



**Universitat de les Illes Balears**

**DEPARTAMENT DE BIOLOGIA**

**GRUP DE RECERCA EN BIOLOGIA DE LES PLANTES EN  
CONDICIONS MEDITERRÀNIES**

**ESTUDIOS DE DIVERSIDAD GENÉTICA EN  
POBLACIONES DE MAÍZ (*Zea mays* L.) EVALUADAS CON  
MICROSATÉLITES**

**TESIS DOCTORAL**

***CLAUDIA A. BEDOYA SALAZAR***

***2012***



MARILYN L. WARBURTON, Investigadora Contratada Doctor en Genética del Departamento de Agricultura de los Estados Unidos – Servicio de Investigación Agrícola ARS, Unidad de Investigaciones de Resistencia Genética del Maíz CHPRRU y Profesor Adjunto de los Departamentos de Bioquímica y Ciencias de las Plantas y el Suelo de la Universidad Estatal de Missisipi

CERTIFICA:

Que Claudia A. Bedoya Salazar, ha realizado bajo su dirección, en los laboratorios del Centro Internacional de Mejoramiento de Maíz y Trigo (CIMMYT), el trabajo que, para optar al TÍTULO universitario de DOCTOR por la Universidad de las Islas Baleares dentro del Programa de Biología de las Plantas, presenta con el título:

ESTUDIOS DE DIVERSIDAD GENÉTICA EN POBLACIONES MAÍZ (*Zea mays L.*)  
EVALUADAS CON MICROSATÉLITES

Considerando concluida la presente memoria, autorizo a su presentación a fin de que pueda ser juzgada por el tribunal correspondiente

Para que así conste, firmo el presente certificado en Starkville, Missisipi a 5 de Marzo de 2012

Marilyn L. Warburton

Claudia A. Bedoya Salazar



## AGRADECIMIENTOS

Primero que todo quiero agradecer especialmente a la Dra. Marilyn Warburton, no solamente por animarme a empezar un PhD antes de que ella dejara el CIMMYT, también por su supervisión, consejo y apoyo en todos los niveles a través de estos años, mil gracias Marilyn. Gracias al Dr. José Luis Araus quien junto con la Dra. Marilyn hicieron posible realizar este doctorado fuera de mi país natal y fuera de mi lugar de residencia en México aprovechando mi trabajo en CIMMYT, idea que para mi era casi inconcebible.

Mi gratitud también es para el CIMMYT, la institución para la que he trabajado por varios años, y aunque no fue fácil balancear las responsabilidades y las colaboraciones con otros proyectos, me permitió en el último año concentrarme en terminar la escritura de este trabajo de Tesis. El trabajar para esta institución y estar involucrada en estudios de diversidad genética de maíz me dieron la oportunidad de conocer muchos investigadores de gran valor científico y personal que siempre tuvieron las puertas abiertas de sus oficinas, su tiempo y disposición para la discusión y asesoramiento. Particularmente, quiero agradecer a la Dra. Susanne Dreisigaker con quien tuve la oportunidad de discutir varios aspectos importantes de esta tesis; al Dr. Raman Babu por su asesoría en análisis de mapeo por asociación y por enseñarme a utilizar GenStat; gracias también al Dr. Suketoshi Taba, a quien le profeso un profundo respeto, el Dr. Taba fue Director del Banco de Germoplasma de maíz del CIMMYT hasta el año pasado, le agradezco por sus regaños, comentarios, e ideas, y por infundirme el valor de que la investigación que hacemos pertenece a una gran cadena de valor que debe tener su impacto final en el agricultor. Gracias también para el Dr. José Crossa por su asesoría en el uso del Índice de Selección, a la Dra. Jill Cairns, por su asesoría cuando empezó a desarrollarse la segunda parte de este trabajo doctoral. No podrían faltar mis agradecimientos para el Dr. Jorge Franco por sus correcciones y comentarios sobre la parte estadística de esta Tesis.

Muchas son las personas que me han ayudado y apoyado en los aspectos técnicos de la realización de esta Tesis, gracias a Hugo López mi compañero en el Laboratorio de Diversidad Genética y mi amigo, quien me apoyó técnicamente cada vez que lo requería.

Gracias a Víctor Chávez y Marcial Rivas del Banco de Germoplasma del CIMMYT, quienes siempre me brindaron rápida colaboración en los requerimientos de semilla, así como la información de pasaporte de las accesiones, Víctor, gracias también por compartir información relacionada con las razas de maíz que no está escrita ni publicada pero que representa el trabajo y la experiencia del día a día con los materiales nativos de maíz. Gracias a los chicos del CRIL (Crop Research Informatics Laboratory) por su paciencia y tiempo en enseñarme a utilizar el Fieldbook, a Víctor Hernández y Juan Arista del GIS (Geographic Information System) por su soporte en la elaboración de los mapas aquí presentados. Gracias especiales al grupo de Fisiología del Programa de Global de Maíz: Ciro, Sotero, y su equipo de la Estación Experimental de Tlaltizapán por su ayuda en la caracterización fenotípica de los materiales tropicales, espero que de alguna manera los resultados presentados en esta Tesis expresen su gran experiencia adquirida por muchos años de trabajo.

Gracias al Generation Challenge Programme que financió la primera parte de este trabajo de investigación, porque así como yo, otros estudiantes de doctorado se han beneficiado de sus talleres y del apoyo que brindan a la investigación en diversidad genética de diferentes cultivos. La segunda parte de esta Tesis en su primera fase fue financiada por El Fondo Regional de Tecnología Agropecuaria (FONTAGRO) y su continuación por la empresa privada AGROBio-México, especialmente gracias a Fabrice Salamanca quien en su condición de Director Ejecutivo de esta empresa me invitó a exponer el proyecto con el Consejo Consultivo obteniendo de esta manera después el financiamiento necesario para poder continuar con mi proyecto.

