15	-98.43	2147	MEXICO	
84	ACATZINGO			18.99	-97.80	2175	MEXICO	
89	ACATZINGO DE HIDALGO			18.98	-97.78	2149	MEXICO	
90	ACAYOCHITLAN			20.15	-98.20	2220	MEXICO	
91	ACAYUCAN		LA PRESA	17.93	-94.58	55	MEXICO	

Figura 8. Página de la base de datos Germinate 3.

Después de obtener la información se realizó un filtrado rescatándose aquellas accesiones o muestras que describieran datos geográficos y genéticos completos, que pertenecieran a las razas Tuxpeño y Cónico Norteño de México (Apéndice 6).

Para posteriormente realizar el cálculo de las distancias puras entre dos puntos determinados (Apéndice 7) entre cada accesión partiendo de la fórmula: $(90 - \text{LatitudA}) * \cos(\text{RADIAN}(90 - \text{LatitudB})) + \sin(\text{RADIAN}(90 - \text{LatitudA})) * \sin(\text{RADIAN}(90 - \text{LatitudB})) (\text{LongitudA} - \text{LongitudB})$.

La fórmula utilizada para determinar la distancia más corta entre dos puntos de la tierra (geodésica) se aproxima al geoide de una esfera de radio $R = 6372,795$ kilómetros (radio cuadrático medio), por lo que el cálculo podría tener un error de distancia del 0.3 %, sobre todo en los extremos polares, y para las distancias largas a través de diversos paralelos. Dados dos puntos A y B en la esfera se indican por su latitud (lat.) y longitud (long.).

Un problema común que suele presentarse al momento de calcular la trayectoria (distancia), es que por lo general se realiza el cálculo de la distancia horizontal más no la topográfica que incluye valores de la pendiente, por ello con base en las coordenadas geográficas, las distancias calculadas previamente de los sitios de colecta y con el soporte del equipo de Sistema de Información Geográfico de CIMMyT, utilizando el software ArcGIS se identificaron las distancias reales (Apéndice 8), que consideran aspectos como el relieve entre cada sitio de colecta de la accesión.

Los resultados de las distancias considerando el relieve o elevaciones se muestran en el Apéndice 9. La matriz de distancias calculadas reales, indicadas como distancias euclidianas e identificadas desde 0 hasta 1965, donde 0 es el punto de origen o primer sitio considerado para esta evaluación, se reportan en metros y kilómetros respectivamente. La matriz consta de 3,861,225 distancias.

Para la representación gráfica se utilizó un primer plano en Google Earth (Figura 9) partiendo de datos del Apéndice 10, el cual se muestra en formato Excel, para transformarlo en un archivo con extensión kml: la primer columna indica la latitud, la segunda columna la longitud, la siguiente el nombre o identificación de los genotipos, posteriormente se coloca una descripción que corresponde con la identificación que se le asignó previo al análisis, y finalmente se indica un código para identificar en el mapa, esto regularmente es a criterio de cada usuario.

En un primer plano gráfico fue posible determinar que, para los 1965 individuos analizados, existen sólo 815 sitios de colecta.

Geográficamente los sitios de colecta de las 1965 accesiones se encuentran distribuidos dentro de la República Mexicana como se muestra en la Figura 9.



Figura 9. Mapa de distribución geográfica de los sitios de colecta, generado en Google Earth.

El segundo método utilizado para representar la distribución geográfica de los sitios de colecta fue mediante las herramientas de ArcGIS, el software utilizado fue ArcMap que generó el mapa mostrado en la Figura 10 donde se indica con puntos en color naranja y negro los 815 sitios, de los cuales fueron colectados los 1965 individuos analizados.

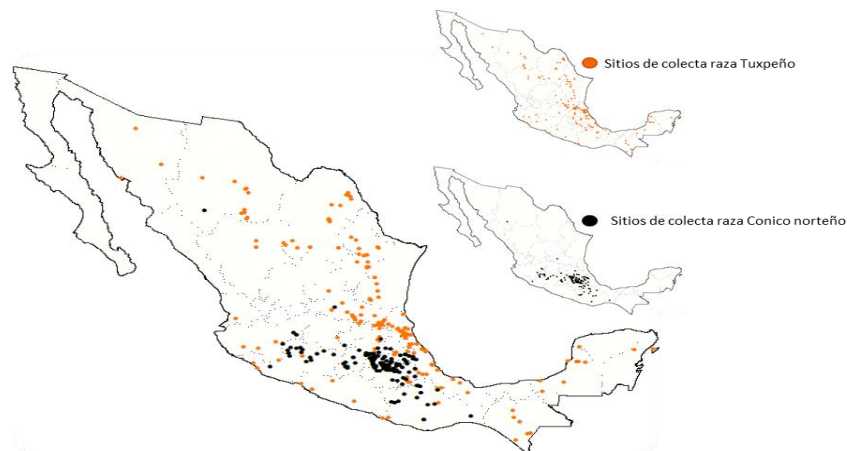


Figura 10. Mapa de distribución geográfica de los sitios de colecta, generado en ArcMap.

Como se observa en el mapa, la costa del Golfo de México es la más poblada por estas dos especies, principalmente; como se sabe la raza Tuxpeño está muy adaptada a altitudes bajas. Además, haciendo honor al nombre y el origen de tal raza, su distribución se observa desde Yucatán por toda la llanura costera al norte de la ciudad de Veracruz, siendo precisamente la posición geográfica de la ciudad de Tuxpan, en donde se localiza el centro de dispersión de esta raza. No obstante, también se pueden observar genotipos Tuxpeños en zonas de Sonora, Chihuahua y Coahuila, que corresponde a lo descrito en la literatura (Figura 11A). Así mismo se observan en el área del Bajío los genotipos correspondientes a la raza Cónico Norteño, principalmente en los estados de Jalisco, Aguascalientes, el norte de Guanajuato y Querétaro (Figura 11B), que se encuentran regularmente en altitudes de entre 1500 y 2000 metros. Correspondiendo con la descripción de esta raza que está adaptada a elevaciones medianamente altas.

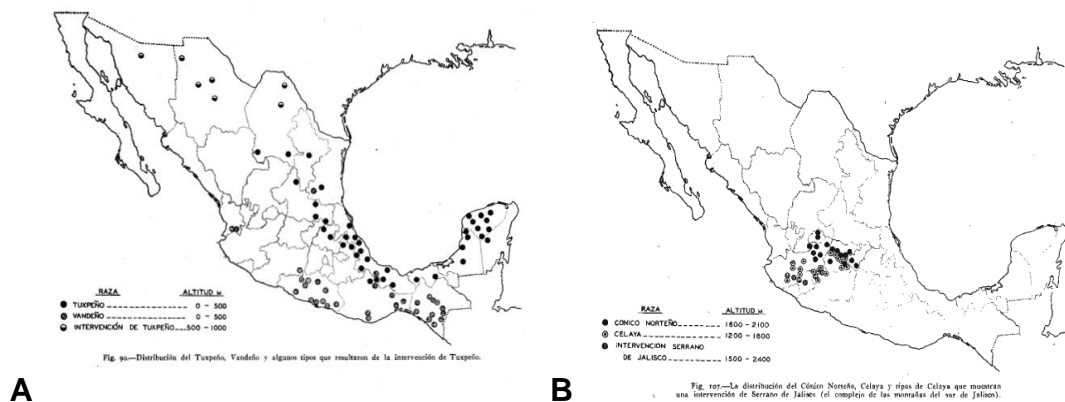


Figura 11. Distribución de las razas en la República Mexicana. A, Tuxpeño y B, Cónico Norteño, según lo descrito por Wellhausen y colaboradores en 1951.

Wellhausen *et al.* en 1952 proponen que la raza Tuxpeño al hibridarse con la raza Cónico dio lugar a la raza Chalqueño, hipótesis que para John F. Doebley no parece prudente analizar, simplemente porque Chalqueño es morfológicamente similar a Cónico, y ambas distintas a Tuxpeño. Cónico y Chalqueño son razas adaptadas a altitudes elevadas, mientras que Tuxpeño es una raza adaptada a

bajas altitudes. Dadas estas consideraciones, información aloenzimática ha demostrado que Cónico y Chalqueño se encuentran estrechamente relacionados y que esta última raza no es intermedia entre Tuxpeño y Cónico Norteño.

Wellhausen *et al.* (1952) exponen una serie de hipótesis muy explícitas concernientes a las relaciones evolutivas existentes entre las razas de maíz de México y concluyen que la mayoría de las razas de maíz mexicanas surgieron a partir de la hibridación entre otras razas. En la Figura 12 se presenta la distribución de accesiones de la raza Chalqueño en la República Mexicana, de las cuales se posee información geográfica.

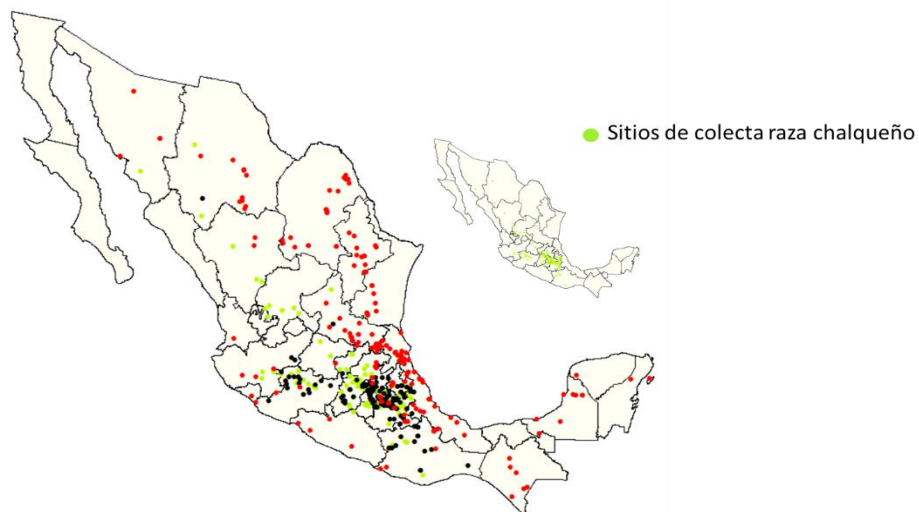


Figura 12. Mapa de distribución geográfica de los sitios de colecta de la raza Chalqueño, generado en ArcMap.

La Figura 12 muestra que las accesiones de la raza Chalqueño geográficamente se localizan muy cercanas a las de Cónico Norteño, pero también se observa algunas muestras en zonas elevadas cercanas a Tuxpeño, de tal modo que en términos geográficos la hibridación entre las dos razas podría ser factible, sin embargo es necesario llevar a cabo estudios de tipo genético para corroborar la teoría de hibridación entre estas razas.

4.2. Análisis comparativo entre distancias genéticas y geográficas

Para llevar a cabo la comparación entre distancias genéticas y geográficas se realizó una alineación análoga de coordenadas calculadas a partir de datos genéticos a lugares de muestreo geográfico, para muestras colectadas en México.

En un análisis de secuencias de SNPs en 1965 individuos de 815 sitios de colecta, se obtuvo un diagrama de MDS bidimensional sobre la base de una matriz de distancias genéticas de compartición de alelos de pares individuales. Cualitativamente, el MDS trazado parecería un mapa geográfico de los lugares de muestreo, con los ejes correspondientes en gran medida a la latitud y longitud. Este mismo fenómeno es visible en el trabajo de Li *et al.* (2008) y Biswas *et al.* (2009).

