



UNIVERSIDAD AUTÓNOMA CHAPINGO

DEPARTAMENTO DE FITOTECNIA

INSTITUTO DE HORTICULTURA



**DIVERSIDAD GENÉTICA EN COLECCIONES NATIVAS
DE MAÍZ EMPLEANDO SNP's**

TESIS

**QUE COMO REQUISITO PARCIAL PARA OBTENER EL GRADO
DE:**

**MAESTRO EN CIENCIAS EN BIOTECNOLOGÍA
AGRÍCOLA**



DIRECCION GENERAL ACADEMICA
DEPTO. DE SERVICIOS ESCOLARES
OFICINA DE EXAMENES PROFESIONALES

**PRESENTA:
GUSTAVO GERMAN VILCHIS**

**DIRECTOR DE TESIS: DR. JUAN PORFIRIO LEGARIA
SOLANO**

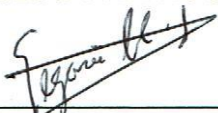
Chapingo, Estado de México, junio de 2016

La presente tesis titulada; Diversidad genética en colecciones nativas de maíz empleando SNP's, fue realizada por el C. Gustavo German Vilchis, bajo la dirección del consejo indicado, ha sido aprobada por el mismo y aceptada como requisito parcial para obtener el grado de:

**MAESTRO EN CIENCIAS
EN BIOTECNOLOGÍA AGRÍCOLA**

CONSEJO PARTICULAR

DIRECTOR:



DR. JUAN PORFIRIO LEGARIA SOLANO

ASESOR:



DRA. SARAH HEARNE

ASESOR:



DR. TERENCE LUKE MOLNAR

ASESOR:



DR. JOSÉ DE JESUS LÓPEZ REYNOSO

ASESOR:



DR. JAIME SAHAGÚN CASTELLANOS

Chapingo, Estado de México. Junio de 2016

ÍNDICE GENERAL

Página.

ÍNDICE DE CUADROS.....	v
ÍNDICE DE FIGURAS.....	vi
ABREVIATURAS USADAS.....	vii
DEDICATORIAS.....	viii
AGRADECIMIENTOS.....	ix
DATOS BIOGRÁFICOS.....	x
RESUMEN.....	xi
ABSTRACT.....	xi
I. INTRODUCCIÓN.....	12
1.1. OBJETIVOS DEL ESTUDIO	14
1.2 HIPÓTESIS.....	14
II. REVISIÓN DE LITERATURA.....	15
III. MATERIALES Y MÉTODOS.....	30
3.1 .Fuente de germoplasma.....	30
3.2 .Caracterización molecular.....	30
3.3 .Control de calidad y análisis de los datos.....	31
3.4 Tendencia inicial de selección dentro de las colecciones de estudio.....	32
IV. RESULTADOS Y DISCUSIÓN.....	34
4.1 Control de calidad y selección de los marcadores.....	34
4.2 Parámetros de diversidad genética.....	34
4.3 Análisis de conglomerados.....	37
4.4 Análisis Escalado Multi Dimensional (MDS).....	40
4.5 Análisis de Componentes Principales (PCA).....	44
4.6 Colecciones nativas con potencial de resistencia a sequía...46	

V.	CONCLUSIONES.....	49
VI.	BIBLIOGRAFÍA.....	51
VII.	APÊNDICE.....	58

ÍNDICE DE CUADROS

	Página
Cuadro 1. Número de marcadores SNP's y su filtrado.....	34
Cuadro 2. Parámetros de diversidad genética para marcadores tipo SNP	36

ÍNDICE DE FIGURAS

	Página
Figura 1. Dendrograma de relaciones entre 290 colecciones nativas de maíz, obtenido a partir de marcadores tipo SNP.....	38
Figura 2. Gráfico tridimensional del análisis MDS por categoría de razas.....	41
Figura 3. Gráfico tridimensional del análisis MDS por categoría de grupos.....	42
Figura 4. Gráfico tridimensional del análisis MDS por categoría de país de origen.....	43
Figura 5. Figura 5. Grafico tridimensional del análisis de componentes principales para los primeros tres componentes.....	45
Figura 6. Gráfico tridimensional de las 20 colecciones nativas seleccionadas por su tolerancia a sequía y su relación en términos de distancias genéticas.....	48

ABREVIATURAS USADAS

AFLPs: Polimorfismos en la longitud de fragmentos amplificados

CIMMyT: Centro Internacional de Mejoramiento de Maíz y Trigo

ddNTPs: di-Deoxinucleótidos

FAO: Organización para la Alimentación y la Agricultura

MAF: Frecuencia del alelo menor

PCR: Reacción en cadena de la polimerasa

PIB: Producto interno bruto

RAPDs: Amplificación aleatoria de ADN polimórfico

RFLPs: Polimorfismos de longitud de fragmentos de restricción

SAGA: Servicio de Análisis Genético para la Agricultura

SIAP: Servicio de Información Agroalimentaria y Pesquera

SIG o GIS: Sistemas de Información Geográfica

SNP: Polimorfismo de nucleótido simple

DEDICATORIAS

A Dios por la vida, la magnificencia de lo creado y por las oportunidades pasadas, presentes y futuras.

A mi esposa Sara y a mi hija Ivka, por su gran apoyo y saber qué de muchas formas creen en mí, su amor incondicional siempre significará un motivo de lucha, este logro es parte de su amorosa contribución.

A mis padres:

Heladio German Enríquez

Mi gran mentor, honorable amigo y buen consejero. Por tu apoyo y fe incondicional.

María de los Ángeles Vilchis Álvarez

Mujer de espíritu incansable, madre consejera y perseverante. Tu sacrificio constante es mi aliento.

A mis hermanos, sobrinos y cuñados; quienes en momentos difíciles me tienden la mano en diversas formas, quiero decirles que esto es suyo también.

AGRADECIMIENTOS

A CONACyT (Consejo Nacional de Ciencia y Tecnología) por otorgarme el financiamiento para cursar mis estudios de maestría.

A la Universidad Autónoma Chapingo que me brindó la oportunidad de ser un estudiante dentro de sus filas y donde pude vivir experiencias en favor de mi desarrollo integral como persona y como mexicano.

Al Centro Internacional de Mejoramiento de Maíz y Trigo –CIMMyT-, de manera especial a MasAgro-Biodiversidad, por el estipendio económico, por los datos aportados y por la asesoría por parte de los investigadores que ahí laboran.

A mi comité asesor: Dr. Juan Porfirio Legaria Solano, Dra. Sarah J. Hearne, Dr. Terence Luke Molnar, Dr. José de Jesús López Reynoso y al Dr. Jaime Sahagún Castellanos, por su confianza, disposición, y compromiso con el proyecto que hoy culmina.

A los maestros del Posgrado en Biotecnología Agrícola, quienes siempre con buena actitud compartieron sus conocimientos para mi formación científica y profesional.

A todos los compañeros e investigadores de CIMMyT y SAGA, que además de cumplir con su noble trabajo en campo o en laboratorio, siempre tuvieron una excelente actitud de colaboración hacia este proyecto. Expreso para ustedes, toda mi amistad y respeto.

DATOS BIOGRÁFICOS

Datos personales

Nombres: Gustavo German Vilchis

Fecha de nacimiento: 22 de agosto de 1984

Lugar de nacimiento: Toluca, Estado de México.

CURP: GEVG840822HMCRLS09

Profesión: Ingeniero Agrónomo Fitotecnista

Cédula profesional: 7572350



Desarrollo académico

Maestría: 2014-2016 Universidad Autónoma Chapingo

Especialidad: 2010-2011 Universidad Autónoma del Estado de México

Licenciatura: 2003-2008 Universidad Autónoma del Estado de México

Bachillerato: 1999-2003 Universidad Autónoma del Estado de México

Experiencia Profesional

Cargo de Superintendente en la empresa denominada: Rancho La Laguna de Jilotepec S.A. de C.V. dedicada a la producción de cereales (maíz, trigo, avena y triticale) y de pollo de engorda a través de tres diferentes fincas productivas en los estados de: San Luis Potosí (Villa de Reyes), Guanajuato (San José Iturbide) y Estado de México (Jilotepec), como responsable de la administración, planeación, asesoría técnica, producción, venta, distribución y cobranza de los diversos productos.

Colaboración en el Instituto Nacional de Investigaciones Nucleares, en la investigación de los recursos naturales y genéticos, a través de programas de exploración, colecta, caracterización, selección y mejoramiento genético mediante mutagénesis radio inducida.

Experiencia en el establecimiento de parcelas experimentales y demostrativas para la evaluación en campo de diversas especies.

DIVERSIDAD GENÉTICA EN COLECCIONES NATIVAS DE MAÍZ EMPLEANDO SNP's

GENETIC DIVERSITY IN MAIZE LANDRACES USING SNP

Gustavo German Vilchis¹ y Juan Porfirio Legaria Solano²

RESUMEN

El maíz (*Zea mays* L.) es una materia prima importante para los humanos y en la producción animal, como fuente de: alimentación, obtención de harina, aceite, proteína y alcohol. La sequía es el factor ambiental que afecta más seriamente la producción de maíz. Las poblaciones nativas poseen características que las hacen muy interesantes para su uso en el mejoramiento. La identificación de germoplasma útil para fines agronómicos es un tema decisivo en mejoramiento genético. El objetivo de este estudio fue estimar la variabilidad genética en 290 colecciones nativas de maíz mediante marcadores SNP, a partir de calcular sus distancias genéticas y sus patrones de agrupamiento. Las accesiones evaluadas pertenecen a una colección núcleo del CIMMYT y que fueron seleccionadas a partir de los Sistemas de Información Geográfica (SIG). La caracterización molecular se hizo a partir de una mezcla compuesta de ADN de 30 plantas sanas al azar, a partir del método CTAB. Los marcadores se obtuvieron a partir de la técnica de genotipificación por secuenciación (GbS). Los resultados de los parámetros genéticos como la heterocigosis, fueron altos, lo que sugiere la existencia de gran variabilidad genética. Las accesiones se agruparon en ocho grupos diferentes, por sus relaciones genéticas, raciales y geográficas. Este estudio aporta información que puede promover un mejor uso y aprovechamiento de la diversidad genética existente que permita ampliar la base genética de la resistencia a sequía.

Palabras clave: SIG (Sistema de Información Geográfica), Genotipificación por secuenciación (GbS), Resistencia a la sequía.

ABSTRACT

Maize (*Zea mays* L.) is an important raw material for humans and in animal production as a source of: food, obtaining flour, oil, protein and alcohol. Drought is the environmental factor that most seriously affects corn production. The landraces have characteristics that make them very interesting for use in breeding programs. Identifying useful germplasm for agricultural purposes is a critical issue in breeding. The objective of this study was to estimate the genetic variability in 290 maize landraces using SNP markers, from calculating their genetic distances and clustering patterns. The accessions belong to CIMMYT's core collection and were selected from Geographic Information Systems (GIS). Molecular characterization was made from a mixture composed (bulk) of DNA from 30 plants, extracted by the CTAB method. The markers were obtained from genotyping by sequencing (GbS) technique. The results of genetic parameters such as heterozygosity were high, suggesting high genetic variability. They were grouped into eight different groups, due their genetic, racial and geographic relationships. This study provides information that can promote a better use and exploitation of the existing genetic diversity that allows expanding the genetic basis of drought resistance.

Keywords: GIS (Geographic Information System), Genotyping by sequencing (GbS), Drought resistance.

¹Estudiante tesista de la Maestría en Ciencias en Biotecnología Agrícola.

²Director de tesis.

I. INTRODUCCIÓN

El maíz (*Zea mays* L.) es una materia prima importante para los humanos y en la producción animal, como fuente de: alimentación, obtención de harina, aceite, proteína y alcohol. La sequía es el factor ambiental que afecta más seriamente la producción de maíz y reduce en gran medida el rendimiento (Chong y Shao, 2010).

La heterogeneidad ambiental, la recombinación entre razas de maíz y la selección por parte de los agricultores contribuyeron al incremento de su diversidad genética (Eagles y Lothrop, 1994).

Uno de los efectos pronosticados del calentamiento global es el incremento de las condiciones de sequía (Cook *et al.*, 2007). Por lo tanto, la tolerancia del maíz ante la sequía resulta una necesidad urgente para mantener la producción mundial (Bruce *et al.*, 2002). Sin embargo, es claro que los mecanismos genéticos de tolerancia al estrés hídrico son muy complejos, de tal forma, los programas de mejoramiento para tolerancia a sequía requieren información respecto al genotipo, el ambiente y la interacción genotipo-ambiente (G X E) (Yue *et al.*, 2005).

La tipificación de genes funcionales o marcadores estrechamente vinculados a tolerancia a sequía, es un paso clave en el mejoramiento genético asistido (Xue *et al.*, 2013).

Las poblaciones nativas de maíz poseen características que las hacen muy interesantes para su uso en el mejoramiento. Sin embargo, el uso directo de variedades locales presenta ciertas desventajas: no existe suficiente documentación con respecto a su comportamiento, baja adaptación local, falta de sincronía en los tiempos de floración y otros (Salhuana y Pollak, 2006). Por lo

tanto, es un reto identificar alelos favorables para estreses abióticos en fondos genéticos no seleccionados para mejoramiento. Contrariamente, en los últimos años se ha determinado que a mayores niveles de productividad se detectan altos niveles de vulnerabilidad genética a cambios ambientales, como sequía o estreses abióticos como enfermedades dentro de poblaciones “elite” a pesar de los esfuerzos en los programas de mejoramiento. Este fenómeno amenaza la meta de aumentos sostenidos en el potencial de rendimiento debido a la estrecha base genética en los cultivares comerciales.

La identificación de germoplasma útil para fines agronómicos es un tema decisivo en el pre-mejoramiento, a respecto existen varias opiniones acerca de la abundancia y disponibilidad de variabilidad genética que se pueda explotar provechosamente, con la idea de que no toda la variabilidad genética es útil, o que a veces hay que “pagar” un precio al incorporar alelos exóticos, ya que en ocasiones se pueden acarrear algunas características indeseables. Usualmente se ha identificado germoplasma exótico útil mediante la evaluación fenotípica dentro de cada programa de mejoramiento o en múltiples ambientes en el caso de las colecciones internacionales (Pritsch, 2001).