En Mallorca, gracias a Isabel Oliver y al Dr. Maurici Mus 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. Agradecimientos también para chicos como Mónica e Igor, Magdalena, Alicia y Cristina que siempre hicieron mi estadía en la Universidad y en la isla más fácil y agradable. Gracias a mis tíos Nena y Jorge que me acogieron como si fuera su hija en su departamento de Barcelona mientras tomaba los cursos, fue súper sentirme en familia, tener su apoyo y todo su cariño.

Mis agradecimientos finales, son algo especial para abordar, mil gracias al Dr. Jean-Marcel Ribaut, para mi ha sido un lujo y un privilegio que revisara esta Tesis, gracias por leerla como si fuera de un estudiante suyo, por todo el tiempo invertido, y por insistir en el conocimiento integrado y aplicado de la genética, la fisiología y el mejoramiento. Ahora, Especialmente gracias por el amor y la gran paciencia, por aguantar mis malos ratos y por animarme a seguir adelante. Gracias por insistirme tanto en que siempre hay que preguntarse: ¿por qué? y en remarcar que hay que ir “más allá”, porque esto es lo que hace la diferencia.





**ESTUDIOS DE DIVERSIDAD GENÉTICA EN POBLACIONES DE MAÍZ (*Zea mays*  
L.) EVALUADAS CON MICROSATÉLITES**

**CONTENIDO**

<b>ABREVIATURAS Y SÍMBOLOS</b> .....	v
<b>RESUMEN</b> .....	viii
<b>1. INTRODUCCIÓN GENERAL</b>	
1.1 Maíz en Latinoamérica .....	1
1.2 Utilización de los recursos genéticos de maíz .....	3
1.3 Análisis de la diversidad genética, marcadores moleculares y consideraciones estadísticas .....	6
1.3.1 Caracterización de la variabilidad genética .....	6
1.3.2 Marcadores Moleculares .....	7
1.3.3 Microsatélites .....	8
1.3.4 Herramientas Estadísticas.....	11
1.3.4.1 Descripción de la diversidad genética .....	11
1.3.4.2 Métodos Multivariados.....	13
1.4 Caracterización genotípica de maíz.....	15
1.5 Importancia de los estudios de diversidad genética de maíz.....	18
<b>2. OBJETIVOS</b> .....	20
<b>3. CARACTERIZACIÓN GENOTÍPICA DE POBLACIONES DE MAÍZ EN LATINOAMÉRICA</b> .....	22
3.1 Introducción.....	22
3.1.1 Maíz en Latinoamérica y culturas precolombinas.....	23
3.2 Materiales y métodos.....	30
3.2.1 Caracterización Genotípica.....	33
3.3.2 Cálculos estadísticos.....	37

3.3 Resultados.....	40
3.3.1 Validación de los marcadores seleccionados con las poblaciones de maíz mexicanas.....	40
3.3.2 Asociaciones entre las 194 poblaciones de maíz latinoamericanas.....	50
3.3.3 Análisis de estructura poblacional en Latinoamérica.....	53
3.3.4 Análisis de estructura subpoblacional en Latinoamérica.....	57
3.3.5 Comparaciones genéticas.....	66
3.3.5.1 Análisis de distancias genéticas.....	67
3.3.5.2 Análisis de diversidad genética.....	68
3.4 Discusión.....	71
3.4.1 Variabilidad de maíz en México.....	71
3.4.2 Variabilidad de maíz en Latinoamérica.....	73
3.4.3 Maíz: historia, arqueología, antropología y asociaciones genéticas en México.....	77
3.4.4 Dispersión del maíz por el litoral, tierra y mar.....	78
3.4.5 Relaciones genéticas y contexto histórico del maíz en Latinoamérica.....	81
3.5 Conclusiones.....	87
<b>4. CARACTERIZACIÓN FENOTÍPICA Y GENÉTICA DE UN PANEL DE MAÍCES TROPICALES PRECOCES DE INTERÉS PARA ESTUDIOS DE ASOCIACIÓN DE TOLERANCIA A LA SEQUÍA.....</b>	<b>89</b>
4.1 Introducción.....	89
4.1.1 Mejora analítica y rendimiento de grano.....	92
4.1.2 Tolerancia a la sequia y análisis de mapeo por asociación.....	94
4.1.3 Consideraciones para un estudio de mapeo por asociación.....	99
4.2 Materiales y Métodos.....	103
4.2.1 Germoplasma.....	103
4.2.2 Condiciones experimentales.....	104
4.2.2.1 Régimen de regadío.....	104
4.2.2.2 Caracteres agronómicos y floración.....	108

4.2.2.3 Crecimiento y senescencia de las plantas .....	109
4.2.3 Caracterización genotípica .....	110
4.2.4 Análisis de datos genotípicos .....	111
4.2.5 Análisis de datos fenotípicos .....	112
4.3 Resultados.....	114
4.3.1 Rendimiento promedio de los diferentes tipos de accesiones .....	114
4.3.2 Clasificación de las accesiones según Índice de Selección evaluado bajo condiciones de estrés .....	116
4.3.3 Caracterización y desempeño agronómico de las accesiones bajo estrés y condiciones normales .....	117
4.3.4 Heredabilidad.....	123
4.3.5 Diversidad genética .....	124
4.3.6 Cambios de frecuencia alélica .....	126
4.3.7 Análisis de estructura poblacional.....	130
4.3.8 Análisis descriptivo del rendimiento basado en el análisis de estructura poblacional.....	137
4.3.9 Efectos en la variabilidad fenotípica debida a la estructura de población .....	139
4.4 Discusión .....	142
4.4.1 Método de Selección del material: el poder del Índice de Selección ....	143
4.4.2 Diversidad genética en el panel inicial, panel seleccionado y poblaciones nativas.....	145
4.4.3 QTLs para sequía y cambios de frecuencia alélica e identificación de <i>loci</i> con potencial agronómico.....	147
4.4.4 Heredabilidad.....	154
4.4.5 Efectos de la estructura de población .....	156
4.4.6 Observaciones carácter-panel de asociación .....	159
4.5 Conclusiones.....	161
<b>5. PERSPECTIVAS.....</b>	<b>164</b>
<b>6. REFERENCIAS .....</b>	<b>170</b>
<b>7. ANEXOS .....</b>	<b>184</b>