Para evaluar cuantitativamente la relación entre distancias genéticas y distancias geográficas se realizó un análisis de Procrustes-transformado considerando las coordenadas MDS de los SNP para producir una alineación óptima con las coordenadas geográficas.

Para este análisis entonces se utilizaron las coordenadas de los sitios de colecta de un análisis MDS interponiéndola sobre la matriz de distancias genéticas de las dos razas. Se utilizó el programa R para obtener la matriz de distancias genéticas de alelos entre las dos razas. Se aplicó un análisis MDS a la matriz, utilizando el comando `cmd scale` en R (Ihaka y Gentleman, 1996). Para las coordenadas geográficas, se utilizaron latitudes y longitudes (ArcGIS). Se graficó las distancias genéticas (Figura13) y las distancias geográficas (Figura14), con los datos ordenados de 0 -1964, de las matrices de MDS para conocer el comportamiento general de la población en estudio.

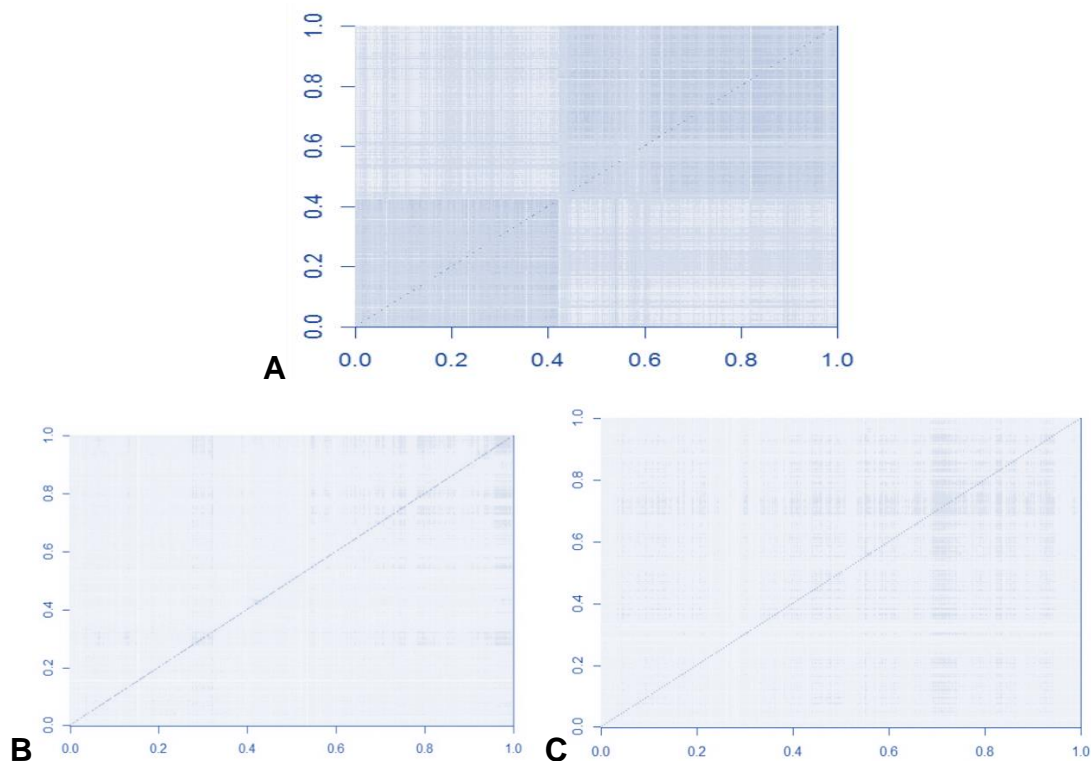


Figura 13. Comparación de distancias genéticas entre las razas de maíz Tuxpeño y Cónico Norteño. A, como un conjunto. B, raza Tuxpeño ordenados de 0-830. C, raza Cónico Norteño ordenados de 0-1135.

La distancia genética puede estimarse por medio de muy diversas funciones, dependiendo del tipo de marcador molecular que se utilice y del tamaño de la población. En la Figura 13 se observa la comparación entre un análisis del conjunto de datos (Tuxpeños-Cónicos Norteños) y las razas por separado. En lo concerniente a estas imágenes (Figura 13) puede verse que ambas razas se agrupan genotípicamente como se esperaba; es decir, por razas.

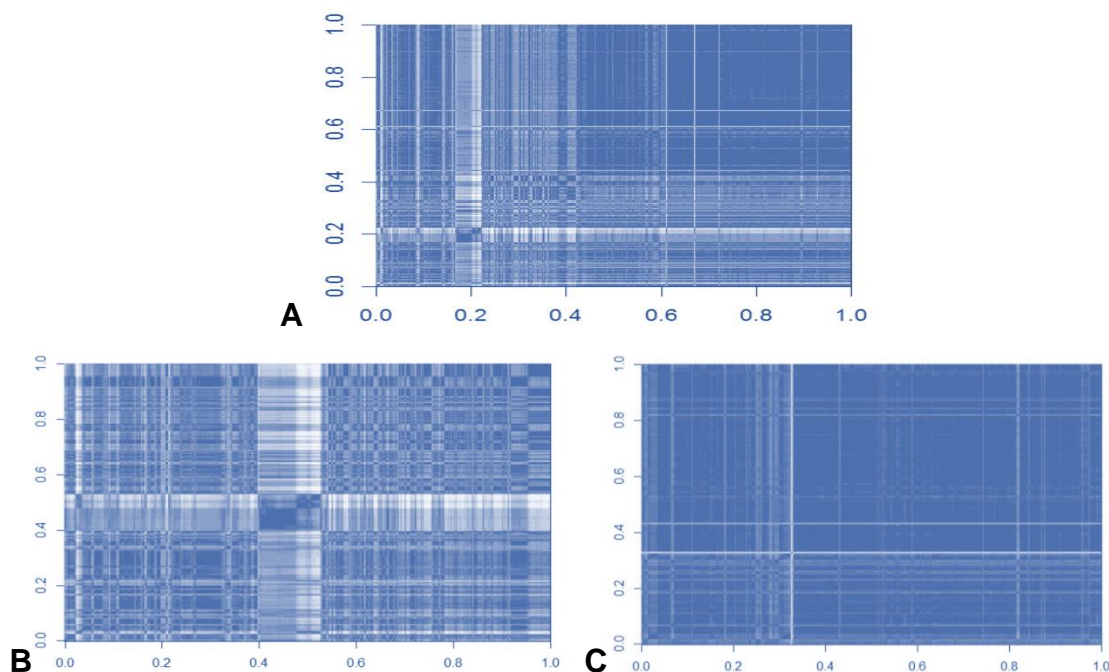


Figura 14. Imagen de distancias geográficas. A, como un conjunto de datos ordenados de 0 -1964. B, raza Tuxpeño ordenados de 0-830. C, raza Cónico Norteño ordenados de 0-1135.

Por otro lado, la distancia geográfica comúnmente se mide con base en la distancia euclidiana —distancia lineal entre 2 puntos—, aunque existen también múltiples algoritmos diferentes que pueden aplicarse. Como en el caso de este análisis (Figura 14) mediante algoritmos de ArcGIS, las variables se representan en un mapa raster, con datos obtenidos de distintas fuentes, y el modelo como tal se realiza en programas de sistemas de información geográfica (SIG), por medio de extensiones como Spatial Analyst en ArcGis v.10.

En cuanto a los atributos involucrados, es decir para representar en este caso la distancia genética como medida de la divergencia genética entre los genotipos y la distancia geográfica como la distancia entre dos puntos sobre la esfera terrestre (como distancia lineal entre dos posiciones dadas (latitud + longitud)), en estas razas de maíz, se realizó un histograma de frecuencias de distancias genéticas y de distancias geográficas (Figuras 15).

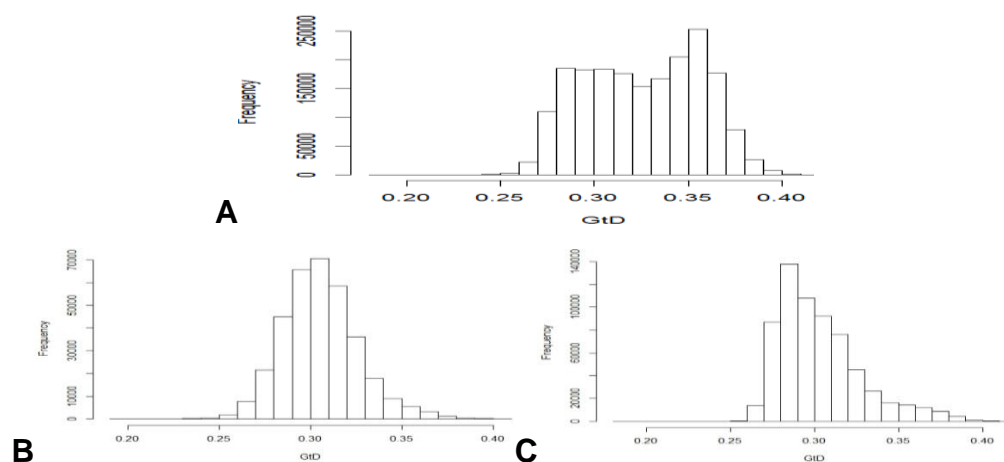


Figura 15. Histograma de distribución de frecuencias de distancias genéticas. A, datos ordenados de 0 -1964. B, distribución de frecuencias de distancias genéticas; raza Tuxpeño ordenados de 0-830. C, distribución de frecuencias de distancias genéticas; raza Cónico

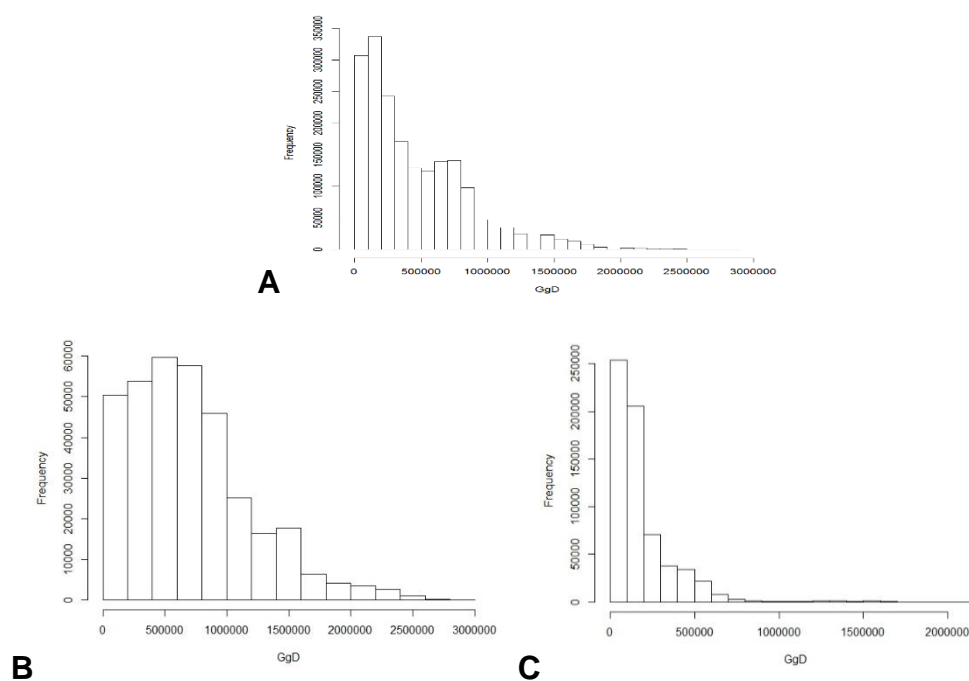


Figura 16. Histograma de distribución de frecuencias de distancias geográficas. A, datos ordenados de 0 -1964. B, distribución de frecuencias de distancias geográficas, raza Tuxpeño ordenados de 0-830. C, distribución de frecuencias de distancias geográficas; raza Cónico Norteño ordenados de 0-1135.