1.1. OBJETIVOS

Conocer la variabilidad genética existente en 290 colecciones de maíz, provenientes del Banco de Germoplasma del CIMMyT; de manera que la información generada pueda en un futuro, contribuir para ampliar la base genética de la resistencia a sequía.

Valorar la diversidad genética a partir de las distancias genéticas entre las colecciones, calculadas a partir de marcadores tipo SNP's y sus patrones de agrupamiento.

Presentar la relación entre las distancias genéticas calculadas y la tendencia inicial de selección fenotípica dentro del programa de mejoramiento genético de MasAgro-Biodiversidad, a partir de los genotipos evaluados en este estudio.

1.2. HIPÓTESIS

El estudio de la variabilidad genética de las poblaciones nativas de maíz de la Colección Núcleo de Maíz del Banco de Germoplasma del CIMMyT puede ser una alternativa importante para la identificación de genotipos resistentes a la sequía.

El análisis molecular de las diversas accesiones de maíz permitirá caracterizarlas y relacionarlas con base en sus diferencias y similitudes genómicas, así como proveer información básica que sugiera la posible ubicación taxonómica de los materiales estudiados.

II. REVISIÓN DE LITERATURA

2.1. Agricultura e importancia del maíz

Actualmente en diversos países del mundo, la agricultura ha dejado de ser vista únicamente como una actividad productiva, para convertirse en una actividad que cumple diversas funciones en la sociedad. De acuerdo con la Organización de las Naciones Unidas para la Agricultura y la Alimentación (FAO, 1999), el carácter multifuncional de la agricultura puede resumirse de la siguiente manera:

Contribución a la seguridad alimentaria. Existe seguridad alimentaria cuando todas las personas tienen en todo momento acceso físico y económico a suficientes alimentos inocuos y nutritivos para satisfacer sus necesidades alimenticias y sus preferencias en cuanto a los alimentos para llevar una vida activa y sana.

Función ambiental. Que comprende el incremento de los efectos positivos y la mitigación de los negativos.

Función económica. La producción primaria de alimentos y otros bienes, así como de productos y servicios relacionados con la capacidad agrícola/empresarial, actividades múltiples con efectos económicos más amplios y efectos directos e inducidos en los sistemas económicos.

Función social. Comprende la viabilidad de las comunidades rurales y los medios de vida, la cultura y los valores culturales.

Bonnal *et al.* (2003) plantean que desde la perspectiva de la multifuncionalidad, la agricultura no sólo contribuye al desarrollo socioeconómico de los países, sino se convierte en un elemento indispensable de la seguridad nacional y del mantenimiento de la paz social de la sociedad rural, cuyo valor es imposible cuantificar y contabilizar en el producto interno bruto (PIB). Que el Estado

mexicano reconozca el carácter multidimensional de la agricultura es de suma importancia debido a que las políticas agrícolas implementadas deben ir más allá de los aspectos productivos y de precios.

Las plantas domesticadas han sido un factor importante en la economía de todas las culturas del mundo; han sido usadas como alimento, bebida, medicina, como materiales en la industria y en gran medida en aspectos religiosos y como estimulantes. El maíz tuvo un papel central en el origen y la difusión de la agricultura en todas las civilizaciones indígenas de Mesoamérica. Las civilizaciones de Mesoamérica dependieron en gran medida del cultivo de maíz; la gran variedad de metates, comales, los instrumentos de cultivo y la cerámica han demostrado su importancia. Adicionalmente, las ceremonias religiosas, festivales seculares, historias y leyendas sobre el origen del maíz y los grupos humanos, y los cientos de usos del grano y otras partes de la planta ilustran su importancia en las culturas antiguas y actuales de México y del resto de Latinoamérica (Sánchez, 2011).

El maíz se cultiva en el mundo con fines distintos; sin embargo, en México es casi exclusivamente para la alimentación humana. La producción mundial de maíz fue de 818 millones de toneladas de grano en 2009; México fue el cuarto país productor, superado por Estados Unidos (40 % de la producción mundial), China y Brasil (FAO, 2009; SIAP, 2011).

2.2. Origen y diversificación del maíz

La domesticación es un proceso evolutivo continuo, mediado por la selección humana, que en las plantas genera fenotipos favorables para el consumo humano y la agricultura pero disminuye o elimina la capacidad de éstas para sobrevivir en condiciones naturales, volviéndolas dependientes del hombre

(Harlan, 1992). La pérdida de la diversidad genética del cultivo en comparación a sus parientes silvestres, es una consecuencia de este proceso, es el resultado de la selección artificial y del cuello de botella genético (Tenaillon *et al.*, 2004).

Las investigaciones sobre el origen del maíz en el tiempo señalan que apareció entre los años 9000 y 5000 A. C., en la región Mesoamericana, y fue el principal cereal utilizado por los Mayas (Turrent y Serratos 2004).

El maíz se cultiva desde las costas de casi todo el continente Americano hasta las tierras altas de alrededor de 4,000 msnm en los Andes; se siembra en las regiones secas con una precipitación media inferior a 400 mm y en regiones con precipitación superior a 4,000 mm. Hay una gran variedad de tipos de maíz: variedades con altura de planta de uno a cinco metros, con distintos grados de tolerancia a la sequía, al calor o a las heladas, con adaptación a las diferentes texturas de suelo, altitud y latitud. Todas las partes de la planta de maíz se utilizan: el grano en la alimentación humana (tortillas, tostadas, atole, tamales, ponteduro, totopos, pinole, pozole, arepas, bollos, chicha, mote, confite, cancha), los tallos para el jugo azucarado, tallos secos para cercas y como combustible, los olotes y las raíces como combustible, los hongos de la mazorca del maíz en alimentación (*Ustilago maydis*), las hojas del tallo y de la mazorca para envolver tamales, entre los usos más comunes. Es importante destacar que el maíz en la actualidad se consume en todo el mundo de diversas formas, como verdura, como elote, el grano seco en diferentes modalidades; sin embargo, en los países desarrollados el maíz es un componente importante de muchos alimentos, bebidas y productos industriales. Se ha calculado que en los supermercados modernos, cerca de 2,500 productos contienen maíz en alguna forma (Perales, 2009).

2.3. Recursos genéticos

En la actualidad es común escuchar que una de las mayores riquezas de un país radica en sus recursos genéticos, de tal forma que países con gran diversidad de dichos recursos están emprendiendo acciones normativas, legales y de investigación a fin de protegerlos. El germoplasma determina el desarrollo de cada ser viviente y no hay recurso que deba ser mejor valorado (Krishnamurthy y Sahagún, 1991).

2.3.1 Definición de recursos genéticos

De acuerdo con Krishnamurthy y Sahagún (1991), Rodríguez (2003) y La Comisión Intersecretarial de Bioseguridad de los Organismos Genéticamente Modificados (2002), el germoplasma es el material que controla la herencia biológica: la suma de las cualidades y potencial que se derivan genéticamente de sus ancestros. Y tienen un valor real o una utilidad potencial (alimentación, medicina, forraje, principios químicos, ornamentales, madereros, artesanal, entre otros).

2.3.2. Categorías de los recursos fitogenéticos

Los recursos fitogenéticos comprenden la diversidad genética del mundo vegetal, misma que constituye un patrimonio de la humanidad de valor incalculable del presente y futuro (Cubillos, 1997).

Krishnamurthy y Sahagún (1991) establecen las categorías siguientes para los recursos fitogenéticos:

Parientes silvestres. Son el producto de la evolución natural y pueden proveer genes importantes a los cultivos. Debido a que estas formas vegetales han permanecido expuestas a la naturaleza, usualmente son portadoras de genes que les permiten afrontar exitosamente los riesgos que la naturaleza les plantea, tales como temperaturas altas, heladas, sequías, suelos salinos, plagas y enfermedades.

Parientes semi-silvestres (weedy relatives). Son plantas que se desarrollan en ambientes modificados por el hombre y sirven de puentes botánicos para la transferencia de genes importantes entre los parientes silvestres y los cultivos.

Cultivares primitivos o tradicionales. Estos han evolucionado durante largos periodos, tanto por el efecto de la selección natural, así como por la que practica el hombre. Estos cultivos han acompañado al hombre en su constante peregrinar sobre el mundo. Este proceso, que se llevó a cabo en diferentes lugares, produjo una tremenda variabilidad genética, misma que ahora puede ser fácilmente utilizada por el fitomejorador.

Cultivares modernos. Son los productos de los métodos modernos de mejoramiento. Las características deseables de estos cultivares, a su vez, pueden servir como recursos genéticos potencialmente útiles en el mejoramiento genético que se haga de los cultivos del futuro.

Material genético élite. Es el germoplasma que ha sido mejorado por los fitomejoradores, y será la base para la creación de los cultivares que serán utilizados por los agricultores.

2.3.3. Utilización de los recursos genéticos de maíz

Las poblaciones nativas de maíz poseen características que las hacen muy interesantes para su uso en el mejoramiento. Entre ellas se pueden citar: el aumento de la variabilidad genética que trae ventajas en la selección tanto para

caracteres de interés agronómico, caracteres de calidad como contenido nutritivo, sabor , textura, así como caracteres de resistencia a enfermedades y plagas. Sin embargo, el uso directo de variedades locales en el mejoramiento presenta ciertas desventajas: no existe suficiente documentación con respecto a su comportamiento con diversos híbridos, probadores o accesiones de control, ya que poseen baja adaptación local, falta de sincronía en los tiempos de floración y características agronómicas indeseables (Salhuana y Pollak, 2006). Por lo tanto, es un reto identificar alelos favorables para parámetros agronómicos como estreses abióticos en fondos genéticos no seleccionados para mejoramiento.

Contrariamente, en los últimos años se ha determinado que a mayores niveles de productividad se detectan altos niveles de vulnerabilidad genética a cambios ambientales, como sequía o estreses bióticos como enfermedades dentro de poblaciones “elite” a pesar de los esfuerzos en los programas de mejoramiento. Este fenómeno amenaza la meta de aumentos sostenidos en el potencial de rendimiento debido a la estrecha base genética en los cultivares comerciales (Pritsch, 2001).

Tomando en cuenta que la estabilidad y productividad de los cultivos agrícolas como el maíz está en parte sujeta a la eficiencia de los métodos para ampliar la base genética del germoplasma utilizado, se ha establecido entonces una etapa previa al mejoramiento: el pre-mejoramiento (pre-breeding) o valoración genética de germoplasma con el fin de posibilitar la transferencia de la variabilidad genética útil al material adaptado y mantener las combinaciones alélicas presentes en el material elite (Duvick, 1990).

2.4. Diversidad genética del maíz

El conocimiento de la diversidad morfológica y genética, su relación con el ambiente y el impacto social determinan la estrategia de conservación y aprovechamiento de la diversidad genética del maíz (Ortega *et al.*, 1991; Castillo, 1993). La riqueza genética contenida en las poblaciones de maíz representa posibilidades inmensas para la selección y recombinación de los recursos genéticos de maíz en diferentes regiones del país (Ortega, 1985), para la solución de diversos problemas del cultivo.

Con niveles de diversidad genética de los más altos del mundo, México es considerado el centro de origen, domesticación y diversidad del maíz, consideración debida principalmente a factores tanto regionales como culturales y sociales en el país (Brush, 1998). En México se tienen documentadas 59 razas, ordenadas en cuatro grupos y algunos subgrupos, según estudios de caracteres morfológicos, iso-enzimáticos y climáticos (Sánchez y Goodman 1992; Sánchez *et al.*, 2000).

Según Sánchez (2011), la diversidad genética es la causa por la que una determinada especie es capaz de evolucionar en diferentes ambientes y bajo presión de selección; así mismo, el conocimiento de tal diversidad, es indispensable para la diversificación de la fuentes de germoplasma, y al mismo tiempo se incrementan las probabilidades de detectar alelos favorables.

En genética, tanto desde el punto de vista del mejoramiento como de la conservación, el principal objetivo es estudiar, determinar y medir la variación existente entre y dentro de poblaciones o accesiones. Estos estudios de diversidad hacen posible la estimación de esta variabilidad genética a varios niveles de análisis más complejos, incluyendo la estimación de la distancia genética entre poblaciones, y pueden extenderse a unidades grandes como

zonas y regiones.

Los análisis de diversidad genética en colecciones de germoplasma pueden facilitar una confiable clasificación de accesiones así como el establecimiento de colecciones núcleo (Mohammadi y Prasanna, 2003; Bedoya *et al.*, 2008). Específicamente, una colección núcleo es un subconjunto de accesiones que representa con un mínimo de repetitividad la diversidad genética de una especie vegetal o de una colección de germoplasma (Frankel y Brown, 1984). De esta manera, con el establecimiento de colecciones núcleo se puede obtener información representativa que permite una utilización más rápida y eficiente de una colección, se logran también mejorar estrategias de manejo y distinguir prioridades de colección o búsqueda de nuevos materiales (Bedoya *et al.*, 2008).

La caracterización morfológica es una metodología muy empleada para conocer la diversidad genética del maíz de forma rápida, sencilla y económica; sin embargo, las condiciones ambientales influyen en la expresión de algunos caracteres morfológicos. Para identificar los caracteres más adecuados que representen la diversidad del maíz, se han estudiado cerca de 50 características de planta, mazorca, grano y olote (Sánchez *et al.*, 1993).

Por otro lado entre los parámetros que permiten evaluar la diversidad están: el número total de alelos en la población, el número medio de alelos por locus, el número medio de alelos por locus polimórfico, la heterocigosidad, los estadísticos F, la distancia genética y la identidad genética. Una medida de la diversidad genética individual son los niveles totales de heterocigosidad, que es definida como la probabilidad de que dos individuos parentales tomados al azar de una población tengan diferentes alelos en ese locus, que pueden ser idénticos para la heterocigosidad esperada bajo el equilibrio Hardy- Weinberg. Cuando se evalúan muchos *loci*, se estima la heterocigosidad promedio sobre el grupo de *loci* muestreados. Este parámetro está desconectado de las frecuencias

genotípicas actuales, así que se entiende como la medición del grado de variación alélica, más que una medición de la frecuencia de heterocigotes, excepto en una población en equilibrio de Hardy- Weinberg (Meffe y Carroll, 1994).