<b>7.1</b>	Datos de pasaporte para las 194 poblaciones de maíz Latinoamericanas .....	184
<b>7.2</b>	Aislamiento del ADN genómico .....	191
<b>7.3</b>	Lista de marcadores moleculares tipo SSR utilizados.....	192
<b>7.4</b>	Reacción de PCR.....	194
<b>7.5</b>	Gráficas del Ln de la probabilidad del análisis de estructura poblacional .....	195
<b>7.6</b>	Cálculo $\Delta K$ basado en Evanno et al. 2005.....	196
<b>7.7</b>	Códigos de raza asignados a las poblaciones de maíz mexicanas.....	197
<b>7.8</b>	Matriz de distancias genéticas entre poblaciones mexicanas basada en la distancia de proporción de alelos compartidos.....	198
<b>7.9</b>	Frecuencias alélicas y cambios de frecuencia alélica entre materiales de pre-mejoramiento del Banco de germoplasma, material mejorado del Programa de Fisiología del CIMMYT y poblaciones nativas .....	199

## ABREVIATURAS Y SÍMBOLOS

ABA: Abscisic Acid (Ácido abscísico)

a.C.: antes de Cristo

ACP: Análisis de Componentes Principales

ACoP: Análisis de Coordenadas Principales

ADN: Ácido Desoxirribonucleico

AM: Association Mapping (Mapeo por Asociación)

AMOVA: Analysis of Molecular Variance (Análisis de Varianza Molecular)

AP: Antes del Presente

ASI: Anthesis-Silking Interval (Intervalo entre la floración masculina y floración femenina)

BC: Backcross (Retrocruza)

CE: Condiciones de Estrés

CH-a: Chapalote y afines

CN: Condiciones Normales

CIMMYT: Centro Internacional de Mejoramiento de Maíz y Trigo

CZ: Complejo Zapalote

d.C.: despues de Cristo

DDS: Días después de la siembra

DTPW: Drought Tolerance Population (Poblaciones tolerantes a la sequía de maíces dentados blancos)

DTPY: Drought Tolerance Population Yellow (Poblaciones tolerantes a la sequía de maíces dentados amarillos)

EH: Enrollamiento de Hoja

ESIM: Eigen Analysis Selection Index (Índice de Selección basado en el Eigen análisis)

GBS: Genotyping by Sequencing (Genotipificación por Secuenciación)

GCP: Generation Challenge Programme (Programa de Reto Generación)

G x E: Gene–Environment Interaction (Interacción Genotipo-Ambiente)

GEM: Germplasm Enhancement of Maize Project (Proyecto de mejoramiento del germoplasma de maíz)

GWAS: Genome Wide Association (Asociación a través de todo el genoma)

H: Heterocigocidad

H<sup>2</sup>: Heredabilidad

HI: Harvest Index (Índice de Cosecha)

INIFAP: Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias de México

INTA: Instituto Nacional de Tecnología Agropecuaria de Argentina

INRA: Institut National de la Recherche Agronomique (Instituto Nacional de Investigación Agronómica de Francia)

IS: Índice de Selección

ITS: Índice de Tolerancia a la Sequía

K: Número de grupos que representaría la estructura poblacional

$\Delta K$ : Razón de cambio en el logaritmo de la probabilidad de los datos entre valores sucesivos de K

LAMP: Latin American Maize Project (Proyecto de maíz para Latinoamérica)

LD: Linkage Disequilibrium (Desequilibrio de ligamiento)

LOcc: Ladera Occidental de la Sierra Madre Occidental

LOri: Ladera Oriental de la Sierra Madre Oriental

NJ: Neighbor-joining Method (Método del vecino mas cercano)

MABC: Marker Assisted Backcross (Retrocruzas asistidas por marcadores)

MPP: Mazorcas por planta

MARS: Marker Assisted Recurrent Selection (Selección recurrente asistida por marcadores)

MAS: Marker Assisted Selection (Selección asistida por marcadores)

MAYG: Mapping As You Go (Mapeo en el camino)

NAM: Nested Association Mapping (Mapeo de asociación anidado)

NCGRP: National Center for Genetic Resources Preservation (Centro Nacional de Preservación de los Recursos Genéticos)

OPV: Open Pollinated Variety (Variedad de polinización abierta)

P: Valores Fenotípicos

PCR: Polymerase Chain Reaction (Reacción en cadena de la Polimerasa)

PIC: Polymorphism Information Content (Contenido de información polimórfica)

Q: Matriz de ancestro

QTL: Quantitative trait loci (*loci* de un carácter cuantitativo)

RAPDs: Random Amplification of Polymorphic DNA (Polimorfismo amplificado aleatoriamente)

RFLPs: Restriction Fragment Length Polimorphism (Polimorfismos de longitud de fragmentos de restricción)

RILs: Recombinant inbred lines (Lineas recombinantes de maíz)

Sen: Senescencia

S: Sur

STS: Sequence Tagged Sites (Secuencias de sitios marcados)

SSRs: Simple Sequence Repeats (Secuencias repetidas cortas)

SMO: Sierra Madre Occidental

SNPs: Single Nucleotide Polymorphisms (Polimorfismos de base única)

UPGMA: Unweighted Pair Group Method with Arithmetic Mean (Método de agrupamiento no ponderado por pares utilizando la media aritmética)