Estos gráficos (Figura 15 y 16) permiten visualizar la distribución de las frecuencias de distancias genéticas y distancias geográficas, representando una función de distribución de frecuencia de variables continuas, cuyos intervalos representan los valores observados y las alturas el número de observaciones en cada intervalo, así observamos que en las frecuencias de distancias genéticas el valor máximo de este análisis fue 0.36 aproximadamente, mientras que para las geográficas los valores máximos se presentan en rangos menores de 500 000 m.

La correlación entre matrices de distancia, genética y geográfica implementa la identidad por descendencia e incluye el efecto del tamaño de la muestra, factores mutacionales y demográficos, al considerarse estos nos permite obtener distancias geográficas más precisas, considerando las variables del entorno.

La Figura 17 muestra que existe poca relación entre las distancias geográficas y genéticas, antes de realizar una prueba de Mantel.

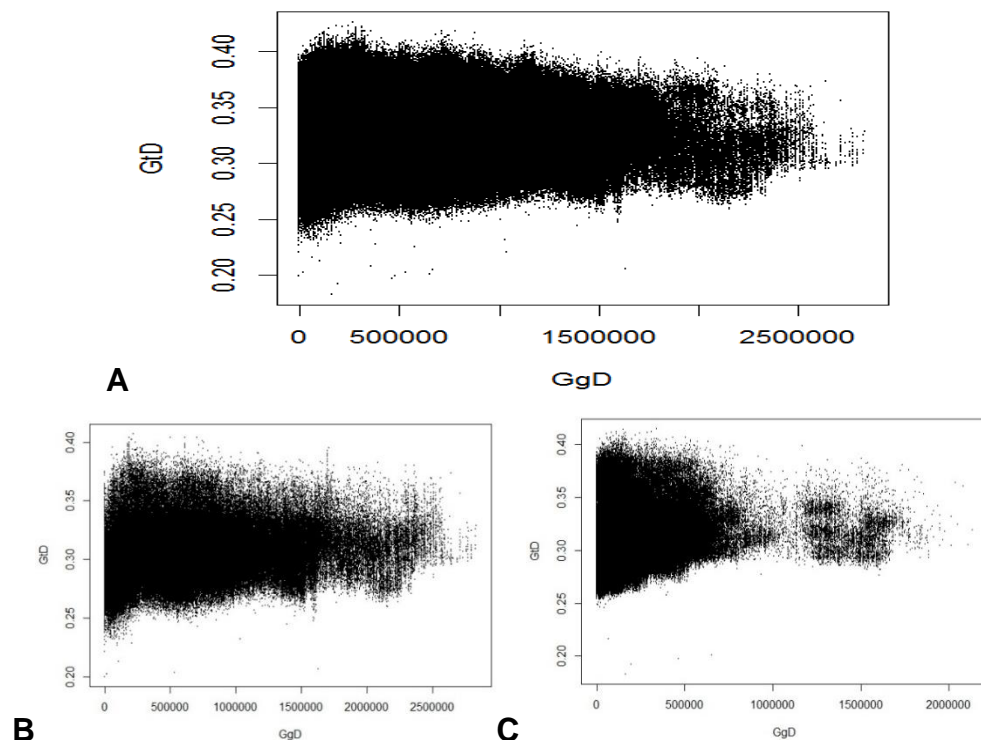


Figura 17. Gráfica de correlación entre las distancias genéticas (Y) y las distancias geográficas (X). A, del conjunto de datos. B, gráfica de correlación raza Tuxpeño. C, raza Cónico Norteño.

Asimismo, nuevos métodos de análisis permiten obtener tanto distancias genéticas como geográficas más precisas, por ejemplo, distancia entre puntos no de forma euclidiana sino considerando aspectos, como la topografía, lo que se traduce en una medida de distancia más real que la lineal entre 2 puntos (Landguth y Cushman, 2010; Cushman *et al.*, 2012).

El análisis estratificado es un tipo de análisis estadístico que estudia la posible asociación entre dos variables categóricas; pero diferenciando sub-muestras de supuestos individuos que presenten distintas categorías en otra tercera variable. Se realizó un análisis de correlación de Pearson que informó sobre la existencia de asociación entre las variables en estudio. Se hizo un análisis estadístico de correlación entre dos matrices del mismo rango, para comparar dos matrices de distancia. En este estudio se comparó también una matriz de distancia genética entre individuos y una matriz de distancias geográficas entre los mismos.

Uno de los métodos clásicos para evaluar aislamiento por distancia es la prueba de Mantel, que estima la correlación entre matrices de distancia, en nuestro caso particular de distancia genética y distancia geográfica. Si el coeficiente de correlación es positivo, indica que existe estructura espacial que puede relacionarse a la distancia geográfica (Guillot *et al.*, 2009).

Indudablemente en el programa R es sencillo realizar esta prueba, en el cual, comúnmente es utilizado el coeficiente de correlación producto-momento de Pearson. Esta prueba nos informa si la relación entre ambos es semejante o no, teniendo en cuenta que la hipótesis nula (H_0) de esta prueba es que las distancias poseen ausencia de correlación.

Los resultados de la aplicación del estadístico o prueba de Mantel al conjunto de datos basado en la correlación de Pearson (producto-momento), se especifican a continuación:

Mantel (xdis = GDm, ydis = GeD), estadístico Mantel r: **0.2329**, significancia: 0.001.

Cuantiles superiores de permutaciones (modelo nulo):

90 % = 0.0119, 95 % = 0.0150, 97.5 % = 0.0173, 99 % = 0.0190.

Número de permutaciones: 999.

Aplicación de prueba de Mantel, a la raza Tuxpeño, basado en la correlación de Pearson (producto-momento), estadístico Mantel r: **0.1252**, significancia: 0.001.

Cuantiles superiores de permutaciones (modelo nulo):

90 % = 0.0264, 95 % = 0.0342, 97.5 % = 0.0421, 99 % = 0.0477.

Aplicación de prueba de Mantel, a la raza Cónico Norteño, basado en la correlación de Pearson (producto-momento), estadístico Mantel r: **0.2777**, significancia: 0.001.

Cuantiles superiores de permutaciones (modelo nulo):

90 % = 0.0340, 95 % = 0.0457, 97.5 % = 0.0550, 99 % = 0.0628.

El análisis proporciona la estadística de correlación r de Mantel, y el p-valor bilateral correspondiente, obtenido a partir de las 999 permutaciones posibles. Como lo indican los resultados, se observa que se debe rechazar la hipótesis nula de ausencia de correlación. En estos términos, se puede considerar que las distancias genéticas y geográficas están correlacionadas en una proporción baja (0.2329), pero positiva.

Con el fin de correlacionar las separaciones físicas (Figura 18) y las distancias genéticas de las razas Tuxpeño y Cónico Norteño (Figura 19), se elaboró una matriz utilizando las distancias geográficas entre los 1965 genotipos (Apéndice 8). Al comparar el dendrograma de distancias genéticas de los SNPs (Figura 6) y el producido con las distancias geográficas, se observa que las accesiones con menor separación física coinciden con las accesiones de menor distancia genética; sin embargo la correlación entre ambas matrices, a pesar de alguna similitud no es significativa ($r = 0.05424$).

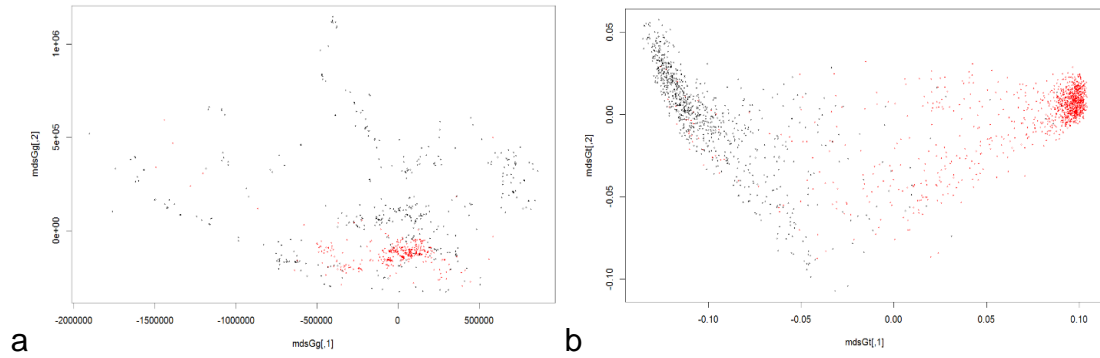


Figura 18. Agrupamiento de accesiones de las razas Tuxpeño y Cónico Norteño. A, con base en el MDS de las distancia geográfica. B, con base en el MDS de las distancias genéticas. Rojo: Cónico Norteño. Negro: Tuxpeño.

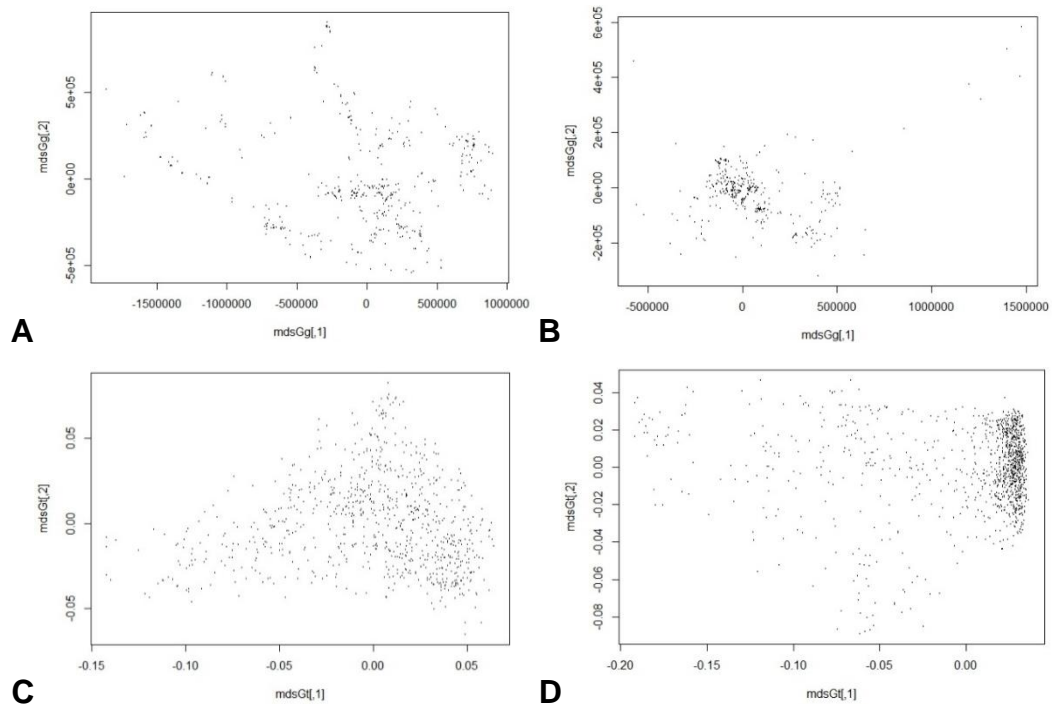


Figura 19. Agrupamiento de accesiones de las razas Tuxpeño y Cónico Norteño. A y B, gráficos con base en el MDS de las distancia geográfica. A. Raza Tuxpeño. B. Raza Cónico Norteño. C y D, gráficos con base en el MDS de las distancias genéticas. C. Raza Tuxpeño. D. Raza Cónico Norteño.