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

2.5. Marcadores moleculares

Los marcadores genéticos se definen como polimorfismos heredables que se han encontrado en una o más poblaciones o individuos y son la herramienta principal de la genética moderna. Dichos marcadores han permitido responder preguntas clave en materia de evolución, genética de poblaciones y expresión de genes. Conforme se han desarrollado las técnicas de la Biología Molecular, se han diseñado varios tipos de marcadores moleculares. La mayoría de estos se encuentran asociados a un locus de carácter cuantitativo (Quantitative Trait Loci o QTL). Muchas veces es necesario utilizar enzimas de restricción para producir una representación reducida del genoma de donde se puedan obtener los marcadores (Davey *et al.*, 2011).

Un buen marcador molecular debe reunir una serie de características para maximizar su utilidad: naturaleza altamente polimórfica, herencia co-dominante (determinación de los estados homocigotos y heterocigotos de los organismos diploides), distribución a lo largo del genoma, comportamiento selectivo neutral (las secuencias de ADN de cualquier organismo son neutrales a condiciones ambientales o prácticas de manejo), fácil acceso y/o disponibilidad, alta reproducibilidad y fácil intercambio de datos entre laboratorios, la técnica para analizar el marcador debe ser rápida, práctica y debe repetirse con fiabilidad en otros laboratorios (Cheng y Crittenden, 1994; Joshi *et al.*, 1999).

Con los avances de la genética moderna, ha sido posible estimar la diversidad genética de individuos y poblaciones a diferentes niveles jerárquicos, ya que ésta es determinada por la información contenida en el DNA, que se traduce de una secuencia de aminoácidos a una cadena polipeptídica, pues la mayoría de las enzimas y proteínas son productos de genes individuales, aunque en algunos casos las unidades que las forman están codificadas por dos o más genes; por ello la variabilidad iso-enzimática o protéica puede equipararse con la diversidad genética (Goodman y Stuber, 1983). Los marcadores moleculares basados en DNA como los microsatélites y otros tipos de marcadores, han ido ganando aceptación como una forma para determinar la diversidad genética. Los marcadores moleculares tienen ciertas ventajas sobre otro tipo de marcadores, ya que en general no son afectados por el ambiente en que las plantas se desarrollan y que son en la gran mayoría de los casos selectivamente neutros (Doebley, 1994).

Para efectuar estos estudios ha sido de gran importancia el desarrollo de la técnica de la reacción en cadena de la polimerasa (PCR, por sus siglas en inglés) (Mullis *et al.*, 1986), la cual es una reacción enzimática catalizada por una ADN polimerasa, que permite la amplificación o reproducción “in vitro” de grandes

cantidades de una región particular del ADN, a partir de una pequeña cantidad original de ADN molde. Esta enzima requiere para la síntesis de un par de “oligos” denominados iniciadores o “primers”, cuyas secuencias son complementarias a la de las regiones flanqueantes 5' y 3' del segmento en particular de ADN que se desea amplificar.

Los marcadores de repeticiones de secuencia simple (SSR) o microsatélites y polimorfismos de un solo nucleótido (SNP) son los preferidos para aplicaciones en genética y mejoramiento (Sharma *et al.*, 2013), pero con el advenimiento de las tecnologías de secuenciación de nueva generación, los marcadores SNPs y genotipificación por secuenciación han sido usados en diferentes especies de plantas.

Los SNP más frecuentes son mutaciones puntuales o pequeñas inserciones o deleciones (indels), las cuales se ha fijado en una parte significativa de la población de una especie ($\geq 1\%$), siendo las mutaciones más comunes las transiciones, donde ocurre un cambio de una purina por otra purina (A/G) o una pirimidina a otra pirimidina (C/T) o con menor frecuencia las mutaciones transversas, un intercambio de una purina por una pirimidina o viceversa (A/C, A/T, G/C, G/T). Los SNPs representan el 90 % de la variación genética total en cualquier organismo (Gupta *et al.*, 2008). Los SNPs pueden ocurrir en regiones codificantes o en regiones con función reguladora, pero son más frecuentes en regiones no codificantes (Borém y Fritsche-Neto, 2014). Realmente la existencia de SNPs no es nada nuevo, de hecho son la base molecular de varios tipos de marcadores moleculares que se desarrollaron con diferentes metodologías en las últimas tres décadas, entre ellos: RFLPs, RAPDs, AFLPs (Caetano, 2009).

2.6. Secuenciación de próxima generación (NGS)

La secuenciación del DNA tiene sus orígenes en la década de los setentas cuando Sanger *et al.* (1975), publicaron un método para determinar la secuencia de un genoma de 5000 pares de bases perteneciente a un virus. El método consistía en detener la fase de extensión en la reacción de síntesis de DNA; para ello utilizaron unas etiquetas moleculares conocidas como dideoxinucleótidos (ddNTPs, por sus siglas en inglés). Después, se separaba el producto de la reacción en un gel mediante electroforesis y la diferencia en el tamaño de los fragmentos así como el tipo de ddNTP permitía deducir la secuencia de la muestra original.

Más de una década después de la publicación original del método de Sanger, el surgimiento de la electroforesis capilar, mejores sistemas de detección para los ddNTPs y la automatización de varios procesos hicieron que la secuencia fuera más rápida y eficiente. Con dichos avances fue posible secuenciar fragmentos de mayor tamaño en un solo experimento (Kircher y Kelso, 2010).

Actualmente, con el surgimiento de nuevas tecnologías de secuenciación se ha logrado incrementar la densidad de los marcadores moleculares para una muestra y también se ha facilitado la producción de bibliotecas de marcadores incluso para organismos que no tienen información genómica (Hamblin *et al.*, 2011).

El lanzamiento a la venta del pirosecuenciador 454 FLX a finales de 2005 marcó el inicio de la oferta comercial de los secuenciadores de próxima generación (NGS); a partir de ese momento varias compañías comenzaron a comercializar nuevos secuenciadores basados en tecnologías diferentes, todos con el objetivo de aumentar la capacidad de procesamiento y reducir los tiempos requeridos para secuenciar fragmentos de DNA (Kircher y Kelso, 2010).

Los métodos de NGS producen lecturas de menor longitud que el método de Sanger. Esto se debe principalmente a un fenómeno conocido como desfaseamiento que es provocado por que las reacciones bioquímicas utilizadas durante la secuenciación, nunca tienen un rendimiento del 100 % y algunas moléculas se “atrasan” quedando fuera de sincronía. Entre mayor sea la longitud del fragmento a secuenciar existe una mayor probabilidad de que el desfaseamiento ocurra, así que los fabricantes de secuenciadores NGS optaron por reducir la longitud de la lectura en lugar de perder calidad de datos (Schadt *et al.*, 2010).

2.7. Genotipificación por secuenciación (GBS)

La disminución del precio de la secuenciación como resultado del surgimiento de los NGS, ha permitido que sea factible la producción de marcadores moleculares de forma directa en un secuenciador. A esto se le conoce como genotipificación por secuenciación (GBS, por sus siglas en inglés). Aunque el GBS es relativamente sencillo para organismos con genomas pequeños, conforme crece el genoma y la complejidad del genoma, como suele ser el caso de las plantas, es necesario recurrir a un método de reducción de complejidad para asegurar que la producción de marcadores sea homogénea y representativa (Elshire *et al.*, 2011).

La técnica de GBS por lo general se basa en el uso de enzimas de restricción tipo II para reducir la complejidad del genoma y de esta forma producir la biblioteca de fragmentos que será secuenciada (Davey *et al.*, 2011). El uso de este tipo de enzimas ofrece muchas ventajas: rápido, sencillo, altamente específico, reproducible y puede alcanzar regiones del genoma que no son fáciles de estudiar utilizando otros métodos. Sin embargo, es importante recalcar

que el éxito de la reducción de complejidad depende en gran medida de la combinación de enzimas de restricción que se selecciona para cada organismo. Se debe buscar una combinación que evite producir muchos marcadores en regiones repetitivas del genoma o zonas de baja expresión. Para ello, por lo general se selecciona al menos una enzima sensible a la metilación ya que las zonas de menor expresión del genoma tienden a estar metiladas (Elshire *et al.*, 2011).

Una vez producida la biblioteca mediante la reducción de complejidad y la amplificación selectiva, se secuencian los fragmentos resultantes. Posteriormente, se debe analizar la información producida para seleccionar los marcadores moleculares que tengan una mayor calidad de datos y que permitan detectar polimorfismos entre las diferentes muestras. Una de las ventajas del método GBS que detecta polimorfismos de nucleótido simple (SNP, por sus siglas en inglés) inclusive en organismos para los cuales no se tiene el genoma de referencia (Davey *et al.*, 2011; Elshire *et al.*, 2011).

2.8. Detección del genotipo y polimorfismo usando NGS

Una de las desventajas de la producción de marcadores moleculares usando NGS es que la calidad de las lecturas disminuye en base a su longitud. Por lo tanto, antes de hacer un análisis de los datos producidos por un secuenciador es necesario determinar si la información cumple con un estándar de calidad. Aquellas secuencias que no cuentan con suficiente calidad para garantizar su utilidad en un análisis posterior deben ser eliminadas. Este proceso de análisis de datos se subdivide en tres etapas principales: convertir la información obtenida del secuenciador de datos útiles (pre-procesamiento), determinar los marcadores utilizando algoritmos que comparten los fragmentos con los resultados de otras muestras obtenidas en la misma corrida (procesamiento) y finalmente comparar

los fragmentos que han pasado los dos filtros anteriores con información reportada anteriormente o incluso un genoma de referencia, con el propósito de obtener únicamente los fragmentos más representativos del genotipo (post-procesamiento) (Nielsen *et al.*, 2011).

III. MATERIALES Y MÉTODOS

3.1. Fuente de germoplasma

Se estudió un conjunto de 290 accesiones nativas pertenecientes a la Colección Núcleo del Banco de Germoplasma de Maíz del CIMMyT (Ver apéndice 7.1), adaptadas a condiciones tropicales y subtropicales y que representan diferentes grupos raciales, las cuales fueron previamente seleccionadas a través de técnicas de los Sistemas de Información Geográfica (SIG), en función de su probable adaptación y resistencia al estrés hídrico. La mayoría de estos genotipos fueron colectados en México, Argentina, y en menor proporción provenientes de otros países de Sudamérica, como Brasil y Perú.

3.2. Caracterización molecular

Las accesiones fueron sembradas en charolas, bajo condiciones controladas en el invernadero de CIMMyT, ubicado en la estación experimental de El Batán, con coordenadas 19°31'46.6"N 98°50'42.5"W. Se usaron 30 plantas elegidas al azar de cada accesión para hacer una muestra compuesta de ADN ("bulk"). Se cosecharon hojas frescas y sanas de plantas de alrededor de 30 días de edad. El tejido vegetal se congeló a -80 °C, luego se liofilizó, para finalmente cortar un disco de igual área de tejido de cada hoja.

Este tejido fue depositado en tubos de extracción, donde cada tubo tenía una esfera metálica pequeña la cual auxilió en el proceso de molienda. Para luego, extraer el ADN genómico mediante la utilización del método CTAB de todas las accesiones (CIMMyT, 2006).

El ADN fue cuantificado por medio de un NanoDrop 8000, con la finalidad de conocer la concentración y calidad del producto obtenido a partir de la extracción, para posteriormente diluirlo a la misma concentración (200 ng/ul) y presentarlo para su procesamiento de genotipificación por secuenciación a través del método DArTseq™, en diferentes organismos como: Diversity Arrays Technology (DArT), Servicio de Análisis Genético para la Agricultura (SAGA) y el Instituto de Biotecnología de la Universidad de Cornell, Ithaca, EUA, a fin de obtener los marcadores tipo SNP.

3.3. Control de calidad y análisis de los datos

Marcadores cuya frecuencia del alelo menor (MAF, por sus siglas en inglés) fue menor de 0.01 fueron descartados, ya que usualmente producen resultados inestables (Corvin *et al.*, 2010; Weng *et al.*, 2011). Con este primer filtrado se eliminaron también aquellos marcadores monomórficos. Un segundo filtrado se realizó para descartar aquellos marcadores con más de 40 % de datos faltantes. De tal forma se obtienen marcadores que ofrezcan información útil para el estudio.

Los parámetros de diversidad genética, como número de alelos, diversidad genética (heterocigosidad esperada) y las distancias genéticas entre accesiones/poblaciones obtenidas mediante el uso de la proporción de alelos compartidos, se calcularon con el paquete estadístico de libre acceso R (*R CoreTeam*, 2015).

El coeficiente de disimilitud usado en el estudio fue la distancia euclidiana conocida como Distancia Modificada de Rogers (MDR, por sus siglas en inglés) (Wright, 1978) y que se define de la manera siguiente:

distancia Modificada de Rogers (MR_{xy})
(entre un par de genotipos x,y)

$$0 \leq MR_{xy} = \frac{1}{\sqrt{2L}} \sqrt{\sum_{l=1}^L \sum_{a=1}^{n_l} (\hat{p}_{lax} - \hat{p}_{lay})^2} \leq 1$$

$\hat{p}_{lax}$ es la frecuencia estimada del alelo a , dentro del locus l , en el genotipo x ; L number de locus, n_l el número de alelos en el l -ésimo locus (si no diploide reemplace $2L$ por Nro. alelos)

Para la representación gráfica de la diversidad genética en las poblaciones de estudio se utilizó el programa DARwin 6.0 (<http://darwin.cirad.fr/Home.php>), para generar el dendrograma por el método Neighbor-Joining basado en la matriz de distancias genéticas, con el propósito de obtener una estructura de las relaciones entre los genotipos tropicales y subtropicales involucrados en este estudio.

También se usaron los métodos de agrupamiento conocidos como Análisis de Componentes Principales (PCA) y Escalamiento Multi Dimensional (MDS), primero a través del programa estadístico R (*R CoreTeam*, 2015) y posteriormente los datos de categorización fueron trazados de manera tridimensional por el programa CurlyWhirly v1.15 (<http://ics.hutton.ac.uk/curlywhirly/>).

3.4 Tendencia inicial de selección dentro de las colecciones de estudio

Durante el año 2014 se evaluaron en campo las 290 colecciones nativas de maíz, tropicales y subtropicales, respectivamente; en dos estaciones experimentales de INIFAP, la primera ubicada en Santiago Ixcuintla, Nayarit y la segunda en El Valle del Fuerte, Sinaloa. También es importante mencionar que durante este mismo año se evaluaron en otra estación de INIFAP ubicada en Iguala, Guerrero;

189 genotipos que corresponden a los de origen y adaptación tropical, y que a su vez, están involucrados en este estudio de diversidad genética.