VAC: Valles Altos Centrales

## RESUMEN

En este trabajo de Tesis se evaluó la diversidad genética presente en los maíces nativos de Latinoamérica y su contextualización en dos diferentes aspectos. El primero en un estudio donde la diversidad genética y la estructura poblacional de los materiales permitieron ligar aspectos históricos, antropológicos y arqueológicos del maíz. En el segundo aspecto se explora el potencial de las razas de maíz que ligada a la evaluación fenotípica con poblaciones de mejoramiento permitieron identificar fuentes de variación alélica importante para algunos caracteres de interés con fines agronómicos. El primer aspecto es presentado como: “Caracterización genotípica de poblaciones de maíz en Latinoamérica”. Se describe la diversidad genética y la estructura poblacional de un germoplasma de maíz representativo de Latinoamérica formado por 194 accesiones de poblaciones nativas en el cual están representadas 131 razas de 23 países. El germoplasma fue genéticamente caracterizado por un grupo de 28 microsatélites en muestras conjuntas de ADN. De esta manera, se deduce que el maíz se dispersó fuera de su área de origen y domesticación en Mesoamérica hacia el Caribe y Sudamérica, y esta dispersión ha estado indudablemente asociada con migraciones humanas tempranas y con movimientos poscolombinos. Los resultados revelaron que eventos prehistóricos e históricos como expansiones de las culturas precolombinas, como dos diferentes patrones de poblamiento del Caribe, las tecnologías marítimas desarrolladas por las culturas precolombinas, y la llegada de los españoles a América pueden ser asociadas con el cultivo del maíz. El segundo aspecto presentado como: “Caracterización fenotípica y genética de un panel de maíces tropicales precoces de interés para estudios de asociación de tolerancia a la sequía” describe la caracterización fenotípica bajo dos regímenes de riego, normal y de estrés hídrico para un panel inicial formado por tres diferentes clases de germoplasma: *pooles* o complejos genéticos desarrollados por el Banco del Germoplasma del CIMMYT, OPVs del Subprograma de Fisiología y poblaciones nativas, donde los valores altos de heredabilidad para los caracteres evaluados demostraron la calidad de los datos fenotípicos. La caracterización genotípica se basó en el análisis de diversidad genética como: índices de diversidad, número de alelos encontrados, alelos únicos dentro de cada clase de germoplasma, índices de diferenciación de Wright y análisis de estructura poblacional utilizando un conjunto de 30 SSRs. Para la



caracterización fenotípica se utilizó un Índice de Selección (IS) tomando en cuenta la evaluación fenotípica del germoplasma en condiciones de estrés, se calculó el índice simple de tolerancia a la sequía (ITS), y se consideró la precocidad del material para diferenciar tolerancia de escape a la sequía. Se identificaron accesiones nativas con gran potencial como fuente de alelos favorables para la expresión de caracteres como tolerancia a la sequía, sincronía floral y contenido de clorofila. Se analizó en detalle los cambios de frecuencia alélica, entre y dentro de las clases de materiales, y se identificaron cambios importantes en *loci* relacionados con resistencia a la sequía y altas temperaturas. Los análisis de estructura poblacional del germoplasma, combinados con los datos fenotípicos permitieron identificar grupos de accesiones con alelos específicos interesantes para vigor, sequía, escape de la sequía, y caracteres secundarios. Basados en estos resultados, se sugiere que el panel inicial aquí estudiado podría utilizarse para un análisis de mapeo por asociación, para identificar *loci* y alelos favorables de manera más sistemática, y utilizando herramientas genéticas para mapeo de alta densidad para diferentes caracteres de interés como componentes del rendimiento y caracteres secundarios de manera independiente.



# 1. INTRODUCCIÓN GENERAL

## 1.1 Maíz en Latinoamérica

En Latinoamérica, el maíz es uno de los principales recursos alimenticios para consumo humano, además de tener una gran importancia socioeconómica y cultural. Durante siglos las actividades de los agricultores, desde las primeras sociedades agrícolas hasta las actuales han permitido la selección y el desarrollo de nuevas variedades de maíz, que se han adaptado a las diversas condiciones ecológicas, agronómicas, de clima y de suelo, cubriendo los requerimientos culturales y nutricionales. Las sociedades precolombinas en todo el territorio de América Latina consideraban que el maíz fue un regalo de los dioses a los hombres y que cultivarlo era sagrado. Según el Popol Vuh (Cardoza-Aragón, 2006), libro sagrado de los mayas, el ser humano fue hecho de maíz luego del fracaso que los dioses tuvieron al probar otros materiales. El maíz acompañó el desarrollo de culturas tan importantes como la olmeca, maya, teotihuacana y la mexica entre otras, y sigue siendo el principal sustento de las culturas rurales indias y mestizas (Barros y Buenrostro, 1997). Estos maíces reciben varios nombres, tales como variedades primitivas, variedades de los agricultores, variedades locales, o criollos y han sido mantenidos y mejorados *in situ* por los agricultores, basados en la percepción de sus necesidades y su experiencia.

El cultivo y la selección de semillas son tradiciones antiguas gracias a las cuales se han generado numerosas variedades locales (las razas) del maíz. Pandey y Gardner (1992) mencionan que de 285 razas de maíz descritas en el hemisferio occidental, 265 están presentes en Latinoamérica, la mayoría de ellas en América del Sur. Más recientemente Dowsell et al. (1996), concluyen que en todo el mundo existen descritas e identificadas cerca de 300 razas de maíz, (involucrando a miles de cultivares diferentes) y que representaban del 90 al 95% de la diversidad genética de la especie. Paterniani y Goodman (1977) estimaron que alrededor del 50% de las razas de maíz en América están adaptadas a los trópicos, 40% a las tierras altas y 10% a los ambientes intermedios. También estimaron que cerca del 40% de las razas de maíz tienen endospermo harinoso, 30% son duros, 20% son dentados y 3% tienen tipos de granos dulces (Figura 1.1).