Usualmente se correlacionan datos genéticos con fenotípicos, no obstante el uso de datos de diversa naturaleza se ha empleado recientemente, rompiendo la linealidad de los datos, es decir, las mediciones a través del rango de mediciones esperado se ven afectados. Trabajos recientes, sugieren que bajo este tipo de violaciones del modelo lineal general, las pruebas de Mantel, tanto simples como parciales, producen un exceso de tasa de error tipo I cuya magnitud se incrementa con la intensidad de la autocorrelación espacial (Raufaste y Rousset 2001; Guillot y Rousset 2012).

Como lo menciona Wang (2010), en el análisis procrustes se consideran las matrices, $X = (x_1, \dots, x_n)^T$ e $Y = (y_1, \dots, y_n)^T$. X es $n \times p$, y cada fila en X corresponde a un punto n , en R_p ; Y es $n \times q$, y cada fila en Y corresponde a un punto n , en R_q . Los puntos se encuentran ajustados, de modo que X_r e Y_r representan vectores de coordenadas en las matrices R_p y R_q , respectivamente.

Las matrices se pueden ver como la descripción de dos conjuntos separados de coordenadas para la misma tasa n (dos "mapas"). Los resultados de un análisis procrustes al conjunto de datos se indican a continuación:

Análisis procrustes ($X = \text{mdsGg}$, $Y = \text{mdsGt}$),

Número de objetos: 1965, Suma de cuadrados procrustes: $3.553874e^{14}$.

Cuadrado medio del error procrustes: 425274.9.

Cuartiles de errores procrustes: mínimo: 13546.23, 1Q 93433.52, mediano: 162593.84, 3Q 397844.56, máximo: 1906689.49.

Rotación de la matriz: [1,] **0.06297703 -0.99801498**, [2,] **0.99801498 0.06297703**

Traslación de los promedios: [1,] **-3.735327e-10 2.085862e-10**

Escalado del objetivo: [1] **1752094.**

El método de procrustes trata de encontrar las transformaciones, cuando se desconoce f y g , ellos minimizan una función $d(f(X), g(Y))$ en todas las opciones f y g que preservan la paridad relativa de las distancias entre puntos en X y entre puntos en Y , respectivamente.

Primero, $|P - q|$. Las columnas cero se agregan al final de la matriz con menos columnas en orden, para colocar ambos conjuntos de puntos en el mismo espacio k -dimensional, con $k = \max(p, q)$.

Por lo tanto, tanto X e Y se convierten en matrices $n \times k$. Sin pérdida de generalidad, $g(Y) = Y$ se puede suponer, por lo que sólo X se transforma. La transformación f puede escribirse ya que $f(xr) = ATxr + b$, donde λ es una dilatación escalar, a es una matriz ortogonal $k \times k$, representando una rotación y posiblemente una reflexión, y b es un vector de traslación $k \times 1$.

La función objetivo d para ser minimizada es la suma a través de la tasa cuadrada de las distancias euclidianas entre coordenadas correspondientes de los taxa en las matrices $F(X)$ e Y .

En la Figura 20 se muestra el último de los pasos del proceso procrustes que consiste en la Rotación de coordenadas de la primer configuracion (Y), matriz genética, hacia la referencia (X) ó matriz geográfica.

Raza Tuxpeño - Análisis Procrustes ($X = \text{mdsGg}$, $Y = \text{mdsGt}$).

Rotación de la matriz: [1,] **0.9425002 0.3342056** [2,] **-0.3342056 0.9425002**

Traslación de los promedios: [1,] **5.029356e-11 -3.50081e-10**

Escalado del objetivo: [1] **6051372**

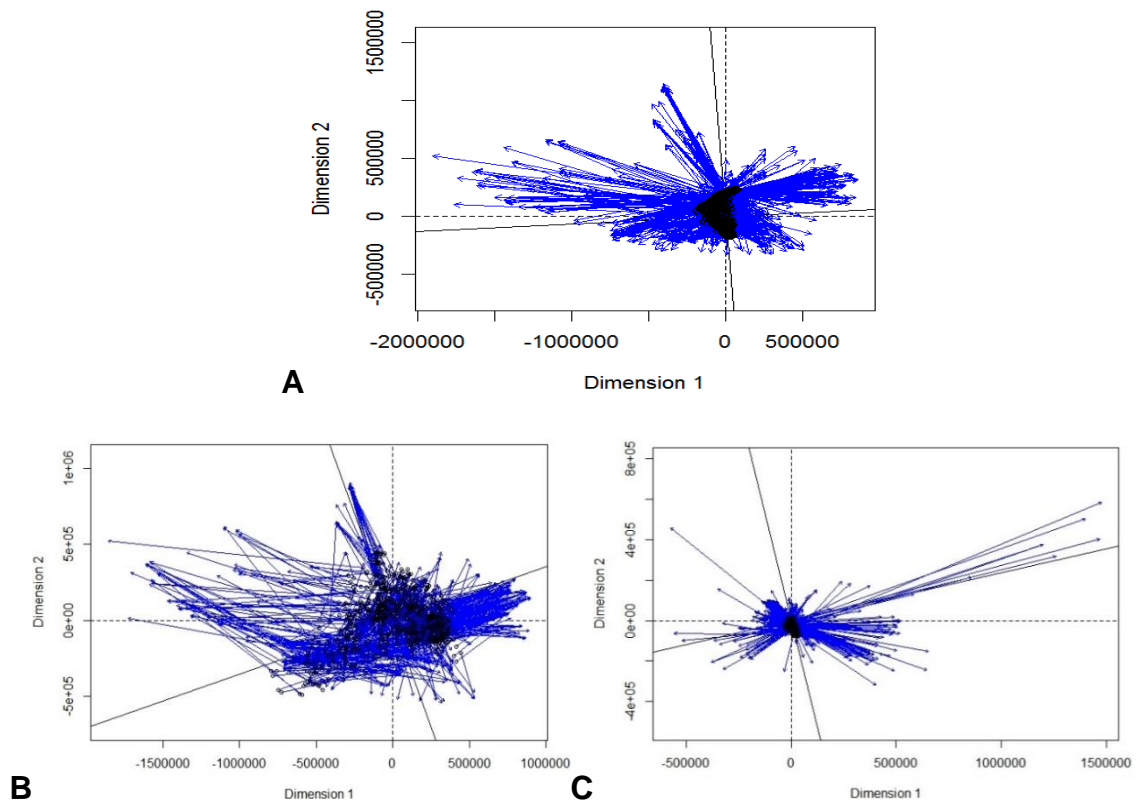


Figura 20. Rotacion de coordenadas en el análisis procrustes, aplicado con la técnica MDS. A, del conjunto de datos. B, raza Tuxpeño. C, raza Cónico Norteño.

Raza Cónico Norteño – Análisis Procrustes ($X = \text{mdsGg}$, $Y = \text{mdsGt}$).

Rotación de la matriz: [1,] **-0.2309835 0.9729577** [2,] **0.9729577 0.2309835**.

Traslación de los promedios: [1,] **9.179656e-12 6.14088e-11**.

Escalado del objetivo: [1] **427263.4**.

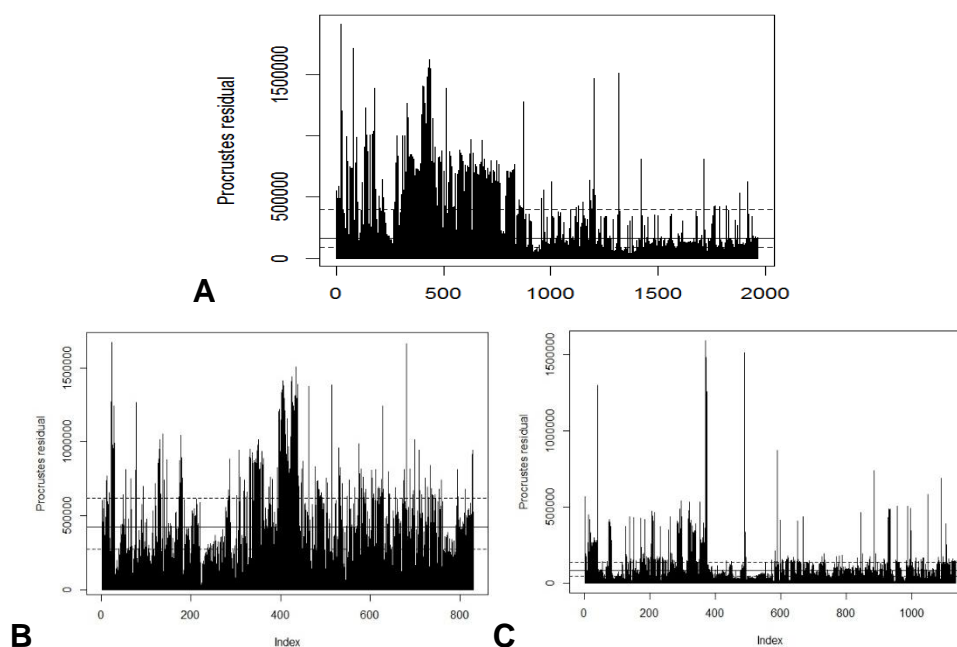


Figura 21. Residuos del análisis procrustes aplicado con la técnica MDS. A, del conjunto de datos. B, residuales de la raza Tuxpeño. C, residuales de la raza Cónico Norteño.

Prieto (2016) utilizó otra forma de presentar el método procrustes que se basa en:

$$\text{Matriz } \mathbf{X} \text{ referencia} = \text{transformación (Matriz } \mathbf{Y}) + \text{residual}$$

En la Figura 21 se gráficán los residuales obtenidos del análisis procrustes en donde se observa que las razas Tuxpeño y Cónico Norteño se ordenan de izquierda a derecha, respectivamente. Es posible deducir que existe mayor nivel de concordancia entre las distancias geográficas y genéticas en la raza Cónico Norteño, que en la raza Tuxpeño, dados los picos del gráfico.

Los residuales (Figura 22) obtenidos según Prieto (2016) indican las diferencias totales entre la matriz de referencia (geográfica) y la matriz a ajustar (genética).