Estas colecciones fueron manejadas a través de riegos controlados durante su etapa de desarrollo y sometidas a un tratamiento de estrés hídrico durante el periodo de prefloración y floración, masculina y femenina, respectivamente. Posteriormente se regó a capacidad de campo para la etapa de llenado de grano.

Con los datos obtenidos, se realizó un ranking de los genotipos con mayor potencial de resistencia a la sequía, tomando como parámetro principal el rendimiento, expresado en gramos; bajo el supuesto que el rendimiento de grano de maíz bajo condiciones de estrés hídrico permite discernir el grado de tolerancia a la sequía (Bazinger *et al.* 2000).

De manera similar, en el año 2015 (ciclo A) se establecieron ensayos para la evaluación de los 121 genotipos seleccionados por su posible fuente de resistencia a la sequía; en tres sitios experimentales: Ciudad Obregón, Sonora (CIMMyT), Valle del Fuerte, Sinaloa (INIFAP) y Santiago Ixcuintla, Nayarit (INIFAP). Esta vez, se establecieron en 2 tratamientos diferentes: riego normal (sin estrés) y riego reducido (con estrés), de igual manera el tratamiento de sequía se aplicó en etapas de prefloración y floración; para finalmente reanudar los riegos en la fase de llenado de grano.

Estos datos permitieron al mejorador, filtrar y avanzar algunos materiales que demuestran ser germoplasma valioso y/o con potencial de utilizar en su esquema de mejoramiento.

IV. RESULTADOS Y DISCUSIÓN

4.1. Control de calidad y selección de los marcadores

Un total de 616,968 SNPs fueron reconocidos y denominados con éxito dentro del conjunto de datos en las colecciones nativas. Los datos derivados de este análisis incluyen tanto la denominación de los SNP, así como también el número de secuencias alternativas leídas por alelo, lo que nos daría la oportunidad de poder identificar también haplotipos. Los marcadores SNPs a ser utilizados fueron condicionados a cumplir con ciertos requisitos en relación a algunas características de calidad y los valores de las mismas que los polimorfismos deberían poseer. Para ello, un proceso de filtrado fue aplicado y se eliminaron los marcadores con una proporción de datos perdidos o no categorizados $>40\%$ y que presentaron una menor frecuencia alélica (MFA) $<1\%$. Estas exigencias eliminaron un 94.04% de los marcadores totales, una cantidad considerable de datos, dejando sólo aquellos que mostraron la calidad necesaria para cumplir con los objetivos de nuestro estudio. Así, se seleccionaron un total de 36,761 SNPs de excelente calidad y que fueron calificados como los más informativos; que para estudios de diversidad genética resultan suficientes para diferenciar y agrupar a las colectas (Cuadro 1).

Cuadro 1. Número de marcadores SNPs y su filtrado.

Parámetros de filtrado de los marcadores	Número de marcadores
Total de marcadores	616,968
- Proporción de datos perdidos > 0.4	363,111
- Marcadores con MFA < 0.01	540,483
Marcadores que cumplan estas dos condiciones:	
Número final de marcadores útiles	36,761

El número total de marcadores (Cuadro 1) responde a lo esperado en muchas plataformas de alto rendimiento para la generación de SNPs, a excepción de que en la tecnología DArTSeq®, no se realiza imputación estadística alguna sobre los datos faltantes o perdidos, de tal forma, se eliminaron aquellos SNPs con más del 40 % de datos perdidos en los genotipos. El número final de marcadores útiles está dado en función del filtrado estricto de la menor frecuencia alélica ($< 1\%$) presente, ya que los marcadores que se encuentran dentro de este rango pueden ser debidos a posibles fallas de lectura en el secuenciador, o ciertamente, a la presencia de alelos raros. Los alelos raros aunque resultan interesantes, no se contemplaron al tiempo del análisis, en función del carácter cuantitativo de estudio al que están relacionados los genotipos, que es la resistencia a sequía, donde se pretende generar la acumulación de muchos efectos aditivos, más que los efectos de tipo dominante.

4.2. Parámetros de diversidad genética

4.2.1. Heterocigosidad esperada y observada

El Cuadro 2 muestra que la heterocigosidad observada (H_{obs}) para las diferentes accesiones analizadas por locus fue alta, con un valor promedio de 0.301. Esto permite confirmar la naturaleza heterocigota de las colecciones nativas de maíz, posiblemente debida al alto grado de recombinación que poseen, propio del sistema de reproducción sexual del cultivo. La heterocigosidad esperada fue menor a la observada, con un valor promedio de 0.215, cuyo valor representa también una alta heterogeneidad genética en las poblaciones involucradas en este estudio, bajo condiciones de apareamiento aleatorio y equilibrio de Hardy-Weinberg.

Los mayores niveles de heterocigosidad esperada (desde 0.45 hasta 0.5) estuvieron presentes en 5192 marcadores y los valores menores de H_{esp} (0.0198 - 0.05) aparecieron en 7684 de los SNPs analizados.

4.2.2. Estadísticos F

En el Cuadro 2 se muestra el valor del coeficiente de endogamia (F_{is}) estimado para las diversas poblaciones estudiadas en este trabajo a partir de los marcadores SNP; cuyo valor negativo (-0.672) expresa de manera importante la presencia de bajos niveles de endogamia en las accesiones preseleccionadas para este análisis de genotipos con posible resistencia a sequía, en función de las desviaciones de los cruzamientos al azar. Este dato también resulta consistente con la alta heterocigosis explicada anteriormente y que permite reconocer la diversidad genética presente en la colección núcleo de estudio.

El coeficiente de fijación F_{st} , permitió medir el grado de reducción en la heterocigosidad de las subpoblaciones o colectas analizadas (0.155), respecto a nuestra población total, debida a los cruces no aleatorios y el nivel de fijación de alelos alternos; de tal forma, se determina que la diferenciación genética entre las colecciones o subpoblaciones es grande.

Cuadro 2. Parámetros de diversidad genética para marcadores tipo SNP.

Heterocigosidad observada (H_{obs})	Heterocigosidad esperado (H_{esp})	Coeficiente de endogamia F_{is}	Coeficiente de fijación F_{st}
0.301	0.215	-0.672	0.155

4.3. Análisis de conglomerados

La representación gráfica de las accesiones al ser comparadas entre sí (Figura 1), se obtuvo a partir del método por agrupamiento de Neighbor Joining y de las distancias genéticas calculadas mediante MDR (Modified Distance Roger's, Distancia Modificada de Roger's) y que representa las relaciones genéticas existentes entre las 290 accesiones analizadas (Ver Apéndice 7.1). Este análisis identificó dos grandes grupos que separan mayoritariamente a las accesiones sudamericanas (argentinas, brasileñas y peruanas) de las mexicanas. El grupo sudamericano a su vez se puede dividir en cinco subgrupos, en términos de similitudes y diferencias moleculares, usando muestras de ADN conjuntas (bulk) y marcadores tipo SNP; y que finalmente, demuestran tener relación con sus grupos raciales y origen geográfico:

El subgrupo 1 se integró por 39 colectas diferentes, la mayoría de procedencia Argentina, excepto una de México, de nombre SONO162 de raza Onaveño. Las razas que predominan en este grupo son: Crisco, Comtro y Cameli.

El subgrupo 2 se formó con 9 genotipos, todos de origen Argentino y de la raza Pisingallo.

El subgrupo 3 fue el más pequeño y poco definido, con solo 5 colecciones nativas, a simple vista no fue evidente la relación entre sus características fenotípicas, geográficas o raciales, aunque genéticamente se encuentran muy cercanas. Sus orígenes corresponden a Argentina (1), Perú (1) y México (3).

El subgrupo 4 quedó integrado por 11 colectas, de dos razas principales, que son: Comtro y Dembla, de Argentina y Brasil como origen, esto puede guardar relación directa entre las distancias genéticas y geográficas.

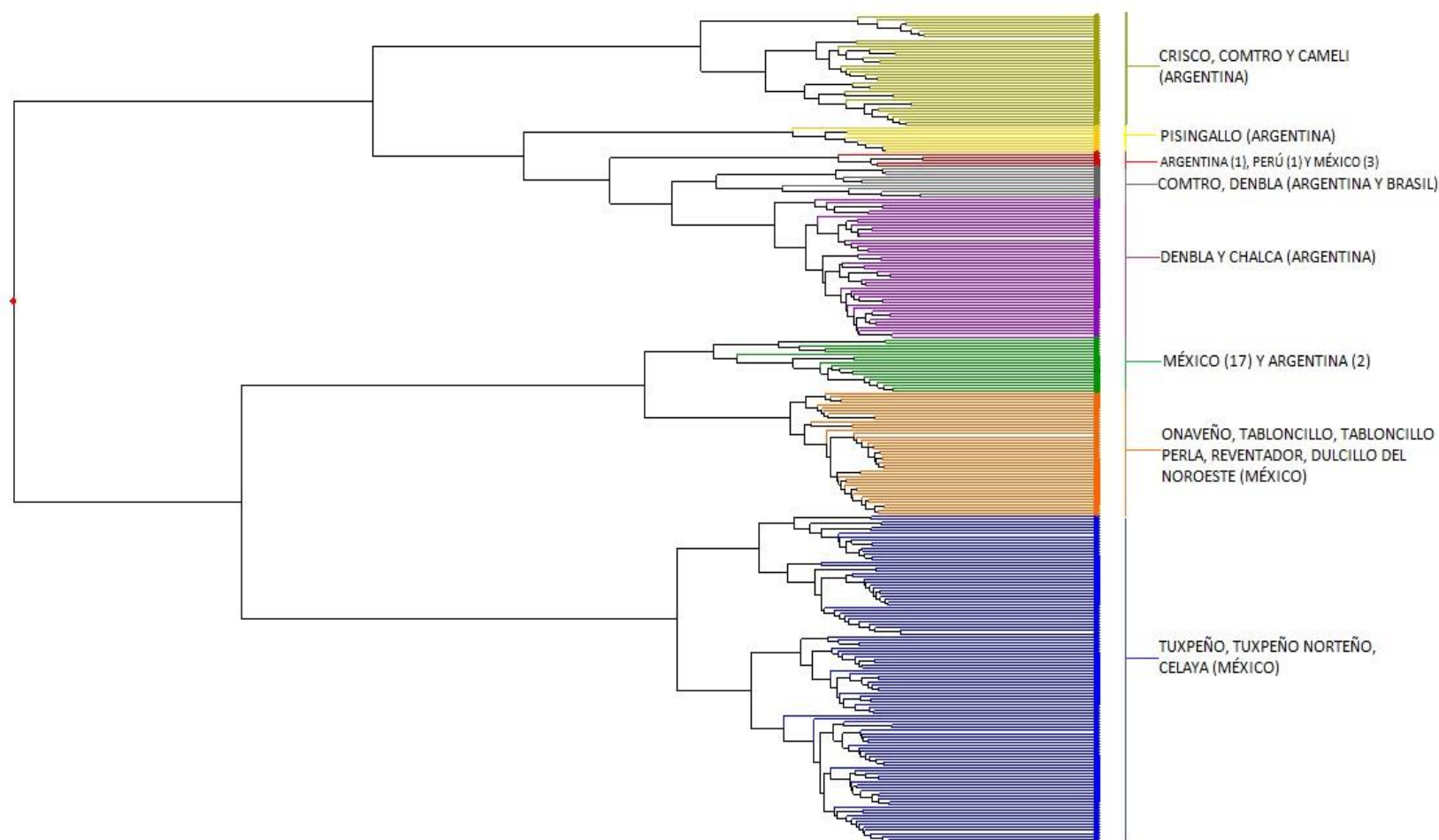


Figura 1. Dendrograma de relaciones entre 290 colecciones nativas de maíz, obtenido a partir de marcadores tipo SNP.

El subgrupo 5 correspondió a 49 colecciones diferentes y que se relacionaron por su origen geográfico, ya que 45 de estas son de Argentina y de las razas Denbla y Chalca, principalmente. Dentro del grupo hay 4 genotipos de México (CHIH70, CHIH117, CHIH116 Y CHIH119), todos procedentes del estado de Chihuahua.

El grupo de accesiones mexicanas se puede clasificar también en tres subgrupos:

El subgrupo 1 quedó formado por 19 genotipos, 17 de México 2 de Argentina. No es posible identificar algún patrón racial, ya que la mitad de las entradas no están clasificadas aún, pero es conveniente mencionar que once de ellas son del Estado de Chihuahua, por lo que su distancia genética y geográfica, pueden guardar mucha relación.

El subgrupo 2 mexicano es el tercer grupo más grande con 43 colecciones, compuesto principalmente por ejemplares de México, de los estados de Sonora y algunos de Sinaloa. Aquí predominan las razas: Onaveño, Tabloncillo y Tabloncillo Perla, que pertenecen al grupo de maíces llamados “Ocho hileras”, propios de la región Occidente de México. También hay presencia de algunas razas conocidas como: Reventador y Dulcillo del Noroeste, dentro de esta agrupación.

Por último, el subgrupo 3 de México resultó ser el más grande de todos con 115 colecciones nativas que comparten relaciones genéticas y raciales; Tuxpeño, Tuxpeño Norteño y Celaya, son las que predominan dentro del grupo. También es posible dilucidar que la mayoría de las colectas tienen origen en los estados mexicanos muy próximos al Trópico de Cáncer.

4.4. Análisis de Escalado Multi Dimensional (MDS)

Se realizó un análisis MDS a los datos obtenidos a partir de la matriz de distancias genéticas modificadas de Roger's, con el propósito de identificar la distribución adecuada de los genotipos y de las cuales se pudo obtener los resultados que se muestran en las Figuras 2, 3 y 4; las imágenes están dadas en 3 dimensiones, y se visualizan de maneras diversas, de acuerdo con la categoría deseada, Ej.: por raza, por país o por grupos de similitud genética.

En la Figura 2, se visualiza la distribución espacial de los diversos genotipos de la colección núcleo estudiada, en función de sus relaciones entre razas y que por sus diferentes tonalidades color verde, es fácil identificar a los siguientes grupos raciales: Comtro, Craman, Crisco, Denama y Denbla. Muy cerca de ellos se observa un grupo color morado, que corresponde a la raza Pisingallo; todas las colecciones mencionadas son razas de origen argentino. De igual manera, se observa otra agrupación importante de razas mexicanas con tendencias de color que van del rojo al rosado, que representan principalmente los siguientes grupos raciales: Reventador, San Juan, Tabloncillo, Tabloncillo Perla, Tuxpeño y Tuxpeño Norteño. Cercanamente también se aprecian algunas colectas en color azul, que corresponden a la raza Onaveño.