**Figura 1.1** Razas nativas de maíces latinoamericanos (Fotografía: Dr. Suketoshi Taba, CIMMYT)



A pesar que hoy día es posible apreciar en su justa medida el valioso recurso genético que representan las razas de maíz, la pérdida de biodiversidad en la especie a nivel global, principalmente durante el siglo XX y también en la actualidad, es un hecho (Bracco et al., 2009). Este problema es particularmente notorio en las explotaciones agrícolas comerciales, donde las razas tradicionales se han ido reemplazando por cultivos genéticamente mejorados a partir de una base genética reducida incluyendo a los cultivos genéticamente modificados.

Aunque en la actualidad los híbridos de maíz representan la parte económica más importante del cultivo en países como Argentina, Brasil, Estados Unidos, y Canadá, utilizan una parte limitada de la diversidad encontrada en variedades de polinización abierta, razas y parientes silvestres como Teocintle, (Fukunaga et al., 2005; Matsuoka et al., 2002b). Para el mantenimiento, conservación y aprovechamiento del potencial genético que representan las razas de maíz y sus parientes silvestres se han formado colecciones que

están depositadas en diferentes bancos internacionales de germoplasma como el del CIMMYT (Centro Internacional de Mejoramiento de Maíz y Trigo) y el NCGRP (National Center for Genetic Resources Preservation, por sus siglas en inglés) en Colorado USA (Taba, 2005). También se han creado bancos de germoplasma a nivel nacional en diferentes países latinoamericanos como el Banco de Germoplasma de Maíz del Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias (INIFAP) de México, el Banco Activo de Germoplasma del Instituto Nacional de Tecnología Agropecuaria (INTA) en Argentina, el Banco Nacional de Germoplasma en Colombia, entre otros (Taba, 2005).

Para el futuro aprovechamiento de estas colecciones preservadas en los bancos de germoplasma en programas de mejoramiento es de gran importancia su caracterización genotípica y fenotípica, a la vez que conocer adecuadamente las condiciones agroecológicas generales (topografía, clima y prácticas de cultivo) de las distintas regiones donde se ha recolectado el germoplasma de maíz. A partir de esta caracterización es posible comprender mejor la distribución y diversidad de las poblaciones de maíz, su migración relacionada con las poblaciones humanas y sus adaptaciones a distintas condiciones ambientales. Mirando los cambios genéticos en respuesta a diferentes tipos de ambientes se genera información muy valiosa para mejoramiento. Explorar la diversidad genética de especies cultivadas de importancia económica, como el maíz, es tarea fundamental en todo programa de investigación agrícola.

## **1.2 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: 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 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 (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). Las fuentes de germoplasma más útiles son los cultivares modernos y líneas avanzadas del mejoramiento, ya adaptados a un ambiente productivo específico, sin embargo, existen condiciones para hacer progreso genético en las cuales es recomendado utilizar germoplasma exótico, incluyendo todas las fuentes de germoplasma no adaptados como por ejemplo materiales presentes en las colecciones internacionales (ejemplo: Banco de Germoplasma de Maíz del CIMMYT), líneas provenientes de programas de mejoramiento, poblaciones nativas, especies y géneros emparentados.

Un ejemplo del gran potencial de los maíces nativos de México fue el descubrimiento de factores de resistencia a la infestación del insecto *Sitophilus zeamais*, una de las plagas más destructivas del maíz almacenado (Serratos et al., 1987). Otro ejemplo, es la formación de los complejos genéticos desarrollados en el CIMMYT, que han servido de base para desarrollar cientos de variedades productivas de maíz. Estos complejos están formados por distintas accesiones de razas con características comunes, material proveniente de programas de mejoramiento estadounidenses y/o de países en desarrollo, teniendo en cuenta la ecología a la cual se adaptan mejor (las tierras bajas tropicales, las tierras subtropicales de altitud media o las tierras altas), así como color y tipo de grano, entre otras

características. Los complejos genéticos son la base del programa de pre-mejoramiento y mejoramiento de maíz en el CIMMYT porque vinculan la enorme diversidad de las colecciones de semilla del banco de germoplasma con el programa de desarrollo de variedades mejoradas, que devuelven esta diversidad a los agricultores en una forma más productiva.

La identificación de germoplasma útil para fines agronómicos es un tema decisivo en el pre-mejoramiento, al respecto existen varias opiniones acerca de la abundancia y disponibilidad de variabilidad genética que se pueda explotar provechosamente, asociado a la idea que no toda la variabilidad genética es útil, o que a veces hay que “pagar” un precio al incorporar alelos exóticos debido a la interacción génica entre diferentes genes para una determinada característica. 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). Adicionalmente a las herramientas utilizadas en el pre-mejoramiento convencional nuevas herramientas biotecnológicas desarrolladas en los últimos años han permitido un gran avance en el estudio del origen, composición, estructura y funcionalidad de los genomas especialmente de los cultivos de interés agrícola como es el maíz. Estos avances en la caracterización de los genomas están permitiendo nuevos abordajes en la utilización y aprovechamiento de los recursos genéticos disponibles (Lanaud y Lebod, 1995). Algunas de estas herramientas son los marcadores moleculares con los que se pueden construir mapas genéticos y caracterizar material biológico a nivel molecular.