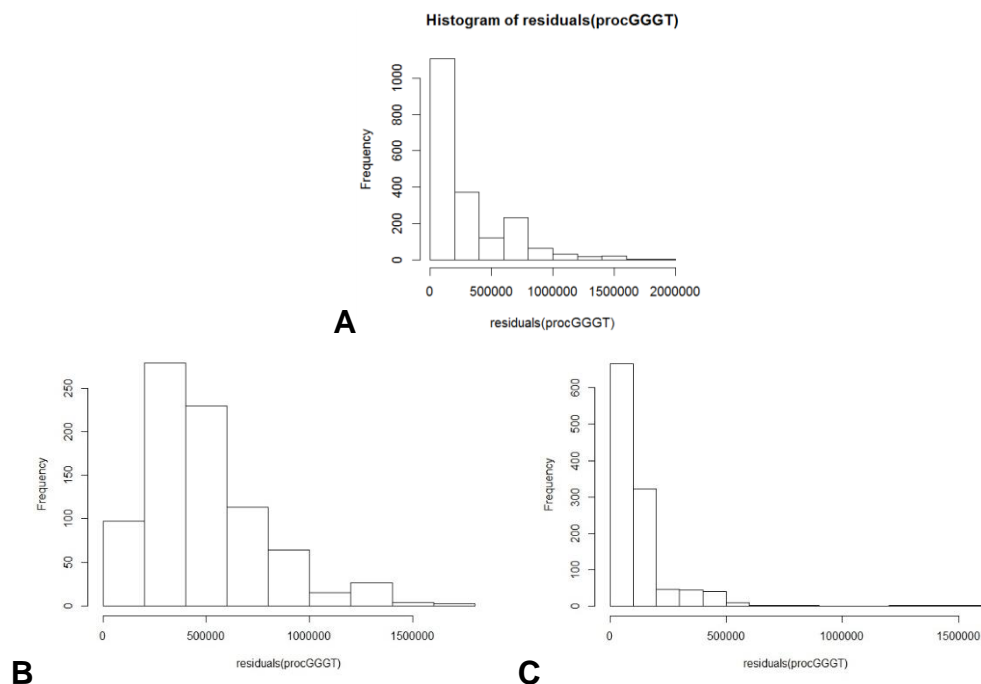


Figura 22. Histograma de residuales del análisis procrustes. A, conjunto de datos. B, raza Tuxpeño. C, raza Cónico Norteño.

Dichos gráficos muestran las diferencias totales entre la matriz de referencia (geográfica) y la matriz a ajustar (genética) para el conjunto de datos (Figura 23A), para la raza Tuxpeño (Figura 23B) y los residuales de la raza Cónico Norteño se muestran en la Figura 23C. Se observa que la variación en la raza Cónico Norteño es menor que en la raza Tuxpeño.

La selección de los métodos y modelos de análisis para un estudio de genética depende primeramente de la pregunta de estudio y del tipo de datos genéticos y espaciales con los que se cuenta. Asimismo, deben considerarse los supuestos y las ventajas y limitaciones de los diferentes métodos y modelos.

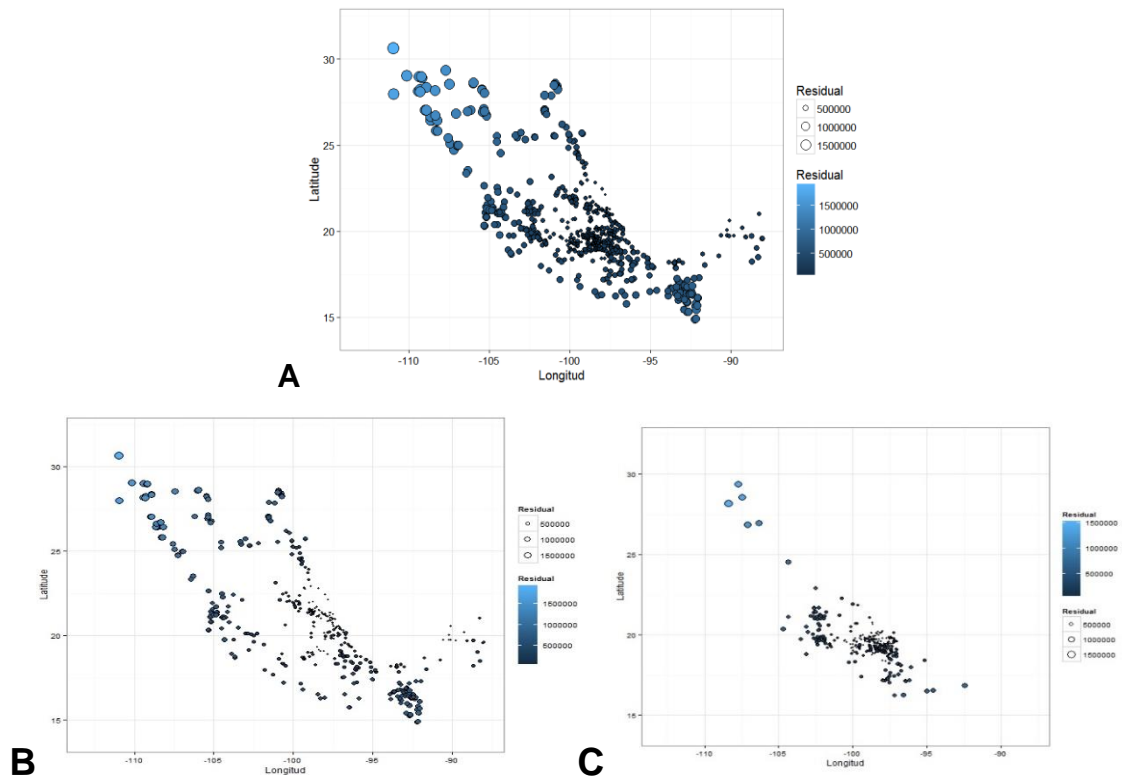


Figura 23. Gráfico de residuales en un mapa de la República Mexicana. A, del conjunto de datos. B, raza Tuxpeño. C, raza Cónico Norteño. Los puntos mas oscuros indican una mayor proximidad con el resto, por lo tanto los de color más claro muestran las relaciones menos cercanas entre la población.

Las diferencias entre la matriz genética y la geográfica (residuales) para la raza Tuxpeño, son muy variadas, a diferencia de la raza Cónico Norteño, donde los valores de residuales son minimos, observándose puntos mas uniformes tanto en tamaño como en color, a excepción de 5 genotipos aislados que se encuentran en la zona superior (norte) de la República Mexicana.

Finalmente, de lo anterior es posible indicar que existe mayor correlacion entre la matriz genética y geográfica para la raza Conico Norteño, que en la raza Tuxpeño.

En las Figuras 24 y 25 se muestra distribución de las distancias geográficas y genéticas promedio, respectivamente, para las accesiones evaluadas.

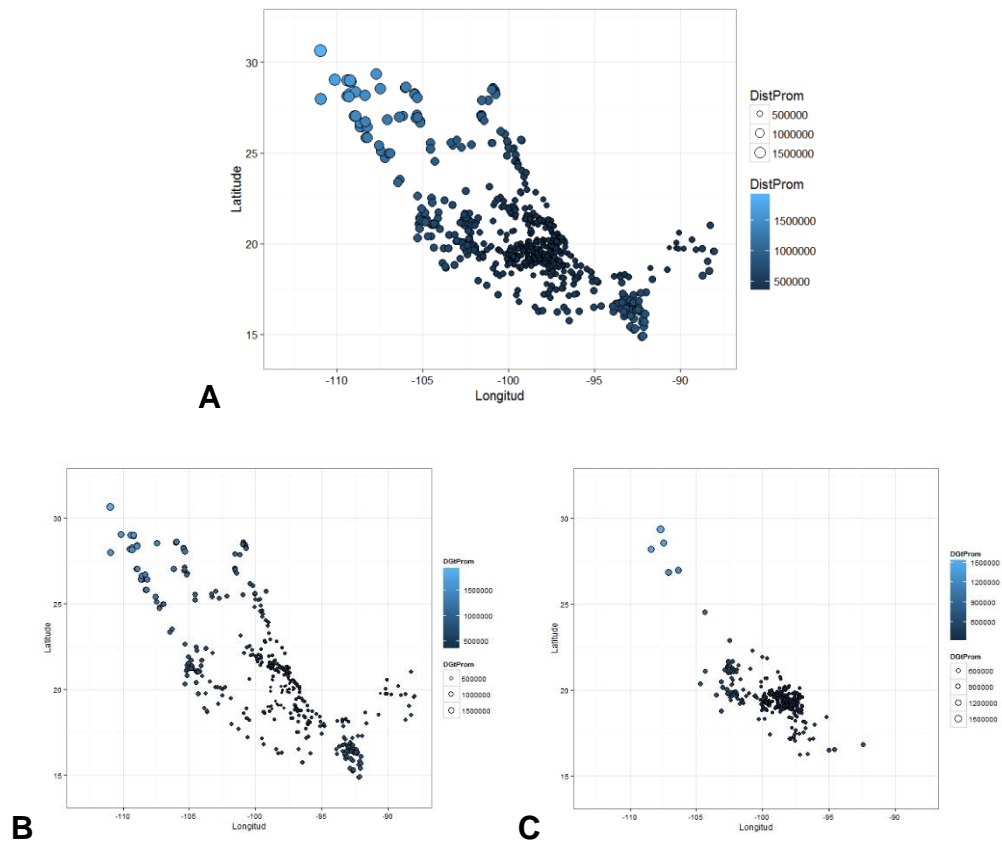


Figura 24. Distancias geográficas promedio para las accesiones de las razas de maíz. A, del conjunto de datos. B, raza Tuxpeño. C, raza Cónico Norteño. Los puntos mas oscuros indican una mayor proximidad con el resto, por lo tanto los de color mas claro muestran las relaciones mas cercanas entre la población.

Sin duda, la correlación de estudios de naturaleza genética y geográfica permite evaluar información genética, ecológica y espacial, a un nivel muy fino y a escalas desde locales hasta regionales. Con ello se han beneficiado planteamientos y evaluaciones, de manera que es posible explicar de una manera mas adecuada los procesos ecológicos y evolutivos que mantienen a las poblaciones a lo largo y ancho del horizonte.