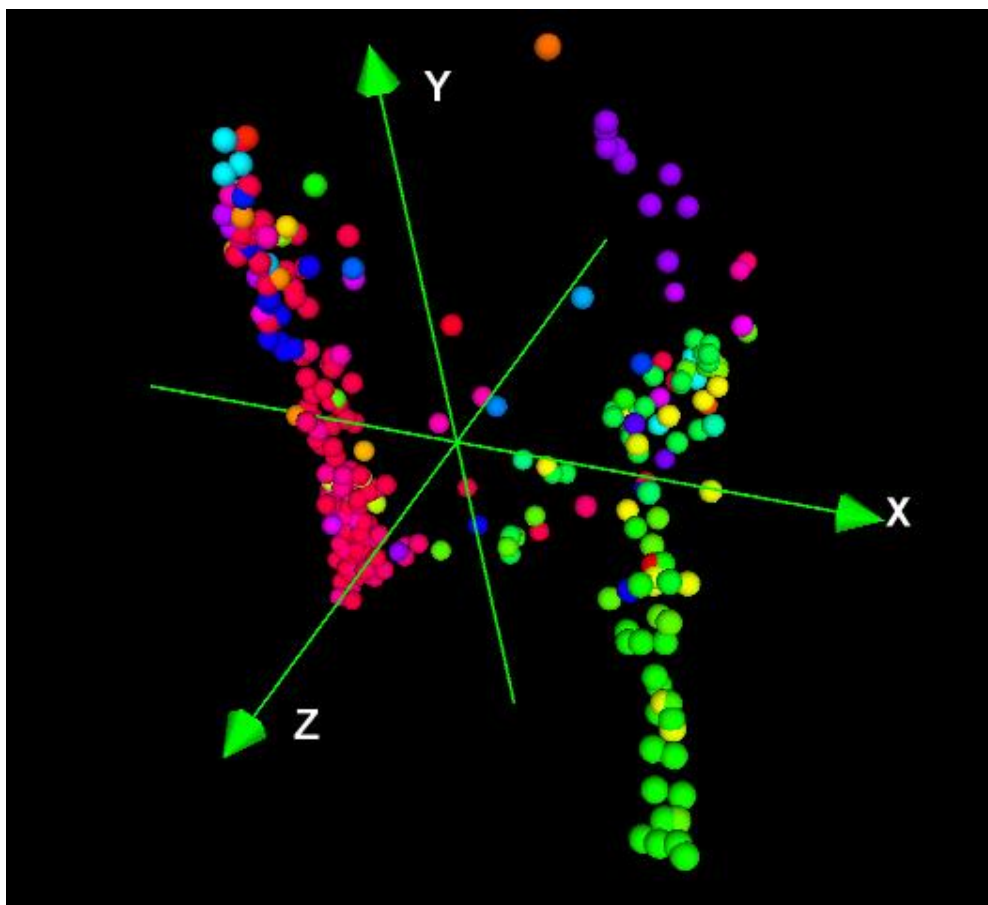


Figura 2. Gráfico tridimensional del análisis MDS por categoría de razas.

En la Figura 2 se observan puntos con diferentes tonalidades de color verde identifican a los grupos raciales Comtro, Craman, Crisco, Denama y Denbla; los de color morado corresponden a la raza Pisingallo; todas las colecciones mencionadas son razas de origen argentino. La agrupación de razas mexicanas se muestra con tendencias de color que van del rojo al rosado, representando principalmente los grupos raciales Reventador, San Juan, Tabloncillo, Tabloncillo Perla, Tuxpeño y Tuxpeño Norteño. Cercanamente también se aprecian algunas colectas en color azul, que corresponden a la raza Onaveño.

Desde una perspectiva diferente, en la Figura 3 es posible observar la conjunción en tres dimensiones de todas las colecciones de estudio, mismas que se encuentran coloreadas por categoría de grupo. Se observan los mismos ocho subgrupos identificados anteriormente en el dendrograma, en términos de sus similitudes y diferencias genéticas. En colores rojo, azul y amarillo, se observa fácilmente un concentrado de genotipos, que en su mayoría son de origen mexicano. El resto de los grupos se encuentran más dispersos dentro y entre sí.

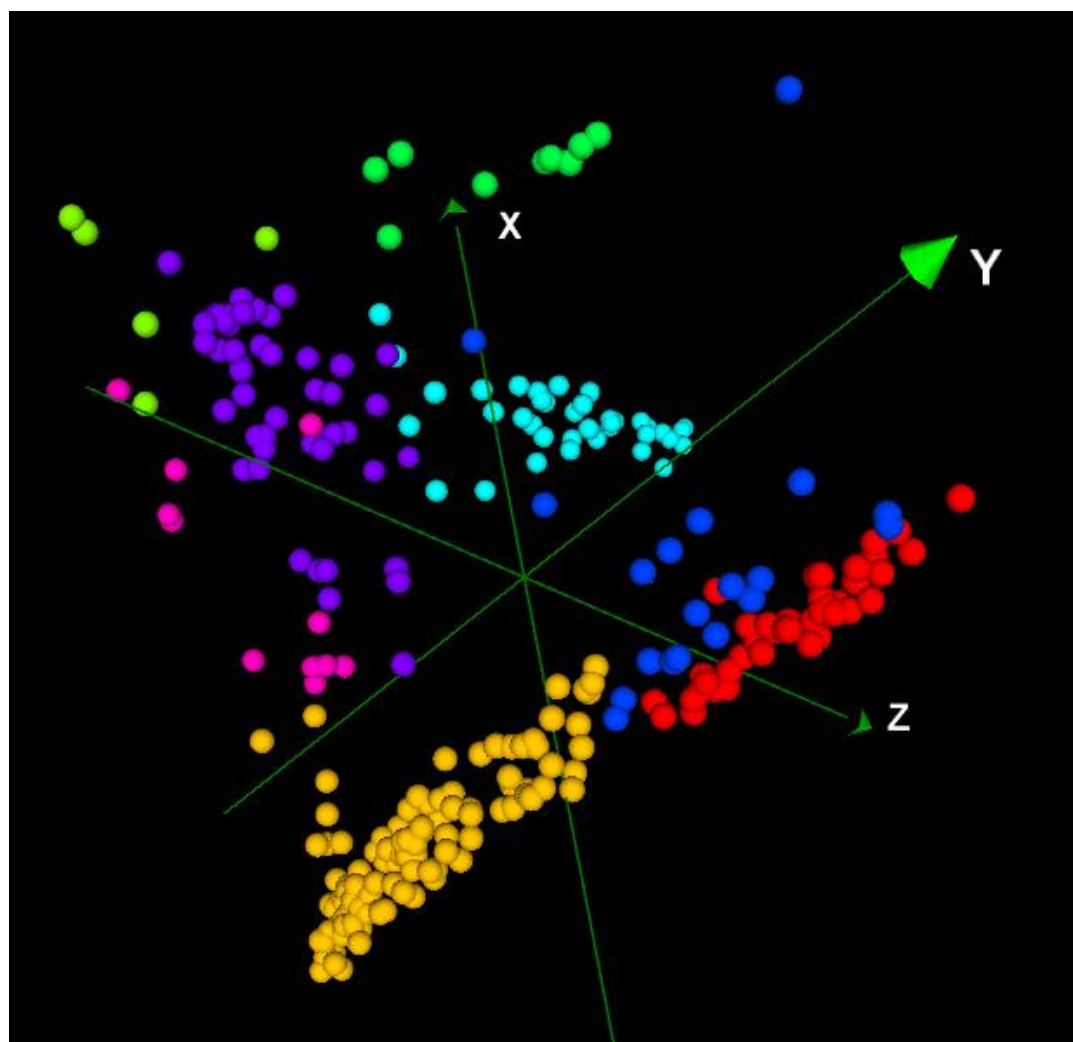


Figura 3. Grafico tridimensional del análisis MDS por categoría de grupos.

Análogamente, en la Figura 4 se aprecia la distribución espacial de las colecciones nativas de maíz, en función de la categoría de país de origen. En color amarillo, se observan las entradas de origen mexicano; en color rojo se encuentran los genotipos argentinos; y como minorías: Brasil en color blanco y Perú en color morado. Es evidente la estrecha relación existente entre las distancias genéticas calculadas a partir de marcadores tipo SNP y el origen geográfico de las colectas.

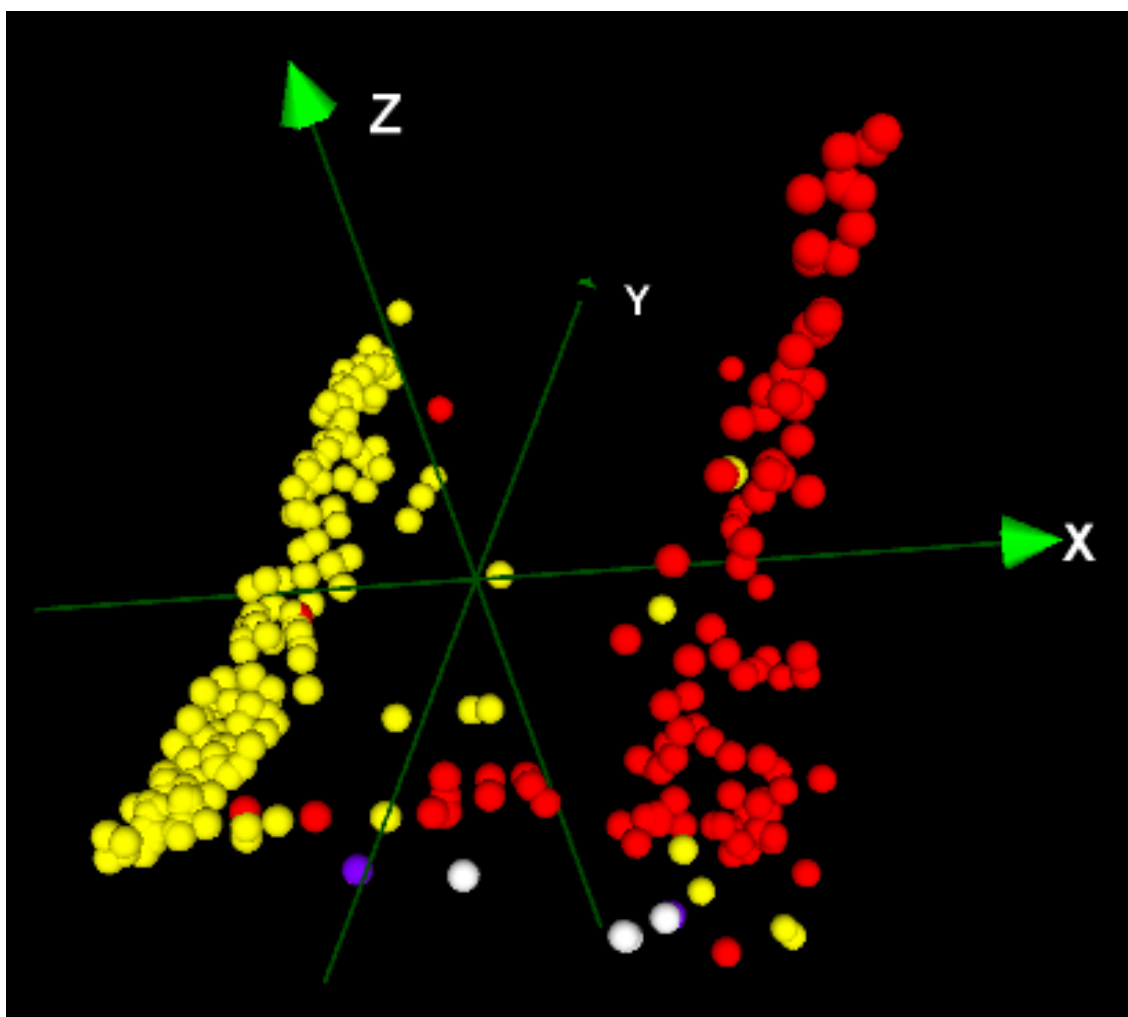


Figura 4. Grafico tridimensional del análisis MDS por categoría de país de origen.

Los puntos en color amarillo representan las accesiones de origen mexicano, los de color rojo a los genotipos argentinos, los de Brasil en color blanco y los de Perú en color morado.

4.5. Análisis de Componentes Principales (PCA)

El análisis de componentes principales (PCA) para las 290 colecciones nativas de maíz, usando los marcadores SNP seleccionados, reprodujeron una separación clara de grupos basados en su raza. De tal forma que, graficando los primeros tres componentes principales para todos los genotipos involucrados en este estudio, se distribuyeron en el espacio tridimensional de manera similar y consecuente al escalamiento multidimensional explicado anteriormente.

La Figura 5 permite dilucidar la consistencia de los datos y los análisis basados en las distancias genéticas de las colectas. Es posible identificar al extremo izquierdo del gráfico, en colores que van del rojo al rosado y también algunos puntos azules, el cúmulo de razas, mismas que, son características representativas de México, como lo son: Reventador, San Juan, Tabloncillo, Tabloncillo Perla, Tuxpeño, Tuxpeño Norteño y Onaveño.

Análogamente, al lado derecho de la Figura 5, se observan concentrados en tonalidades verdes, amarillas y algunas moradas, la distribución de los genotipos argentinos, haciendo evidente la relación estrecha entre sus grupos raciales y su origen geográfico. Las razas Comtro, Craman, Crisco, Denama, Denbla.y Pisingallo, se aprecian a primera vista.