De esta manera, la re-valorización de los recursos genéticos disponibles en los bancos de germoplasma de maíz se ha visto enriquecida por los avances tecnológicos en la biología molecular. La utilización de marcadores moleculares ha resultado una herramienta valiosa, que ha permitido caracterizar la diversidad genética presente en las colecciones, complejos genéticos, etc. En este respecto, en el caso específico del CIMMYT, los investigadores de la institución y colaboradores han aumentado las actividades de pre-mejoramiento tratando de sacar provecho de las herramientas biotecnológicas. Por ejemplo, en la última década se realizaron varias evaluaciones de la diversidad genética entre y dentro de poblaciones de

maíz tropical, subtropical y templado con marcadores moleculares como los microsatélites (Reif, et al 2006), lo mismo para la estimación de la diversidad genética de razas, poblaciones mejoradas y líneas con el fin de examinar fuentes potenciales de diversidad alélica existente en estos materiales para su futuro uso programas de mejoramiento (Warburton et al., 2008), y en la caracterización genética de líneas, variedades de polinización abierta (OPVs) (Warburton., et al 2002) usando métodos de caracterización genética.

### **1.3 Análisis de la diversidad genética, marcadores moleculares y consideraciones estadísticas**

#### **1.3.1 Caracterización de la variabilidad genética**

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).

Estudios intrapoblacionales permiten evaluar la diversidad genética entre individuos de una misma accesión determinando, entre otras cosas, su potencial como nueva fuente de alelos para mejoramiento. Estos estudios también permiten entender la estructura genética del



maíz, trazar rutas de migración del maíz a través del tiempo, identificar a través de los cambios de frecuencia alélica la selección natural como resultado de la adaptación a un ambiente ecológico nuevo, o la selección dirigida como resultado de la preferencia de los agricultores (rendimiento, textura, color, etc.). Desde un punto de vista de mejoramiento, el último objetivo es determinar la fenología de estas fuentes diferentes de diversidad para identificar los alelos nuevos que pueden impactar de manera favorable los parámetros agronómicos.

El reciente desarrollo de la biología molecular y las nuevas herramientas estadísticas en genómica y poblaciones genéticas han permitido explorar las opciones para identificar y utilizar esta variación genética para la mejora de plantas (Aranguren-Méndez et al., 2005). Desde el punto de vista molecular, esta variación también denominada polimorfismo genético, hace referencia a la existencia en una población de múltiples alelos, es decir, un polimorfismo es una variación en la secuencia de un lugar determinado del ADN entre los individuos de una población y puede medirse a través de diferentes técnicas disponibles hoy en día. Estas diferencias en los polimorfismos permiten identificar las variaciones alélicas dentro de un gen.

Para efectuar estos estudios ha sido de gran importancia el desarrollo de la técnica de la reacción en cadena de la polimerasa (PCR) (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.

### **1.3.2 Marcadores moleculares**

En años recientes, diferentes sistemas de marcadores, tales como, Polimorfismos de Longitud de Fragmentos de Restricción (RFLPs), Polimorfismos Amplificados Aleatoriamente (RAPDs), Secuencias de sitios marcados (STS), Secuencias Repetidas Cortas (SSRs), Polimorfismos de base única (SNPs) y otros, han sido desarrollados y

aplicados a una variedad de especies de cultivos, incluyendo cereales, como es el caso del maíz (Joshi et al., 1988; Schlötterer, 2004). Los marcadores moleculares son útiles en el diseño de nuevos esquemas de mejoramiento y selección, como por ejemplo: la selección asistida por marcadores (Marker Assisted Selection, MAS) (Ribaut y Hoisington, 1998; Ribaut et al., 2002), en la creación de catálogos de genes funcionales, clonación de secuencias de ADN que codifican genes y la secuenciación completa de las especies de interés (Lee, 1998), también se han vuelto muy populares para análisis filogenéticos añadiendo una nueva dimensión a las teorías evolutivas (Joshi et al., 1988)

La caracterización molecular se está consolidando como uno de los métodos más apropiados para llevar a cabo estudios de identificación de genotipos y similitud genética en numerosas especies. Las nuevas tecnologías de marcadores moleculares permiten acelerar y optimizar los métodos actuales de caracterización lo que puede permitir detectar la variabilidad genética en un conjunto determinado de germoplasma de interés.

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), (Tabla 1.1).

Aunque es difícil encontrar un marcador molecular que se reúna todos los criterios anteriores, un sistema de marcadores se puede identificar si cumple al menos algunas de las características anteriores dependiendo del tipo de estudio a realizar. En este sentido, el uso de marcadores moleculares tipo microsatélite se ha generalizado en estudios de diversidad genética y de estructura genética de poblaciones por su herencia mendeliana, alta confiabilidad, repetitividad y automatización, y pueden detectar homocigosis y heterocigosis en organismos diploides como el maíz (Matsuoka et al., 2002a).

### **1.3.3 Microsatélites**

En detalle, los microsatélites (Simple Sequence Repeats, SSR) son secuencias de uno a diez o más pares de bases repetidas y adyacentes distribuidas en el genoma (Ciofi et al., 1988), que generalmente consisten en dinucleótidos, por ejemplo: (AC)<sub>n</sub>, (AG)<sub>n</sub>, (AT)<sub>n</sub>; trinucleótidos, por ejemplo: (TCT)<sub>n</sub>, (TTC)<sub>n</sub>; tetranucleótidos, por ejemplo: (TATG)<sub>n</sub>, donde n es el número de unidades repetidas dentro de un *locus* de microsatélite en un genotipo de referencia (Figura 1.2). También pueden presentar una forma más compleja, mezclando di, tri, o más nucleótidos en una patrón repetitivo, por ejemplo (GA)<sub>n</sub>(GAA)<sub>n</sub>.

El análisis con este tipo de marcadores involucra la detección de fragmentos específicos de ADN y se obtiene la medida de los alelos en pares de bases para cada una de las regiones. Presentan ventajas respecto a otros marcadores moleculares por su capacidad para analizar un único *locus* de posición conocida en el cromosoma y su distribución homogénea en el genoma (Matsuoka et al., 2002b; Schlotterer, 2004). La variabilidad en el número de estas secuencias repetidas permite la visualización de diferencias genéticas entre individuos. En este sentido, si se dispone de diferentes SSRs distribuidos uniformemente en el genoma, los polimorfismos de los SSRs causados por las diferencias en el número de unidades repetidas permiten discriminar de manera precisa entre individuos altamente emparentados.