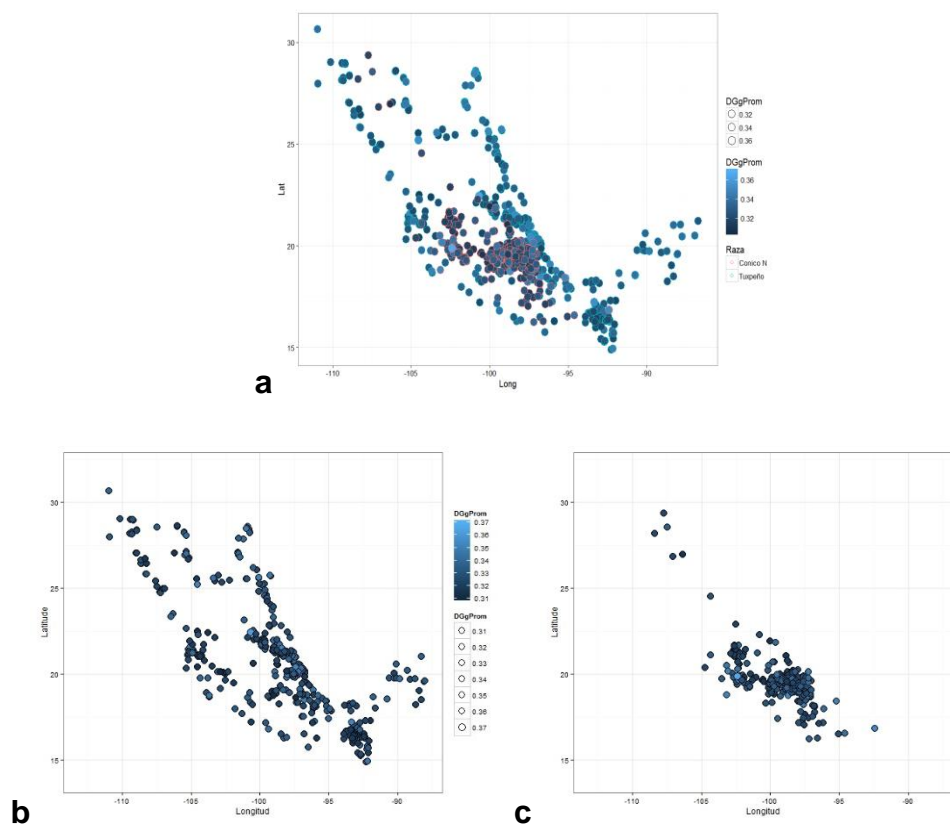


Figura 25. Distribución genética promedio de las accesiones de las razas de maíz. A, del conjunto de datos. B, raza Tuxpeño. C, raza Cónico Norteño. Los puntos más oscuros indican una mayor proximidad con el resto, por lo tanto los de color más claro muestran las relaciones más cercanas entre la población.

En lo referente a las distancias genéticas promedio (Figura 25), la raza Tuxpeño mostró mayor variación (desde 0.31 hasta 0.37), indicando la existencia de mayor diversidad en esta raza, mientras que para Cónico Norteño se observó una menor diversidad genética, con distancias entre 0.32-0.36, indicada como puntos mas uniformes en cuanto a tamaño y color.

5. CONCLUSIONES

El análisis de los genotipos y su clasificación indicó que la raza Cónico Norteño se encuentra más conservada y genéticamente homogénea que la raza Tuxpeño.

Las accesiones de las razas formaron cinco grupos, donde los Tuxpeños se distribuyeron principalmente en los grupos dos , tres y cuatro y los cónico Norteño en el uno y el cinco.

Se detectó la presencia de genotipos intermedios entre las dos razas, lo que sugiere que puede tratarse de productos de hibridación entre las mismas.

El análisis geográfico indicó que la distribución actual de las accesiones en la República Mexicana corresponde con la reportada en 1951 por Wellhausen..

El análisis general de accesiones de las dos razas arrojó un valor de correlación bajo y significativo entre las distancias geográficas y las distancias genéticas.

Las mayores distancias tanto genéticas como geográficas se encontraron para accesiones de la raza Cónico Norteño.

El análisis individual de las razas mostró una mayor correlación entre distancias genéticas y geográficas para la raza Cónico Norteño que para la raza Tuxpeño.

Cónico y Chalqueño provienen de altitudes elevadas, mientras que Tuxpeño se caracteriza por provenir de bajas altitudes. Temporada semiárida, déficit de humedad, precipitación de 409 a 496 mm al año.

Como perspectiva para este trabajo se propone incluir en un estudio genético posterior a la raza Chalqueño debido a que algunos autores reportan que esta raza es el resultado de la hibridación entre Tuxpeño y Cónico Norteño.

6. LITERATURA CITADA

- Aranguren-Méndez, J. A., Román-Bravo, R., Isea, W., Villasmil, Y., & Jordana J. (2005). Los microsatélites (STR's marcadores moleculares de ADN por excelencia para programas de conservación: una revisión. *Arch. Latinoam. Prod. Anim*, 13, 30-42.
- Avendaño, A., Trejo, C. H., López, L. C., Molina, C. J., Santacruz A. V., y Castillo, F. G. (2005). Comparación de la tolerancia a la sequía de cuatro variedades de maíz (*Zea mays* L.) y su relación con la acumulación de prolina. *Interciencia*, 30(9), 560-564.
- Azofeifa-Delgado, Á. (2006). Uso de marcadores moleculares en plantas; aplicaciones en frutales del trópico. *Agronomía mesoamericana*, 17(2), 221-242.
- Badstue, L. B., Bellon, M. R., Berthaud, J., Juárez, X., Rosas, I. M., Solano, A. M., & Ramírez, A. (2006). Examining the Role of Collective Action in an Informal Seed System: a Case Study from the Central Valleys of Oaxaca, Mexico. *Hum. Ecol*, 34, 249-273.
- Balzarini, M. G., Di Rienzo J. A. (2011). InfoGen versión 2011. FCA, Universidad Nacional de Córdoba, Argentina, disponible en: <http://www.info-gen.com.ar>, consultada en noviembre de 2016.
- Balzarini, M. G., Teich I., Bruno, C., Peña, A. (2011). Making genetic biodiversity measurable: a review of statistical multivariate methods to study variability at gene level. *Fac. Cienc. Agrar., Univ. Nac*, 43, 261-275.
- Bellon, M. R., Aguirre, J. A., Smale, M., Berthaud, J. (2002). Intervenciones participativas para la conservación de maíz en fincas en los Valles Centrales de Oaxaca, México. In: Chávez-Servia, J. L., L. M. Arias-Reyes, D.I. Jarvis, J. Tuxill, D. Lope-Alzina y C. Eyzaguirre (eds.). *Resúmenes del Simposio: Manejo de la Diversidad Cultivada en los*

Agroecosistemas Tradicionales. 13-16 de febrero del 2002, (p. 33 y 34). Mérida, México. Instituto Internacional de Recursos Fitogenéticos. Roma, Italia.

- Bedoya C., A. (2012). Estudios de diversidad genética en poblaciones de maíz (*Zea mays* L.), evaluadas con microsatélites. (Tesis Doctoral, Universitat de les Illes Balears, Departamento de Biología), disponible en: <http://www.tdx.cat/handle/10803/116859>, consultada en octubre de 2015
- Biswas, S., Scheinfeldt, L. B., Akey, J. M. (2009). Genome-wide insights into the patterns and determinants of fine-scale population structure in humans. *Am. J. Hum. Genet*, 84, 641–650.
- Brachi, B., Faure, N., Horton, M., Flahauw, E., Vazquez, A. (2010). Linkage and Association Mapping of *Arabidopsis thaliana* Flowering Time in Nature. *PLoS Genet*, 6(5): e1000940. doi: 10.1371/journal.pgen.1000940
- Bramardi, S. J., Bernet G. P., Asíns M. J., Carbonell E. A. (2005). Simultaneous agronomic and molecular characterization of genotypes via the Generalised *Procrustes Analysis*. *Crop Sci*, 45, 1603-1609.
- Casas F., M. (2012). El análisis de escalamiento multidimensional: una alternativa y un complemento a otras técnicas multivariantes. *La sociología en sus escenarios*, 25. 3-5
- Carrari F., Schauer N., Willmitzer L., Fernie, A. (2006). Systems biology: a renaissance of the top-down approach for plant analysis. In: Saito K., Dixon R.A., Willmitzer L. (Eds), *Plant Metabolomics* (pp.185-198). Springer Berlin Heidelberg.
- Cushman, S. A., Shirk, A., Landguth, E. L. (2012). Separating the effect of the habitat area, fragmentation and matrix resistance on genetic differentiation in complex landscapes. *Landscape Ecology*, 27, 369-380.

- CIMMyT (Centro Internacional de Mejoramiento de Maíz y Trigo). (1994). *The CIMMyT Maize Germplasm Bank: Genetic Resource Preservation, Regeneration, Maintenance, and Use*. México, D.F.: Maize Program Special Report. 88 p.
- CIMMyT (Centro Internacional de Mejoramiento de Maíz y Trigo). (2006). *Protocolos de Laboratorio. Laboratorio de Genética Molecular Aplicada del CIMMyT* (3a ed.). El Batán, Texcoco, México. 62 p.
- CIMMyT (Centro Internacional de Mejoramiento de Maíz y Trigo). (2014), *investigación sobre mexican maize landraces*, disponible en: <http://www.CIMMyT.org/es/que-hacemos/investigacion-sobre-maiz/item/mexican-maize-landraces-eroding-but-not-lost>, consultada en octubre de 2015
- CIMMyT (Centro Internacional de Mejoramiento de Maíz y Trigo). (2016), disponible en: <http://www.CIMMyT.org/es/banco-de-germoplasma/>, consultado en octubre de 2016.
- Cockram, J., White, J., Zuluaga, D. L., Smith, D., Comadran, J., Macaulay, M., Luo, Z., Kearsy, M. J., Werner, P., Harrap, D., Tapsell, C., Liu, H., Hedley, P. E., Stein N., Schulte D., Steuernagel, B., Marshall, D. F., Thomas, W. T. B., Ramsay, L., Mackay, I., Balding, D. J., Consortium T. A., Waugh R., O'Sullivan D. M. (2010). Genome-wide association mapping to candidate polymorphism resolution in the unsequenced barley genome. *Proc. Natl. Acad. Sci. USA*, 107, 21611-21616.
- Collard, B. C. Y., Jahufer, M. Z. Z., Brouwer, JB & Pang, E. C. K. (2005). An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concepts. *Euphytica*, 142, 169-196. doi:10.1007/s10681-005-1681-5
- CONABIO (Comisión nacional para el conocimiento y uso de la biodiversidad). (2016). Base de datos del proyecto global "Maices nativos", disponible

en: <http://www.biodiversidad.gob.mx/genes/proyectoMaices.html>, consultado en mayo de 2016.

- Contreras M., V. H. (1990). *Distribución de las zonas áridas y semiáridas en la República Mexicana en base a cinco índices de aridez*. (Tesis profesional, URUZA-UACH, Bermejillo, Durango, México). 51-60
- Corvin, A., Craddock, N., & Sullivan, P. F. (2010). Genome-wide association studies: a primer. *Psychological Medicine*, 40(07), 1063-1077.
- Cuevas S., J. A. (2003). Domesticación inicial del maíz: evaluación experimental de algunos móviles. (Tesis Doctoral, Colegio de Postgraduados, Montecillo, México). 222 p.
- Demarchi. (2009). Microsatélites, distancias genéticas y estructura de poblaciones nativas sudamericanas. *Revista Argentina de Antropología Biológica*, 11(1), 73-88.
- Doyle, J. J. & Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull*, 19, 11-15.
- Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. *PLoS ONE*, 6(5), 10.
- Fernández I., A., (2001). Localizaciones Geográficas. Las Coordenadas Geográficas y la Proyección UTM. (Universal Transversa Mercator), Ingeniería Cartográfica, Geodesia y Fotogrametría, Departamento de Ingeniería Agrícola y Forestal, Universidad de Valladolid, España. 86 p.
- Goodman, M. M. y R. McK. Bird. (1977). The races of maize IV: Tentative grouping of 219 Latin American races. *Economic Botany*, 31, 204-221.
- Gower, J. (1975). Generalized procrustes analysis. *Psychometrika*, 40, 33-51.