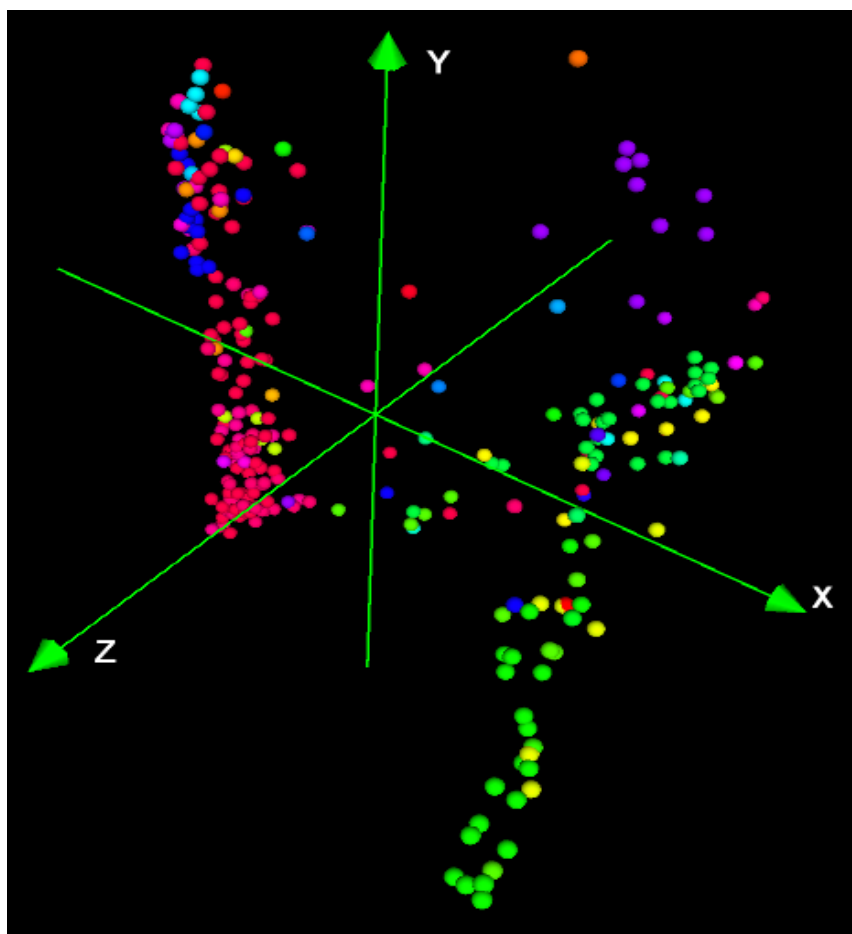


Figura 5. Grafico tridimensional del análisis de componentes principales para los primeros tres componentes

En el extremo izquierdo del gráfico anterior, en puntos de colores que van del rojo al rosado y también algunos puntos azules, se muestra el cúmulo de razas representativas de México, como lo son Reventador, San Juan, Tabloncillo, Tabloncillo Perla, Tuxpeño, Tuxpeño Norteño y Onaveño. Al lado derecho de la figura, con puntos concentrados en tonalidades verdes, amarillas y algunas moradas se muestra la distribución de los genotipos argentinos, las razas Comtro, Craman, Crisco, Denama, Denbla y Pisingallo, se aprecian a primera vista.

4.6. Colecciones nativas con potencial de resistencia a sequía

A partir de datos derivados de una investigación paralela al estudio y aprovechamiento de la diversidad genética en 290 genotipos nativos de maíz y su evaluación en campo, fue posible identificar a las accesiones con mayor potencial de resistencia a sequía, a través de un ranking de rendimiento en diferentes ambientes durante los años 2014 y 2015; y tratamientos de estrés controlado.

De tal forma que, en el sitio experimental de Santiago Ixcuilitla (2014) se observaron rendimientos muy bajos en algunas accesiones bajo condiciones de estrés hídrico, ej.: las colectas con menor rendimiento fueron CHZM04021 (76.67 g) y ARZM17027 (116.67 g). También fue posible identificar rendimientos altos, de más de 900 g/parcela en colecciones como: SONO153, SINA 8 y SINA 9.

Para el ensayo evaluado en INIFAP, Valle del Fuerte (2014); también se observaron valores que permiten dilucidar algunos genotipos que toleran la sequía; manifestando rendimientos de hasta 1153.33 g por parcela (COAH16).

En el sitio experimental de Iguala, Gro. (2014), donde se evaluaron de manera exclusiva los materiales de origen tropical, los valores máximos de rendimiento oscilaron de 491 a 603 gramos, siendo más bajos en comparación a los otros ambientes de evaluación durante el mismo año, esto puede estar dado por la severidad con que fue aplicado el tratamiento de estrés hídrico; ya que algunos materiales no produjeron ni 5 g/parcela.

Con muchos de estos datos de rendimiento, fue posible filtrar las colecciones nativas que parecen poseer un potencial genético que permita ampliar la base de la resistencia a sequía en maíz. Permitiendo avanzar hacia una fase nueva de evaluación, con 121 genotipos seleccionados de un total inicial de 290; de los

cuales 41, corresponden a colecciones subtropicales y el resto a de estas son de adaptación tropical.

De las evaluaciones fenotípicas del año 2015, no fue posible obtener datos de la localidad de Santiago Ixcuintla, Nayarit; en función del intenso periodo de lluvias presente durante la fase de prefloración y floración de las accesiones sometidas a estrés. En los sitios experimentales de Ciudad Obregón, Son. y Valle del Fuerte, Sin., fue posible seleccionar genotipos que manifestaran potencial de rendimiento en ambas condiciones de humedad, en condiciones de riego reducido y de riego normal; a través del ranking de rendimiento y otros criterios de selección, elegidos por el mejorador.

Los seis rendimientos máximos bajo condiciones de riego normal en el sitio experimental de Ciudad Obregón, Son., se ubicaron en el rango de 1,943 a 1,503 gramos; mientras que en riego reducido, el grano producido en esa proporción de genotipos osciló de 978 a 730; estimando una reducción de casi el 50 % del rendimiento, a causa del estrés inducido.

Análogamente, en el Valle del Fuerte, Sin. (2015), la producción de grano se redujo bajo condiciones de riego reducido. Los rendimientos más elevados para los genotipos sometidos a estrés hídrico corresponden a colecciones subtropicales, principalmente; con pesos dentro del rango siguiente: 1780 a 1153 gramos, para las primeras seis accesiones. Mientras que, en condiciones de riego normal, la misma proporción de genotipos tuvieron valores de 2270 a 1791 g; siendo también las colecciones subtropicales, las mejor ubicadas dentro del ranking de rendimiento

De tal manera, hasta el momento se han seleccionado 20 genotipos subtropicales del panel inicial de 290 colecciones nativas, para su uso y aprovechamiento,

como fuente de diversidad genética para ampliar la base de la resistencia a sequía en maíz.

En la Figura 6, se observan las 20 accesiones seleccionadas a partir su comportamiento y rendimiento en campo en condiciones de estrés hídrico y su descripción gráfica en función de las distancias genéticas propias: ARZM12193, ARZM12236, ARZM12237, ARZM12240, CHIH333, CHIH338, CHIH59, CHIH85, CHIH87, COAH117, COAH18, COAH19, COAH20, COAH21, COAH78, NVOL19, NVOL46, SNLP166, SNLP169, SNLP17; en color verde se aprecian las colecciones de origen argentino; de las cuales, tres de ellas son de raza Dembla y una pertenece al grupo racial Crisco. El resto de los genotipos son de origen mexicano y todas corresponden al grupo donde se catalogaron las razas Tuxpeño Norteño, Tuxpeño y Celaya.

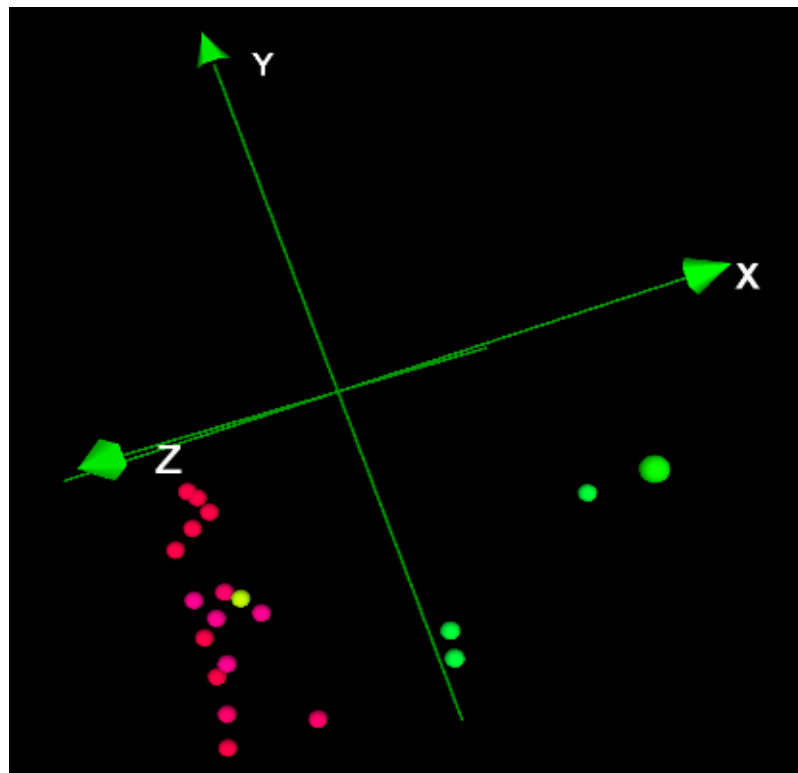


Figura 6. Gráfico tridimensional de las 20 colecciones nativas seleccionadas por su tolerancia a sequía y su relación en términos de distancias genéticas.

V. CONCLUSIONES

De los resultados obtenidos a partir de 36,761 marcadores tipo SNP, producto de la genotipificación por secuenciación (GbS) de la plataforma DArTseq™ y con la aplicación de los diferentes análisis estadísticos se derivan las conclusiones siguientes:

Existe diversidad genética importante en las 290 colecciones nativas de maíz provenientes del Banco de Germoplasma del CIMMYT evaluadas, en el presente estudio.

Los parámetros de diversidad genética calculados indican la presencia de una alta variabilidad en los genotipos tropicales y subtropicales.

A partir del cálculo de las distancias genéticas (MDR) y el análisis de conglomerados se diferenció a las accesiones sudamericanas de las mexicanas definiéndose un total de ocho subgrupos en base a su distribución geográfica y racial.

Los análisis de escalado multidimensional (MDS) y el de componentes principales (PCA), resultaron eficientes para identificar la distribución espacial en tres dimensiones de cada uno de los genotipos, lo que permitió observar y precisar el grado de divergencia genética entre cada una de las colecciones nativas.

La relación entre las distancias genéticas de las colecciones nativas y su selección en campo en condiciones de estrés hídrico, permite concluir que la diversidad existente en las razas Tuxpeño y Tuxpeño Norteño parece seguir siendo clave en los avances del mejoramiento genético moderno en maíz.

Este trabajo da información que puede servir para el mejor uso y aprovechamiento de la diversidad genética existente en los bancos de

germoplasma de maíz, de tal forma que, acompañado de un esquema de mejoramiento y evaluaciones fenotípicas se pueden identificar genotipos o alelos favorables que amplíen la base genética de la resistencia a sequía o a algunos otros factores tanto bióticos como abióticos.

VI. BIBLIOGRAFÍA

Bazinger M, Edmeades GO, Beck D, Bellon M (2000) Breeding for drought and nitrogen stress tolerance in maize: from theory to practice. CIMMYT, México, D.F.

Bedoya, C.A., Zaharieva, M., Dreisigacker, S., Cortes-Cruz, J. y Warburton, M.L. (2008). Marcadores moleculares: Aplicaciones y Avances en el CIMMYT. *Claridades Agropecuarias* 184:36-47.

Bonnal, P., Bosc, P. M., Díaz, J. M., & Losch, B. (2003). Multifuncionalidad de la agricultura y nueva ruralidad. Reestructuración de las políticas públicas a la hora de la globalización. Ponencia presentada en el Seminario Internacional "El mundo rural: transformaciones y perspectivas a la luz de la nueva ruralidad", desarrollado en la Universidad Javeriana, CLACSO, REDCAPA, Bogotá.

Borém, A., & Fritsche-Neto, R. (2014). *Biotechnology and Plant Breeding: Applications and Approaches for Developing Improved Cultivars*. Elsevier. 270 p.

Bruce, W. B., Edmeades, G. O., & Barker, T. C. (2002). Molecular and physiological approaches to maize improvement for drought tolerance. *Journal of experimental botany*, 53(366), 13-25.

Brush, S. B. (1998). Bio-cooperation and the benefits of crop genetic resources: the case of Mexican maize. *World Development*, 26(5), 755-766.

Brush, S. B., & Perales, H. R. (2007). A maize landscape: Ethnicity and agrobiodiversity in Chiapas Mexico. *Agriculture, Ecosystems & Environment*, 121(3), 211-221.

Caetano, A. R. (2009). Marcadores SNP: conceitos básicos, aplicações no manejo e no melhoramento animal e perspectivas para o futuro. *Revista*

Brasileira de Zootecnia, 38(8), 64-71.

Castillo, G. F. (1993). La variabilidad genética y el mejoramiento genético de los cultivos. In: Ciencia. Revista de la Academia de la Investigación Científica. México. No. Especial. pp: 69-79.

Cheng, H. H., & CRITTENDEN, L. B. (1994). Microsatellite markers for genetic mapping in the chicken. Poultry Science, 73(4), 539-546.

ChongTao, W., & ShaoKun, L. (2010). Assessment of limiting factors and techniques prioritization for maize production in China. Scientia Agricultura Sinica, 43(6), 1136-1146.

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. 92 p.

Comisión intersecretarial de Bioseguridad y organismos genéticamente modificados. México. (2002). Marco regulatorio en organismos genéticamente modificados. 27 p.

Cook, E. R., Seager, R., Cane, M. A., & Stahle, D. W. (2007). North American drought: reconstructions, causes, and consequences. Earth-Science Reviews, 81(1), 93-134.

Corvin, A., Craddock, N., & Sullivan, P. F. (2010). Genome-wide association studies: a primer. Psychological medicine, 40(7), 1063.

Corvin, A., Craddock, N., & Sullivan, P. F. (2010). Genome-wide association studies: a primer. Psychological Medicine 40(07):1063–1077.

Cubillos, P. A. (1997). Aspectos ambientales, ideológicos y políticos en el debate de la bioprospección y uso de los recursos genéticos en Chile. 3 p. Sociedad de Botánica de Chile. Viña del Mar, Octubre 9 de 1996

CurlyWhirly v1.15 (<http://ics.hutton.ac.uk/curlywhirly/>).