**Figura 1.2.** Ejemplo de un microsatélite di-nucleótido A-C(n).



Otra aplicación de los microsatélites es la construcción de mapas de ligamiento más completos y detallados, con lo cual se hace la identificación de genes de interés; también permiten seleccionar en una población en segregación los individuos que presentan el máximo de alelos favorables en los genes de interés agronómico (Ribaut et al., 2002).

**Tabla 1.1** Comparación de diferentes tipos de marcadores moleculares\*

<b>Tipo de Marcador</b>	<b>Acrónimo</b>	<b>Ventajas</b>	<b>Desventajas</b>
Polimorfismo de base única	SNPs	Baja tasa de mutación	Costosos de aislar y caracterizar Bajo contenido de información para un único SNP
		Abundantes en el genoma	
		Facilidad de caracterización Nuevos enfoques analíticos se están desarrollando en el presente Repetitividad entre laboratorios Bases de datos existentes y accesibles	
Secuencias Repetidas Cortas - Microsatélites	SSR	Altamente informativos	Alta tasa de mutación
		Fácil de aislar	Comportamiento complejo de mutación
		Interpretación sencilla de los resultados	Comparación entre laboratorios requiere preparación
Tecnología de matrices de diversidad	DArT	Tecnología de alto rendimiento	Debe estar disponible la información sobre la secuencia de los genes Costosa Tiene requisitos técnicos considerables La cantidad y el tipo de los datos producidos requiere de un alto nivel de pericia en computación y de un equipo avanzado
		Exploración del genoma entero	
		Resultados altamente reproducibles	
Secuenciación de ADN		El contenido de información es máximo	Todavía significativamente más costoso que las otras técnicas
		Fácil comparación entre estudios	Presenta exigencias técnicas

\* Basada en información de Schlötterer (2004), Aranguren-Méndez et al. (2005) y Diversity Arrays, (2011)

En el área de genética poblacional los microsatélites están siendo ampliamente utilizados porque permiten estimar los niveles de variabilidad genética dentro de poblaciones y analizar las relaciones genéticas existentes entre las mismas (Aranguren-Méndez et al., 2005). Los SSRs permiten la obtención de datos poblacionales y el cálculo de frecuencias alélicas. De esta manera se pueden estimar las distancias genéticas entre poblaciones o

entre individuos (Matsuoka et al., 2002a), así como realizar análisis filogenéticos (Fukunaga et al., 2005; Vigourox et al., 2008) y de estructura de población (Reif et al., 2004; Warburton et al., 2008).

### **1.3.4 Herramientas estadísticas**

#### **1.3.4.1 Descripción de la diversidad genética**

El estudio de la diversidad genética es el proceso por el cual la variación entre individuos, grupos de individuos o poblaciones es analizada por un método específico o por una combinación de métodos (Mohammadi y Prasanna, 2003). Se puede definir, como diversidad genética a la multiplicidad de frecuencias alélicas presentes en un grupo de individuos, que junto con el ambiente provee la naturaleza del fenotipo y es base fundamental en programas de mejoramiento (Cardona, 2010). Aunque el uso de marcadores moleculares permite una confiable diferenciación de genotipos, y la información que proveen es de diferentes tipos, la selección de los métodos de análisis depende de los objetivos del experimento, del nivel de resolución requerida, de los recursos tecnológicos e infraestructura disponible (Mohammadi y Prasanna, 2003).

Para analizar esta variabilidad genética existe una variedad de estadísticas que resumen la información en términos manejables. Los estadísticos más empleados y que se describen a continuación son: porcentaje de *loci* polimórficos, el número medio de alelos por *locus*, y la heterocigosidad esperada y observada:

#### **Porcentaje de *loci* polimórficos:**

Un *locus* se considera polimórfico si detecta más de un alelo en una población. Generalmente el criterio más utilizado es el 1%: un *locus* será polimórfico si el alelo más común presenta una frecuencia menor al 99% en la población bajo estudio.

#### **Número medio de alelos por *locus*:**

Esta estadística indica el número medio de alelos que presenta un *locus* en una población. Esta medida requiere un número mínimo de individuos analizados, ya que cuando el número es grande, mayor es la probabilidad de detectar alelos suplementarios en los *loci*.

**Heterocigosidad (H):**

Representa una mejor medida de la diversidad genética. La heterocigosidad se puede medir como heterocigosidad observada ( $H_O$ ) y/o heterocigosidad esperada ( $H_E$ ).  $H_O$  se define como la frecuencia relativa de individuos heterocigotos observados en la muestra para cualquiera de los *loci*, y es influenciada por el número de alelos y la frecuencia alélica. Mientras que la  $H_E$ , desde el punto de vista matemático, es la probabilidad de que dos alelos tomados al azar de la población sean diferentes (Crow y Kimura, 1970).

**Índice de Contenido Polimórfico (PIC):**

Este índice evalúa cuan informativo es un marcador en la población de acuerdo a las frecuencias de los alelos y oscila entre 0 (no informativo) y 1 (muy informativo), este índice se basa en el número de alelos y las frecuencias alélicas.