- Guillot, G., Leblois, R., Coulon, A., Frantz, A. C. (2009). Statistical methods in spatial genetics Molecular. *Ecology*, 18, 4734-4756.
- Hernández X., E. 1985b. Agricultura tradicional y desarrollo. *Xolocotzia I, Revista*.
- Hernández X., E. y Alanís G., F. (1970). Estudio morfológico de cinco razas de maíz de la Sierra Madre Occidental de México: Implicaciones filogenéticas y fitogeográficas. *Agrociencia*, 5 (1), 3-30.
- Houle, D., Govindaraju, D. R., Omholt, S. (2010). Phenomics: the next challenge. *Nat. Rev. Genet*, 11, 855-866.
- Instituto Geográfico Nacional. (2016), disponible en: <http://www.ign.es/ign/layoutIn/actividadesSistemaInfoGeografica.do>, consultado en octubre de 2016.
- James Hutton Institute. (2016), disponible en: <http://www.hutton.ac.uk/>, consultado en octubre de 2016.
- Jombart, T., Pontier, D., Dufour, A. B. (2009). Genetic markers in the playground of multivariate analysis. *Heredity*, 102, 330-341.
- Jolliffe, I. (2002). Principal component analysis. John Wiley & Sons, Ltd. (487p). ISBN 978-0-387-22440-4
- Kato, Y., T. A, (1984). Chromosome morphology and the origin of maize and its races. *Evol. Biol.*, 17, 219-253.
- Kato, Y., T. A, (2005). Cómo y dónde se originó el maíz. *Investigación y Ciencia*, agosto 2005, 68-72.
- Keitt, T. H., Bjørnstad, O. N., Dixon, P. M., Citron-Pousty, S. (2002). Accounting for spatial pattern when modeling organism-environment interactions. *Ecography*, 25, 616-625.

- LAMP (Proyecto Latinoamericano de Maíz). (1991). ARS-USDA, CIMMyT, Pioneer Hi-Bred International Inc., Universidad Agraria La Molina (Perú).
- Li, J. Z., Absher, D. M., Tang, H., Southwick, A. M., Casto A, M., Ramachandran, S., Cann H. M., Barsh, G. S., Feldman, M., Cavalli-Sforza, L. L., Myers, R. M. (2008). Worldwide human relationships inferred from genome-wide patterns of variation. *Science*, 319,1100-1104.
- Linhart, Y. B, & Grant, M. C. (1996). Evolutionary significance of local genetic differentiation in plants. *Annu. Rev. Ecol. Syst.*, 27, 237-277.
- Littell, R. C., Milliken, G. A., Stroup, W. W., Wolfinger, R. D. (2006). SAS System for Mixed Models. SAS Institute Inc., Cary, NC, USA.
- Luna F, M. (1993). Mejoramiento genético de maíz para condiciones adversas. In: *Simposium internacional IV Nacional, El maíz en la década de los 90. Memorias in extenso*. (pp. 161-170). Zapopan, Jalisco. México.
- Luna F, M., Gutiérrez-Sánchez J. R. (1996) Mejoramiento de maíz y estrategias de producción para temporal de la región semiárida y árida del Centro-Norte. Developing Drought- and Low N-Tolerant Maize. Proceedings of a Symposium, March 25-29, CIMMyT, El Batán, Mexico. Mexico, D.F.: CIMMyT. Disponible en: [Zea mays%Plant breeding%Hybrids%Selection%Agricultural policies%Farming systems%Production factors%Rainfed farming%Semiarid zones%Arid zones%Mexico](#), consultado en octubre de 2016.
- Landguth, E. L., Cushman, S. A. (2010). CDPOP: A spatially explicit cost distance population genetics program. *Molecular Ecology Resources*, 10, 156-161.
- Márquez Sánchez, F. (2008). De las variedades criollas de maíz (*Zea mays*) a los híbridos transgénicos: I. Recolección de germoplasma y variedades mejoradas. *Agricultura, Sociedad y Desarrollo*, 5(2), 151-166.

- Menozzi, P., Piazza, A., Cavalli-Sforza, L. (1978). Synthetic maps of human gene frequencies in Europeans. *Science*, 201, 786-792.
- Mercer K, A., Martínez-Vásquez & Perales R. (2008). Asymmetrical local adaptation of maize landraces along an altitudinal gradient. *Evolutionary Applications*, 1, 489-500.
- Montañez, C., & Warman, A. (1985). Los productores de maíz en México: Restricciones y alternativas. México, D.F: Centro de Ecodesarrollo.
- Muñoz O., A., Pérez, G., López, P. A., Salvador, R. J. (2002). Maíz de cajete: agrosistema y resistencia a sequía. In: Palerm, J (Ed.). *Antología sobre pequeño riego. Vol. III. Sistemas de riego no convencionales*. (pp.137-164). Colegio de Postgraduados, Texcoco, México.
- Ortega P., R. (1985^a). Variedades y razas mexicanas de maíz y su evaluación en cruzamientos con líneas de clima templado como material de partida para mejoramiento. Traducción al español de la tesis de doctorado abreviada. Instituto de Plantas N. I. Vavilov. Leningrado, URSS. 22 p.
- Ortega P., R. (1985b). Recursos genéticos para el mejoramiento de maíz en México, *Revista Germen (Sociedad mexicana de Fitogenética)*, 3, 18-36.
- Ortega P., R. A, Sánchez G., J. J, Castillo G., F, y Hernández C., J. M. (1991). Estado actual de los estudios sobre maíces nativos en México. In: Ortega P., R. A., G. Palomino H., F. Castillo G., V. A, González H. y M. Livera M. (Eds.). *Avances en el Estudio de los Recursos Fitogenéticos de México*. (pp.161-185). SOMEFI. Chapingo, México.
- Paliwal R., L. (2011). El maíz en los trópicos: Mejoramiento y producción. In: *Colección FAO: producción y protección*. Organización de las Naciones Unidas para la Agricultura y la Alimentación. Roma, Italia. 376 p.
- Paredes Gavidia, D. E. (2014). *Recolección y caracterización morfológica y molecular de la colección Nacional del Camote*. (Tesis Doctoral

Dissertation, Universidad de las Fuerzas Armadas ESPE. Carrera de Ingeniería en Biotecnología). 177 p.

Patterson, N., Price, A., Reich, D. (2006). Population structure and eigenanalysis. *PLoS Genet* 2, e190.

Paterniani, E. (2000). Evolución del maíz. In: Fontana, N, H., González, N. c, (Eds). 2000. *El maíz en Venezuela*. (530 p.) Fundación Polar. Caracas.

Perales R., H. R. (2009). Maíz, riqueza de México. *Ciencias*, 92-93, 46-55.

Perales R., H y Duncan G. (2011). Modelos de distribución para las razas de maíz en México y propuesta de centros de diversidad y de provincias bioculturales. Informe técnico preparado para la CONABIO, Octubre. 161 p.

Phillips, W., FRITZ, P. (1995). Cacao y marcadores de ADN en el CATIE. Mejoramiento genético y desarrollo de los cultivos tropicales. Turrialba, CR, CATIE. 25 p.

Piñero, D. (2008). La variabilidad genética de las especies: aspectos conceptuales y sus aplicaciones y perspectivas en México. In: *Capital natural de México, vol. I: Conocimiento actual de la biodiversidad*. (pp. 415-435). CONABIO, México.

Pressoir, G., Berthaud, J. (2004). Population structure and strong divergent selection shape phenotypic diversification in maize landraces. *Heredity*, 92, 95-101.

Rincón S., F., F. Castillo G., M. A. Ruíz T. (2010). Diversidad y distribución de los maíces nativos en Coahuila. México. Sociedad Mexicana de Fitogenética. Chapingo, México. 26-114

- Ruíz C., J., Durán A., N., Sánchez J., J., Ron, J., González D., R., Holland, J. B. (2008). Climatic adaptation and ecological descriptors of 42 Mexican races. *Crop. Sci*, 48, 1502-1512.
- Ruíz C, J. A., Sánchez G., J. J., Hernández J., M., C., Willcox., M. C., Ramírez O., G., Ramírez D., J. L. y González E., D. R. (2013). Identificación de razas mexicanas de maíz adaptadas a condiciones deficientes de humedad mediante datos biogeográficos. *Revista Mexicana de Ciencias Agrícolas*, 4, 829-842.
- SAGA (Servicio de análisis genético para la agricultura). (2016). disponible en: <http://www.masagro.gob.mx/Acciones/Resultados/Descubriendo-la-diversidad-genetica-de-la-Semilla/Paginas/SAGA.aspx>, consultado en octubre de 2016.
- Sánchez, J. J., Goodman, M. M. (1992). Relationships among the Mexican races of maize. *Econ. Bot.*, 46(1), 72-85.
- Soleri, D., & Cleveland, D. A. (2001). Farmer's genetic perception regarding their crop populations: An example with maize in the Central Valleys of Oaxaca, México. *Econ. Bot.*, 55(1), 106-128.
- Schlötterer, C. (2004). The evolution of molecular markers-just a matter of fashion. *Nature Reviews Genetics*, 5, 63-69.
- Serratos J. A., Willcox, M. C, Castillo F. (1997). Flujo genético entre maíz criollo, maíz mejorado y teocintle: implicaciones para el maíz transgénico. CIMMYT, México, DF. 138 p.
- SIAP (Servicio de Información Agroalimentaria y Pesquera). (2016), disponible en: <http://www.siap.gob.mx/>, consultado en octubre de 2016.
- Stretta, E. J. P., Mosiño., P. A. (1963). Distribución de las zonas áridas de la República Mexicana, según un nuevo índice de aridez derivado del de Emberger. *Ingeniería hidráulica (México)*, 16, 40-47.

- Taba, S. (1995a). Maize germplasm: its spread, use and strategies for conservation. In: S. Taba, (Ed), *Maize genetic resources*. (pp.7-58). Mexico, DF, CIMMYT.
- Taba, S. (1995b). Teosinte: geographic variations and conservation. In: S. Taba, (Ed.). *Maize genetic resources*. (pp.59-72). México, DF, CIMMYT.
- Teich I., Planchuelo A. M., Balzarini M. (2011). Análisis de asociaciones en escenarios de datos masivos. In: *40 JAIIO Congreso Argentino de AgroInformática*. (pp.166-177), Argentina.
- Teich, I., Mottura, M., Verga, A., & Balzarini, M. (2015). Asociación entre variabilidad genética y fenotípica con ajuste por autocorrelación espacial en Prosopis. BAG. *Journal of basic and applied genetics*, 26(2), 63-74.
- USDA (Departamento de Agricultura de Estados Unidos). (2016). disponible en: <http://www.usda.gov/wps/portal/usda/usdahome>, consultado en junio de 2016.
- Vekemans X., Hardy O. J. (2004). New insights from fine-scale spatial genetic structure analyses in plant populations. *Mol. Ecol.*, 13, 921-935.
- Wang, C., Szpiech, Z. A., Degnan, J. H., Jakobsson, M., Pemberton, T. J., Hardy, J. A., & Rosenberg, N. A, (2010). Comparing spatial maps of human population-genetic variation using Procrustes analysis. *Statistical applications in genetics and molecular biology*, 9(1), 20
- Weng, J., Xie C, Hao Z., Wang, J., Liu, C., Li, M. (2011). Genome-wide association study identifies candidate genes that affect plant height in Chinese elite maize (*Zea mays* L.) inbred lines. *PLoS One*, 6(12): e29229.
- Wellhausen, E. J., Roberts, L. M., Hernández X., E., en colaboración de Mangelsdorf, P. C. (1951). Razas de maíz en México. Su origen, características y distribución. México D. F. *Oficina de Estudios*

Especiales-Secretaría de Agricultura y Ganadería. Folleto técnico Núm. 55, 237 p.