DARwin 6.0 (<http://darwin.cirad.fr/Home.php>).

Davey, J. W., Hohenlohe, P. A., Etter, P. D., Boone, J. Q., Catchen, J. M., & Blaxter, M. L. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics*, 12(7), 499-510.

Doebley, J. (1994). Genetics and the morphological evolution of maize. In *The maize handbook* (pp. 66-77). Springer New York.

Duvick, D. N., Janick, J., & Simon, J. E. (1990). Genetic enhancement and plant breeding. In *Advances in new crops. Proceedings of the first national symposium's new crops: research, development, economics'*, Indianapolis, Indiana, USA, 23-26 October 1988. (pp. 90-96). Timber Press.

Eagles, H. A., & Lothrop, J. E. (1994). Highland maize from Central Mexico—its origin, characteristics, and use in breeding programs. *Crop Science*, 34(1), 11-19.

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PloS one*, 6(5), e19379.

FAO, 1999. Análisis del carácter multifuncional de la agricultura y la tierra. Documento preparado para la Conferencia FAO/Países Bajos sobre el carácter multidimensional de la agricultura y la tierra. Maastricht, Países Bajos. 12-17 de septiembre.

FAO. (2009). Base de Datos de Estadísticas Agropecuarias. (Disponible en: www.faostat.fao.org. Noviembre 2011).

Frankel, O. H., & Brown, A. H. D. (1984). Plant genetic resources today: a critical appraisal. *Crop genetic resources: conservation & evaluation*/edited by

JHW Holden and JT Williams pp. 249-257

Goodman, M. M., & Stuber, C. W. (1983). Races of maize. 6: Isozyme variation among races of maize in Bolivia (No. 84-050731. CIMMYT.).

Gupta, P. K., Rustgi, S., & Mir, R. R. (2008). Array-based high-throughput DNA markers for crop improvement. *Heredity*, 101(1), 5-18.

Hamblin, M. T., Buckler, E. S., & Jannink, J. L. (2011). Population genetics of genomics-based crop improvement methods. *Trends in Genetics*, 27(3), 98-106.

Harlan, J. R. (1992). *Crops and man*, Madison, WI. American Soc. Agronomy.

Joshi, S. P., Ranjekar, P. K., & Gupta, V. S. (1999). Molecular markers in plant genome analysis. *Curr Sci*, 77(2), 230-240.

Kircher, M., & Kelso, J. (2010). High-throughput DNA sequencing—concepts and limitations. *Bioessays*, 32(6), 524-536.

Krishnamurthy, L. y Sahagún C. J. (1991). Recursos fitogenéticos, su conservación para un desarrollo sostenible. Universidad Autónoma Chapingo. pp: 13-19.

Meffe G K, CR Carroll. (1994). Genetics conservation of diversity within species: In: *Principles of Conservation Biology*, Groom M.J., G.K. Meffe and Carroll (Eds). Sinauer Associates, Inc. Sunderland, Massachusetts, USA. pp: 43-178.

Mullis, K. B. F. F., Faloona, F., Scharf, S., Saiki, R. K., Horn, G. T., & Erlich, H. (1986, January). Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. In *Cold Spring Harbor symposia on quantitative biology* (Vol. 51, pp. 263-273). Cold Spring Harbor Laboratory Press.

Nielsen, R., Paul, J. S., Albrechtsen, A., & Song, Y. S. (2011). Genotype and SNP calling from next-generation sequencing data. *Nature Reviews Genetics*, 12(6), 443-451.

Ortega P. R. (1985). Recursos genéticos para el mejoramiento del maíz en México. Primera parte: Análisis general. *Boletín de intercambio técnico y científico de la SOMEFI*. Germen 3:19-36.

Ortega, P. R., Sánchez, G. J.J., Castillo, G., Hernandez, C. J. M. (1991). Estado actual de los estudios sobre maíces nativos de México. In: Avance en el Estudio de los Recursos Fitogenéticos de México. Ortega P R, G Palomino H, F Castillo G, V A González H, M Livera M (eds). SOMEFI. Chapingo, México. pp: 161-185.

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

Pritsch, C. (2001). El pre-mejoramiento y la utilización de los recursos. En: Estrategia en recursos fitogenéticos para los países del Cono Sur. Documentos en PROCISUR, 2001. Montevideo, Uruguay. Disponible en línea:

<http://www.fagro.edu.uy/~fitotecnica/docencia/materiales%20apoyo/Premejoramiento.pdf>

R-Core Team. (2015). R: A language and environment for statistical computing. Vienna, Austria. R Foundation for Statistical Computing. (<https://www.r-project.org/>. Consultado el 12 de marzo).

Rodríguez, B. M. (2003). Aspectos jurídicos internacionales del acceso a los recursos genéticos que componen la diversidad biológica. Tesis de Doctorado. Universidad de Perú. 25 p.

Salhuana, W. y Pollak, L. (2006). Latin American Maize Project (LAMP) and Germplasm Enhancement of Maize (GEM) project: generating useful breeding

germplasm. *Maydica* 51:339-355.

Sánchez, G. J. J., Goodman, M.M. & Rawlings, J.O. (1993). Appropriate characters for racial classification in maize. *Economic Botany* 47: 44-59.

Sánchez, G. J. J., Goodman, M.M. & Stuber, C.W. (2000). Isozymatic and morphological diversity in the races of maize of Mexico. *Economic Botany* 54: 43-59.

Sanchez, J. J., & Goodman, M. M. (1992). Relationships among the Mexican races of maize. *Economic Botany*, 46(1), 72-85.

Sánchez-González, J. J. (2011). Diversidad del maíz y el teocintle. Informe preparado para el proyecto: “Recopilación, generación, actualización y análisis de información acerca de la diversidad genética de maíces y sus parientes silvestres en México”. Comisión Nacional para el Conocimiento y Uso de la Biodiversidad. Manuscrito. (Disponible en: http://www.biodiversidad.gob.mx/genes/pdf/proyecto/Anexo9_Analisis_Especialistas/Jesus_Sanchez_2011.pdf. Consultado en mayo 2016.

Sanger, F., & Coulson, A. R. (1975). A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *Journal of molecular biology* 94(3):441-448.

Schadt, E. E., Turner, S., and Kasarskis, A. (2010). A window into third-generation sequencing. *Hum. Mol. Genet.* 19:227–240. doi: 10.1093/hmg/ddq416

Schlötterer C. (2004). The evolution of molecular markers – just a matter fashion? *Nature Reviews Genetics* 5: 63-69.

Schnable, P. S., Ware, D., Fulton, R. S., Stein, J. C., Wei, F., Pasternak, S., & Minx, P. (2009). The B73 maize genome: complexity, diversity, and dynamics. *Science* 326(5956):1112-1115.

Sharma, S., Upadhyaya, H. D., Varshney, R. K., & Gowda, C. L. L. (2013). Pre-breeding for diversification of primary gene pool and genetic enhancement of grain legumes. *Frontiers in Plant Science* 4:09

SIAP. (2011). Anuario Estadístico de la Producción Agrícola. Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación (Disponible en: www.siap.gob.mx, noviembre 2011).

Tenaillon, M. I., U'Ren, J., Tenaillon, O., & Gaut, B. S. (2004). Selection versus demography: a multilocus investigation of the domestication process in maize. *Molecular Biology and Evolution*, 21(7), 1214-1225.

Turrent, F. A. y Serratos, J. A. (2004). Contexto y antecedentes del maíz silvestre y el cultivado en México, Capítulo I, Comisión para la Cooperación Ambiental de América del Norte (CCA). Simposio sobre Maíz y Biodiversidad, Memorias, Oaxaca, México, 11 de marzo de 2004.

Weng, J., Xie, C., Hao, Z., Wang, J., Liu, C., Li, M., Li, X. 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): 8 p.

Wright, S. (1978). Vol. 4: Variability within and among natural populations. *Evolution and the genetics of populations: a treatise in four volumes*: University of Chicago Press.

Xue, Y., Warburton, M. L., Sawkins, M., Zhang, X., Setter, T., Xu, Y., & Zhang, X. (2013). Genome-wide association analysis for nine agronomic traits in maize under well-watered and water-stressed conditions. *Theoretical and applied genetics*, 126(10), 2587-2596.

Yue, B., Xue, W., Xiong, L., Yu, X., Luo, L., Cui, K., & Zhang, Q. (2006). Genetic basis of drought resistance at reproductive stage in rice: separation of drought tolerance from drought avoidance. *Genetics*, 172(2), 1213-1228.

VII. APÉNDICE

7.1. Listado de 290 colecciones nativas de maíz involucradas en el estudio, así como algunas características de identificación y forma de agrupación en los análisis.

Grupo	Identificador Seed of Discovery	Identificador Banco de Germoplasma de CIMMYT	Raza	Adaptación
1	SEEDDIV18492	ARGE307	TUXPEN	Subtropical
1	SEEDDIV12666	ARZM12082	CRISCO	Subtropical
1	SEEDDIV15257	ARZM12186	COMTRO	Subtropical
1	SEEDDIV8330	ARZM12193	CRISCO	Subtropical
1	SEEDDIV6065	ARZM12206	CRISCO	Tropical
1	SEEDDIV17201	ARZM12209	CRISCO	Tropical
1	SEEDDIV6804	ARZM13018	CRISCO	Tropical
1	SEEDDIV6821	ARZM13019	COMTRO	Tropical
1	SEEDDIV6812	ARZM13024	COMTRO	Tropical
1	SEEDDIV6723	ARZM13025	CRISCO	Tropical
1	SEEDDIV12626	ARZM13070	CRISCO	Tropical
1	SEEDDIV10606	ARZM13071	COMTRO	Tropical
1	SEEDDIV26806	ARZM13072	COMTRO	Tropical
1	SEEDDIV7132	ARZM13077	CRISCO	Tropical
1	SEEDDIV7135	ARZM13084	DENSDA	Tropical
1	SEEDDIV12652	ARZM13087	CRAMAN	Tropical
1	SEEDDIV20328	ARZM13089	DENAMA	Tropical
1	SEEDDIV8231	ARZM13113	CRISCO	Tropical
1	SEEDDIV7624	ARZM13114	COMTRO	Tropical
1	SEEDDIV7634	ARZM13132	CRISCO	Tropical
1	SEEDDIV7702	ARZM13136	CRISCO	Tropical
1	SEEDDIV7108	ARZM14056	CRISCO	Tropical
1	SEEDDIV9864	ARZM17013	CRISCO	Tropical
1	SEEDDIV7589	ARZM17020	CRISCO	Tropical
1	SEEDDIV7426	ARZM17025	AMARGO	Tropical
1	SEEDDIV15434	ARZM17026	CRISCO	Tropical
1	SEEDDIV7002	ARZM17029	CRISCO	Tropical

1	SEEDDIV9884	ARZM18015	CRISCO	Tropical
1	SEEDDIV5596	ARZM18033	CRISCO	Tropical
1	SEEDDIV21943	ARZM18043	CRISCO	Tropical
1	SEEDDIV15692	CHZM04011	CAMELI	Tropical
1	SEEDDIV22954	CHZM04021	CAMELI	Tropical
1	SEEDDIV7556	CHZM04023	DIDECA	Tropical
1	SEEDDIV7552	CHZM04024	CAMELI	Tropical
1	SEEDDIV15882	CHZM04028	CAMELI	Tropical
1	SEEDDIV10963	CHZM04029	CAMELI	Tropical
1	SEEDDIV15854	CHZM04030	CAMELI	Tropical
1	SEEDDIV7909	CHZM04044	CAMELI	Tropical
1	SEEDDIV20646	SONO162	ONAVEN	Tropical
2	SEEDDIV13059	ARZM12125	PISINGALLO	Subtropical
2	SEEDDIV6784	ARZM12228	PISINGALLO	Subtropical
2	SEEDDIV7126	ARZM13022	PISINGALLO	Tropical
2	SEEDDIV23470	ARZM13094	PISINGALLO	Tropical
2	SEEDDIV15242	ARZM13095	PISINGALLO	Tropical
2	SEEDDIV1238	ARZM17015	PISINGALLO	Tropical
2	SEEDDIV22083	ARZM17021	PISINGALLO	Tropical
2	SEEDDIV22082	ARZM17022	PISINGALLO	Tropical
2	SEEDDIV22436	ARZM17027	PISINGALLO	Tropical
3	SEEDDIV12178	ARZM12079	SOCORR	Subtropical
3	SEEDDIV15797	LIM.78		Tropical
3	SEEDDIV7093	SINA9	TABPER	Tropical
3	SEEDDIV15421	SNLP171	TUXPEN	Subtropical
3	SEEDDIV21902	SONO168	REVENT	Tropical
4	SEEDDIV7473	ARZM07068	AVMORO	Tropical
4	SEEDDIV6832	ARZM12236	DENBLA	Subtropical
4	SEEDDIV6769	ARZM12237	DENBLA	Subtropical
4	SEEDDIV6800	ARZM12239	COMTRO	Subtropical
4	SEEDDIV6780	ARZM12242	COMTRO	Subtropical
4	SEEDDIV10632	ARZM13111	COMTRO	Tropical
4	SEEDDIV12156	ARZM18035	CRISCO	Tropical
4	SEEDDIV1015	BRAZ2932		Tropical
4	SEEDDIV2405	BRAZ2933	PEROLA	Tropical
4	SEEDDIV6582	BRAZ2934	PADENT	Tropical
4	SEEDDIV2426	BRAZ2957		Tropical
5	SEEDDIV15317	ARZM12021	DENBLA	Subtropical
5	SEEDDIV7242	ARZM12080	DENBLA	Subtropical
5	SEEDDIV12251	ARZM12081	CALCHA	Subtropical