- Wellhausen, E. J., Roberts, L. M. & Hernández X, E. with Mangelsdorf, P. C. (1952). *Races of Maize in Mexico: Their Origin, Characteristics and Distribution*. Bussey Institution, Harvard University, Cambridge.
- Wen W, Franco J, Chavez-Tovar V. H., Yan J, Taba S. (2012). Genetic Characterization of a Core Set of a Tropical Maize Race Tuxpeño for Further Use in Maize Improvement. *PLoS ONE*, 7, 1-10.
- Zapata A., R. J., y Luna F., M. (1989). *Investigación de maíz en el CIANOC: Resultados y Avances hasta 1985*. Publicación Especial No. 2. CIANOC-INIA-SARH. Calera, Zacatecas. México.
- Zhivotovsky, L. A, Rosenberg, N. A, Feldman, M. W. (2003). Features of evolution and expansion of modern humans, inferred from genome wide microsatellite markers. *Am. J. Hum Genet.*, 72, 1171-1186.

7. APÉNDICES

Apéndice 1. Listado de las 830 accesiones de la raza Tuxpeño evaluadas en este estudio.

El resto puede consultarse en: <https://www.dropbox.com/s/tgztznw5k86w3sv/Tuxpe%C3%B1o.docx?dl=0>

GID	Bank_number	Lat New	Long New	Location_id
33	CIMMYTMA-BANK-000003	19.18	-102.83	9264
41	CIMMYTMA-BANK-000005	19.50	-101.87	9388
52	CIMMYTMA-BANK-000006	18.68	-103.66	2992
243371	CIMMYTMA-BANK-000022	17.97	-101.80	4721
243372	CIMMYTMA-BANK-000023	17.17	-99.54	9354
277	CIMMYTMA-BANK-000036	18.23	-100.52	9453
738	CIMMYTMA-BANK-000088	18.33	-97.35	7999
243416	CIMMYTMA-BANK-000089	18.33	-97.35	7999
243419	CIMMYTMA-BANK-000092	19.83	-98.03	1731
790	CIMMYTMA-BANK-000097	20.03	-97.45	539
243424	CIMMYTMA-BANK-000100	16.28	-98.27	2806
844	CIMMYTMA-BANK-000117	16.33	-98.03	6469
879	CIMMYTMA-BANK-000120	17.99	-96.10	2605
243432	CIMMYTMA-BANK-000121	18.04	-96.11	5362
243461	CIMMYTMA-BANK-000178	17.71	-101.30	3005
2002	CIMMYTMA-BANK-000233	18.33	-97.24	8441
2039	CIMMYTMA-BANK-000234	18.33	-97.24	8441
243507	CIMMYTMA-BANK-000243	18.10	-95.93	2624
2291	CIMMYTMA-BANK-000285	20.61	-100.48	4397
2821	CIMMYTMA-BANK-000300	20.15	-103.05	9413
243558	CIMMYTMA-BANK-000324	19.75	-103.75	1893
3089	CIMMYTMA-BANK-000371	28.98	-109.40	646
243594	CIMMYTMA-BANK-000372	28.98	-109.40	646
3188	CIMMYTMA-BANK-000376	30.64	-110.96	1949
243603	CIMMYTMA-BANK-000388	28.27	-105.47	5446

Apéndice 2. Listado de las 1135 accesiones de la raza Cónico Norteño evaluadas en este estudio.

El resto puede consultarse en: <https://www.dropbox.com/s/nx9hg2ypb4fas6n/Connor.docx?dl=0>

GID	Bank_number	Lat New	Long New	Location_id
96	CIMMYTMA-000017	18.80	-103.14	3011
807	CIMMYTMA-000101	16.58	-94.60	7581
817	CIMMYTMA-000103	17.99	-96.10	2605
5775	CIMMYTMA-000544	20.20	-100.15	336
7711	CIMMYTMA-001375	19.25	-99.62	4256
7756	CIMMYTMA-001404	19.26	-100.03	311
7771	CIMMYTMA-001409	19.73	-99.23	119
7774	CIMMYTMA-001416	19.33	-98.57	9428
7782	CIMMYTMA-001433	19.06	-97.96	8609
7789	CIMMYTMA-001459	19.83	-101.02	283
7791	CIMMYTMA-001460	19.04	-98.78	5096
7793	CIMMYTMA-001462	20.07	-102.73	5026
7794	CIMMYTMA-001464	19.92	-102.00	6768
7795	CIMMYTMA-001467	19.65	-100.59	7892
7877	CIMMYTMA-001555	20.65	-98.65	10050
7965	CIMMYTMA-001665	22.27	-100.76	7863
8486	CIMMYTMA-001961	18.88	-98.68	3868
9336	CIMMYTMA-002336	19.62	-97.08	4923
16985	CIMMYTMA-005572	19.42	-101.53	8005
17033	CIMMYTMA-005645	19.62	-97.08	4923
17602	CIMMYTMA-005888	18.13	-96.82	7828
17613	CIMMYTMA-005898	18.17	-96.91	8058
17614	CIMMYTMA-005899	18.17	-96.91	8058
17615	CIMMYTMA-005900	18.17	-96.95	8616
17754	CIMMYTMA-005915	18.17	-96.95	8616
19632	CIMMYTMA-006039	17.19	-97.87	8027

Apéndice 3. Ejemplo de matriz de distancias calculadas mediante el método modificado de Roger (1983).

Accesión	X653792	X618033	X554127	X554060	X612194	X650336	X692955
X653792	0	0.345747	0.341364	0.354127	0.265507	0.27664	0.377751
X618033	0.345747	0	0.28699	0.285748	0.347757	0.361999	0.312974
X554127	0.341364	0.28699	0	0.295081	0.341326	0.358026	0.330488
X554060	0.354127	0.285748	0.295081	0	0.354499	0.372436	0.328506
X612194	0.265507	0.347757	0.341326	0.354499	0	0.282284	0.379508
X650336	0.27664	0.361999	0.358026	0.372436	0.282284	0	0.391827
X692955	0.377751	0.312974	0.330488	0.328506	0.379508	0.391827	0

Apéndice 4. Ejemplo de matriz de salida del Análisis de Componentes Principales (PCA).

Taxa	PC1	PC2	PC3	PC4	PC5	PC6	PC7
X65379	-	-	0.01870	-0.0047	0.01469	-	0.02049
2	1.06538	0.22162	3		3	0.00961	8
X61803	1.27284	-	0.08201	0.03119	0.00320	0.05067	0.01375
3	3	0.23741	1	8	4	1	9
X55412	1.02028	-	0.01687	-	-	0.00761	0.03367
7		0.07867	7	0.08499	0.03961	6	5
X55406	1.20584	0.06983	0.01995	-	-	-	0.01680
0	3	9		0.01325	0.01895	0.01365	5
X61219	-	-	0.02558	0.01883	-	-	0.00769
4	1.04932	0.14906	6	7	0.00093	0.01179	1
X65033	-	0.17748	0.02189	0.04007	-	-	-
6	1.00102	4	7		0.00423	0.01327	0.01321

Apéndice 5. Ejemplo de matriz Clúster (análisis de asociación).

X653792	1	1	1	1	1	1	1	1
X618033	1	2	2	2	2	2	2	2
X554127	1	2	2	3	3	3	3	3
X554060	1	2	2	2	2	2	2	2

Apéndice 6. Filtrado de la Información rescatada de la base de datos Germinate 3.

GID	Race	Target ID	Jorge's ID	SalD	SNP files (from 0)	PAV files (from 0)
33	Tuxpeño	553259	X553259	SEEDDIV1250	1356	1339
41	Tuxpeño	553283	X553283	SEEDDIV1252	17172	17155
52	Tuxpeño	553295	X553295	SEEDDIV1253	23783	23766
243371	Tuxpeño	553344	X553344	SEEDDIV1265	27163	27146
243372	Tuxpeño	553261	X553261	SEEDDIV1266	1605	1588
277	Tuxpeño	553274	X553274	SEEDDIV1275	20137	20120
738	Tuxpeño	553386	X553386	SEEDDIV1315	13882	13865
243432	Tuxpeño	552853	X552853	SEEDDIV762	32117	32100
1466	Tuxpeño	#N/A	#N/A	#N/A	#N/A	#N/A
243461	Tuxpeño	692984	X692984	SEEDDIV887	26188	26171
2002	Tuxpeño	553395	X553395	SEEDDIV1387	29434	29417

Apéndice 7. Ejemplo de matriz de distancias puras.

Accesión	1	2	3	4	5	6
1	0.00	107.36	103.68	173.23	414.08	265.87
2	107.36	0.00	209.61	170.07	357.35	200.04
3	103.68	209.61	0.00	211.64	467.53	335.40
4	173.23	170.07	211.64	0.00	255.96	138.72
5	414.08	357.35	467.53	255.96	0.00	157.49
6	265.87	200.04	335.40	138.72	157.49	0.00

Apéndice 8. Ejemplo de matriz de identificación de distancias calculadas reales.

FI	Shape	N	Target_ID	countries			OI
D	*	r	D	_	Lat_New	Long_New	D
					19.18333	-	55325
0	Point	1	553259	MEXICO	3	102.833333	0 9
						-	55328
1	Point	2	553283	MEXICO	19.5	101.866667	1 3
						-	55329
2	Point	3	553295	MEXICO	18.6781	103.661841	2 5
					17.97179	-	55334
3	Point	4	553344	MEXICO	8	101.799996	3 4
					17.16736		55326
4	Point	5	553261	MEXICO	3	-99.537659	4 1
					18.23333	-	55327
5	Point	6	553274	MEXICO	3	100.516667	5 4
					18.33462		55338
6	Point	7	553386	MEXICO	4	-97.353287	6 6
					18.33462		70585
7	Point	8	705853	MEXICO	4	-97.353287	7 3

Apéndice 9. Ejemplo de matriz de distancias calculadas reales, determinando la longitud en metros y kilómetros, respectivamente.

OID_OBSERV	OID_TARGET	longitude (m)	(km)
0	0	0	0
0	1	114085.34908400000	114.0853491
0	2	110182.73461300000	110.1827346
0	3	183419.02360000000	183.4190236
0	4	439352.55697000000	439.352557
0	5	281551.93779000000	281.5519378
0	6	619630.69693400000	619.6306969
0	7	619630.69693400000	619.6306969

Apéndice 10. Ejemplo de archivo Excel para transformar en un archivo de extensión kml.

Latitude	Longitude	Name	Description	Icon
19.183333	-102.833333	1	553259	1
19.5	-101.866667	2	553283	1
18.6781	-103.661841	3	553295	1
17.971798	-101.799996	4	553344	1
17.167363	-99.537659	5	553261	1