5	SEEDDIV6129	ARZM12084		Subtropical
5	SEEDDIV6738	ARZM12086	CALCHA	Subtropical
5	SEEDDIV6741	ARZM12117	COMTRO	Tropical
5	SEEDDIV6666	ARZM12127	COMTRO	Tropical
5	SEEDDIV6712	ARZM12185	DENBLA	Subtropical
5	SEEDDIV6668	ARZM12189		Subtropical
5	SEEDDIV17394	ARZM12196	DENBLA	Subtropical
5	SEEDDIV6713	ARZM12202	DENBLA	Tropical
5	SEEDDIV11778	ARZM12203	DENBLA	Tropical
5	SEEDDIV6837	ARZM12204	CALCHA	Tropical
5	SEEDDIV13931	ARZM12205	CALCHA	Tropical
5	SEEDDIV6838	ARZM12210	COMTRO	Tropical
5	SEEDDIV11732	ARZM12214	DENBLA	Subtropical
5	SEEDDIV12627	ARZM12215	DENBLA	Tropical
5	SEEDDIV6891	ARZM12216	DSDBMF	Subtropical
5	SEEDDIV7851	ARZM12217	DENBLA	Subtropical
5	SEEDDIV6727	ARZM12218	CALCHA	Subtropical
5	SEEDDIV6793	ARZM12223	DENBLA	Subtropical
5	SEEDDIV12302	ARZM12224	DENBLA	Tropical
5	SEEDDIV6760	ARZM12225	CHAUCH	Subtropical
5	SEEDDIV6559	ARZM12226	CHAUCH	Subtropical
5	SEEDDIV8004	ARZM12229	DENBLA	Tropical
5	SEEDDIV6701	ARZM12230	DENBLA	Subtropical
5	SEEDDIV22807	ARZM12231	DENBLA	Subtropical
5	SEEDDIV6806	ARZM12233	DSDBMF	Subtropical
5	SEEDDIV8179	ARZM12238	DSDBMF	Subtropical
5	SEEDDIV11688	ARZM12240	DENBLA	Subtropical
5	SEEDDIV6931	ARZM12241	DENBLA	Subtropical
5	SEEDDIV10612	ARZM13069	DENBLA	Tropical
5	SEEDDIV11734	ARZM13074	DENBLA	Tropical
5	SEEDDIV12694	ARZM13078	DENBLA	Tropical
5	SEEDDIV16857	ARZM13093	CRISCO	Tropical
5	SEEDDIV17498	ARZM13096	DENBLA	Tropical
5	SEEDDIV9866	ARZM13104	DENBLA	Tropical
5	SEEDDIV10683	ARZM13119	DENBLA	Tropical
5	SEEDDIV10678	ARZM13120	PERLA	Tropical
5	SEEDDIV7160	ARZM13127	CALCHA	Tropical
5	SEEDDIV10655	ARZM13134	DENBLA	Tropical
5	SEEDDIV10668	ARZM13135	CALCHA	Tropical
5	SEEDDIV6951	ARZM17028	CALCHA	Tropical

5	SEEDDIV10617	ARZM17033	MOROCH	Tropical
5	SEEDDIV22704	ARZM17034	DENSDB	Tropical
5	SEEDDIV6482	CHIH116	TABPER	Tropical
5	SEEDDIV6046	CHIH117	TABPER	Tropical
5	SEEDDIV6333	CHIH119	SNJUAN	Tropical
5	SEEDDIV3139	CHIH70	MAIZON	Subtropical
6	SEEDDIV16773	ARZM13133	CRISCO	Tropical
6	SEEDDIV12272	ARZM19005	BL8HIL	Subtropical
6	SEEDDIV22941	CHIH118	SNJUAN	Tropical
6	SEEDDIV6107	CHIH126	USPOP	Tropical
6	SEEDDIV24012	CHIH127	MDULCE	Tropical
6	SEEDDIV5793	CHIH264		Subtropical
6	SEEDDIV19248	CHIH329		Subtropical
6	SEEDDIV23560	CHIH339		Subtropical
6	SEEDDIV25582	CHIH342		Subtropical
6	SEEDDIV12044	CHIH345		Subtropical
6	SEEDDIV12076	CHIH348		Subtropical
6	SEEDDIV973	CHIH38	APACHI9	Subtropical
6	SEEDDIV8143	CHIH53		Subtropical
6	SEEDDIV17211	DURA245	BOLITA	Subtropical
6	SEEDDIV11123	DURA246	CHALQU	Subtropical
6	SEEDDIV8215	SNLP168		Subtropical
6	SEEDDIV14468	SNLP3		Subtropical
6	SEEDDIV21531	SONO147	LADYFI	Tropical
6	SEEDDIV22087	SONO4	TUXPEN	Tropical
7	SEEDDIV11413	SINA176		Tropical
7	SEEDDIV11328	SINA177		Tropical
7	SEEDDIV4892	SINA7	HARINO8	Tropical
7	SEEDDIV15746	SINA75	BLANSO	Tropical
7	SEEDDIV20144	SINA79	DULNOR	Tropical
7	SEEDDIV7087	SINA8	TABPER	Tropical
7	SEEDDIV20150	SINA80	BLANSO	Tropical
7	SEEDDIV22422	SONO142	TABLON	Tropical
7	SEEDDIV20316	SONO143	REVENT	Tropical
7	SEEDDIV22533	SONO144	DULNOR	Tropical
7	SEEDDIV22592	SONO145	ONAVEN	Tropical
7	SEEDDIV16334	SONO146	TABPER	Tropical
7	SEEDDIV21863	SONO148	ONAVEN	Tropical
7	SEEDDIV21884	SONO150	ONAVEN	Tropical
7	SEEDDIV21433	SONO151	ONAVEN	Tropical

7	SEEDDIV21400	SONO152	ONAVEN	Tropical
7	SEEDDIV22531	SONO153	ONAVEN	Tropical
7	SEEDDIV22399	SONO154	ONAVEN	Tropical
7	SEEDDIV21843	SONO155	ONAVEN	Tropical
7	SEEDDIV20682	SONO156	TUXPEN	Tropical
7	SEEDDIV22471	SONO157	TABLON	Tropical
7	SEEDDIV21861	SONO158	ONAVEN	Tropical
7	SEEDDIV22945	SONO159	DULNOR	Tropical
7	SEEDDIV24281	SONO161	TABPER	Tropical
7	SEEDDIV22423	SONO163	TABPER	Tropical
7	SEEDDIV20436	SONO164	DULNOR	Tropical
7	SEEDDIV19991	SONO165	TABPER	Tropical
7	SEEDDIV21867	SONO166	ONAVEN	Tropical
7	SEEDDIV20657	SONO167	ONAVEN	Tropical
7	SEEDDIV20326	SONO169	ONAVEN	Tropical
7	SEEDDIV22508	SONO170	ONAVEN	Tropical
7	SEEDDIV4526	SONO177		Tropical
7	SEEDDIV5165	SONO21		Tropical
7	SEEDDIV1552	SONO23	REVENT	Tropical
7	SEEDDIV5184	SONO24	ONAVEN	Tropical
7	SEEDDIV4339	SONO26	NALTEL	Tropical
7	SEEDDIV1553	SONO27	REVENT	Tropical
7	SEEDDIV2615	SONO28		Tropical
7	SEEDDIV783	SONO29	BLANSO	Tropical
7	SEEDDIV4292	SONO31		Tropical
7	SEEDDIV19732	SONO34		Tropical
7	SEEDDIV19419	SONO40		Tropical
7	SEEDDIV1199	SONO75	TABPER	Tropical
8	SEEDDIV6603	ARZM12083	COMTRO	Subtropical
8	SEEDDIV16768	ARZM12207	PISING	Tropical
8	SEEDDIV16756	ARZM13091	COMTRO	Tropical
8	SEEDDIV1808	BLANCODEJUNIO	BLJUNI	Tropical
8	SEEDDIV2100	CHIH17		Subtropical
8	SEEDDIV1467	CHIH19	CELAYA	Subtropical
8	SEEDDIV785	CHIH20	CELAYA	Subtropical
8	SEEDDIV15312	CHIH262		Subtropical
8	SEEDDIV5592	CHIH265		Subtropical
8	SEEDDIV21462	CHIH31		Subtropical
8	SEEDDIV21207	CHIH330		Subtropical
8	SEEDDIV19380	CHIH331		Subtropical

8	SEEDDIV5654	CHIH333		Subtropical
8	SEEDDIV5877	CHIH334		Subtropical
8	SEEDDIV5648	CHIH335		Subtropical
8	SEEDDIV5645	CHIH336		Subtropical
8	SEEDDIV5602	CHIH337		Subtropical
8	SEEDDIV10325	CHIH338		Subtropical
8	SEEDDIV19427	CHIH343		Subtropical
8	SEEDDIV11458	CHIH345		Subtropical
8	SEEDDIV2617	CHIH35	TUXPEN	Subtropical
8	SEEDDIV25423	CHIH50	TUXNOR	Subtropical
8	SEEDDIV1474	CHIH55	CELAYA	Subtropical
8	SEEDDIV8193	CHIH56	TUXNOR	Subtropical
8	SEEDDIV13353	CHIH59		Subtropical
8	SEEDDIV13350	CHIH63	TUXNOR	Subtropical
8	SEEDDIV8241	CHIH64	TUXNOR	Subtropical
8	SEEDDIV1473	CHIH68	CELAYA	Subtropical
8	SEEDDIV3140	CHIH71		Subtropical
8	SEEDDIV1475	CHIH73	TUXPEN	Subtropical
8	SEEDDIV22010	CHIH84	TUXNOR	Subtropical
8	SEEDDIV12439	CHIH85	TUXNOR	Subtropical
8	SEEDDIV22298	CHIH86	TUXNOR	Subtropical
8	SEEDDIV12098	CHIH87	TUXNOR	Subtropical
8	SEEDDIV9222	CHIH89	TUXNOR	Subtropical
8	SEEDDIV22541	COAH109		Tropical
8	SEEDDIV17923	COAH11	TUXPEN	Tropical
8	SEEDDIV17199	COAH110		Tropical
8	SEEDDIV11352	COAH111		Tropical
8	SEEDDIV8071	COAH113		Tropical
8	SEEDDIV10362	COAH116		Tropical
8	SEEDDIV16432	COAH117		Subtropical
8	SEEDDIV13498	COAH119		Tropical
8	SEEDDIV8726	COAH123		Tropical
8	SEEDDIV1005	COAH15	TUXPEN	Tropical
8	SEEDDIV1123	COAH16	TUXPEN	Subtropical
8	SEEDDIV26354	COAH18		Subtropical
8	SEEDDIV1497	COAH19	TUXPEN	Subtropical
8	SEEDDIV19617	COAH20	TUXPEN	Subtropical
8	SEEDDIV4971	COAH21	TUXNOR	Subtropical
8	SEEDDIV12122	COAH52	TUXNOR	Tropical
8	SEEDDIV1498	COAH53	TUXPEN	Tropical

8	SEEDDIV18314	COAH54	TUXPEN	Tropical
8	SEEDDIV17796	COAH55	TUXPEN	Tropical
8	SEEDDIV4327	COAH56	TUXPEN	Tropical
8	SEEDDIV19545	COAH58	TUXPEN	Tropical
8	SEEDDIV3109	COAH59	TUXPEN	Tropical
8	SEEDDIV836	COAH60		Tropical
8	SEEDDIV18416	COAH61	TUXPEN	Tropical
8	SEEDDIV17046	COAH62		Tropical
8	SEEDDIV18481	COAH63	TUXPEN	Tropical
8	SEEDDIV5152	COAH64	TUXPEN	Tropical
8	SEEDDIV1500	COAH65	CELAYA	Tropical
8	SEEDDIV11548	COAH66		Tropical
8	SEEDDIV12176	COAH69	TUXNOR	Tropical
8	SEEDDIV19631	COAH72	TUXNOR	Subtropical
8	SEEDDIV23472	COAH76	TUXNOR	Subtropical
8	SEEDDIV23503	COAH78	TUXNOR	Subtropical
8	SEEDDIV19017	COAH8		Tropical
8	SEEDDIV3110	COAH81	TUXPEN	Tropical
8	SEEDDIV8351	COAH84		Tropical
8	SEEDDIV4216	COAH9	TUXPEN	Tropical
8	SEEDDIV11542	COAH90		Tropical
8	SEEDDIV19576	COAH91		Tropical
8	SEEDDIV12396	COAH97		Tropical
8	SEEDDIV1117	NVOL10	TUXPEN	Tropical
8	SEEDDIV5326	NVOL17	TUXPEN	Tropical
8	SEEDDIV12304	NVOL19	TUXPEN	Subtropical
8	SEEDDIV17894	NVOL22	TUXPEN	Tropical
8	SEEDDIV3209	NVOL23	TUXPEN	Tropical
8	SEEDDIV19501	NVOL45		Tropical
8	SEEDDIV15720	NVOL46		Subtropical
8	SEEDDIV5271	NVOL5	TUXPEN	Tropical
8	SEEDDIV8906	NVOL51		Tropical
8	SEEDDIV26355	NVOL52		Tropical
8	SEEDDIV3205	NVOL6	TUXPEN	Tropical
8	SEEDDIV5929	NVOL66		Tropical
8	SEEDDIV5727	NVOL67		Tropical
8	SEEDDIV5322	NVOL7	TUXPEN	Tropical
8	SEEDDIV13455	NVOL88		Tropical
8	SEEDDIV12574	PIUR173		Tropical
8	SEEDDIV7098	SINA10	TABPER	Tropical

8	SEEDDIV1197	SINA74	TUXPEN	Tropical
8	SEEDDIV11930	SINA76	TUXPEN	Tropical
8	SEEDDIV19838	SINA77	TUXPEN	Tropical
8	SEEDDIV21984	SINA78	SNJUAN	Tropical
8	SEEDDIV20131	SINA81	TUXPEN	Tropical
8	SEEDDIV19834	SINA82	TUXPEN	Tropical
8	SEEDDIV9741	SNLP165		Subtropical
8	SEEDDIV18858	SNLP166		Subtropical
8	SEEDDIV11764	SNLP167		Subtropical
8	SEEDDIV9566	SNLP169		Subtropical
8	SEEDDIV3052	SNLP17	CELAYA	Subtropical
8	SEEDDIV9121	SNLP170		Subtropical
8	SEEDDIV11957	SNLP179		Subtropical
8	SEEDDIV9091	SNLP187		Subtropical
8	SEEDDIV20654	SONO149	ONAVEN	Tropical
8	SEEDDIV3233	SONO20	BLANSO	Tropical
8	SEEDDIV740	SONO25	TUXPEN	Tropical
8	SEEDDIV5181	SONO7		Tropical
8	SEEDDIV16347	SONO76	TABLON	Tropical
8	SEEDDIV22030	TAMA118		Subtropical
8	SEEDDIV8084	TAMA119		Subtropical
8	SEEDDIV21061	TAMA120		Subtropical
8	SEEDDIV2688	TAMA46		Tropical