Varios enfoques se han propuesto para estimar la cantidad de diferenciación genética entre poblaciones y en subdivisiones de una población en base a los supuestos relacionados con el modelo evolutivo y las consecuentes modificaciones en los algoritmos (Wright, 1951; Nei, 1973). En este sentido, Wright (1951) describe Estadísticas-F o Índices de Fijación, como las correlaciones entre las frecuencias de los alelos relativos a una subpoblación y a la población total respectivamente ( $F_{IS}$  y  $F_{IT}$ ), y las correlaciones entre las frecuencias de dos alelos tomados al azar de cada subpoblación ( $F_{ST}$ ), indicando la pérdida de heterocigotos por la subdivisión en subpoblaciones (S) respecto a la población total (T).  $F_{ST}$  suele utilizarse como medida de diferenciación genética entre subpoblaciones o entre poblaciones. En general, los Índices de Fijación miden la reducción de heterocigosis esperada con apareamiento al azar a cualquier nivel de una jerarquía poblacional. De otro lado, mediante un Análisis de Varianza Molecular (AMOVA) se puede evaluar la estructuración o distribución de la diversidad genética dentro y entre grupos de razas, poblaciones, aplicándose también a diferentes niveles jerárquicos, (Excoffier et al., 1992).

Dado que la distancia genética puede definirse como la diferencia entre dos entidades que puede ser descrita por la variación alélica (Nei, 1973), o más ampliamente como "una medida cuantitativa de diferenciación genética ya sea a nivel de secuencia o de frecuencias

alélicas que se calcula entre individuos, poblaciones o especies" (Beaumont et al., 1998), las distancias genéticas pueden ayudar a entender las relaciones evolutivas entre poblaciones y permiten obtener información para caracterización de razas (Nagamine y Higuchi, 2001). La distancia genética entre dos genotipos, poblaciones o individuos puede calcularse por varias medidas estadísticas y depende de la naturaleza del conjunto de datos (Mohammadi y Prasanna, 2003), en este respecto, en la literatura disponible se encuentran varias discusiones relacionadas con la selección y uso de las medidas de distancias genéticas (Beaumont et al., 1998; Mohammadi y Prasanna, 2003).

#### **1.3.4.2 Métodos multivariados**

Los marcadores moleculares proveen datos multidimensionales que generalmente son complementados con otros tipos de información, por ejemplo datos morfológicos o fisiológicos. Así, los estudios sobre variabilidad genética están frecuentemente asociados a la medición de muchos caracteres en una misma entidad biológica. En este aspecto, son de especial interés los métodos multivariados altamente útiles para analizar la diversidad entre entidades caracterizadas por múltiples variables que permiten resumir la información en pocas variables sintéticas informativas de la variabilidad total. Algunos de estos métodos son: análisis de conglomerados o agrupamiento, Análisis de Componentes Principales (ACP) y/o Análisis de Coordenadas Principales (ACoP), y análisis de estructura de población.

El propósito principal de un análisis de agrupamiento es agrupar individuos o entidades en función de las características que poseen, individuos con descriptores similares son matemáticamente agrupados dentro de un mismo grupo (Hair et al., 1995). Con los datos obtenidos de caracterizaciones morfológicas, bioquímicas, o con un conjunto de marcadores moleculares estos análisis permiten la asignación de un individuo a un grupo según su composición genética determinada (Mohammadi y Prasanna, 2003; Aranguren-Méndez et al., 2005). En términos generales existen dos tipos de métodos de agrupación, los métodos basados en distancia y métodos en que las observaciones de cada grupo se supone que son aleatorias de algún modelo paramétrico, las inferencias sobre los parámetros correspondientes a cada grupo y la pertenencia de cada individuo a un grupo

especifico se llevan a cabo conjuntamente con métodos estadísticos como máxima-verosimilitud o métodos Bayesianos (Mohammadi y Prasanna).

En el método de agrupamiento basado en distancias se requiere definir una distancia entre el individuo y los individuos de los diferentes grupos posibles en el estudio, obteniendo así una matriz de distancias genéticas que puede representarse gráficamente según el algoritmo de agrupamiento seleccionado (Johnson y Wichern, 1992). Algunos algoritmos de agrupamiento son: el método de agrupamiento no ponderado por pares utilizando la media aritmética (UPGMA), método sugerido cuando se suponen en el modelo de análisis diferentes tasas de evolución (Eding y Laval, 1999; Takezaki y Nei, 1996); el método del vecino más cercano (Neighbor Joining, NJ) que se caracteriza por construir árboles mediante sucesivos agrupamientos de alineaciones tomando en cuenta las distancias entre un par de individuos o grupos y las distancias entre estos y los restantes individuos o grupos simultáneamente donde el resultado es un árbol no rotado (sin raíz u origen evolutivo) (Saitou y Nei, 1987). Es así como los dendogramas son las representaciones graficas más utilizadas de una matriz de distancias y bajo ciertas circunstancias esta representación puede ser tomada como una estimación de la filogenia (Aranguren-Méndez et al., 2005). Estos métodos son dinámicos y el hecho de asignar más individuos a los grupos puede influir en la composición y número de grupos que resultan del estudio de diversidad genética. Por ejemplo, añadir un individuo puede generar el agrupamiento de 2 grupos cercanos o dividir un grupo en varios grupos nuevos.

Los Análisis de Componentes y Coordenadas Principales (ACP y ACoP) son métodos de reducción de datos para clarificar las relaciones entre dos o más caracteres y dividir el total de la varianza de los caracteres originales en un número limitado de nuevas variables correlacionadas (Wiley, 1981; Mohammadi y Prasanna, 2003). Este análisis permite la visualización de las diferencias entre individuos e identificar posibles grupos, así como la determinación del número óptimo de grupos en un estudio. La reducción de variables es lograda mediante una transformación lineal de las variables originales en un nuevo conjunto de variables correlacionadas llamadas componentes principales (CP). Tanto el ACP como el ACoP son métodos de análisis de datos que presentan los resultados en forma de diagramas de dispersión, la diferencia entre ellos se basa en que un ACP parte de una



matriz completa de datos, como presencia versus ausencia de alelos en datos de marcadores moleculares, mientras que en ACoP se parte de una matriz de distancias entre un conjunto de entidades (Mohammadi y Prasanna, 2003). En un conjunto de datos que posee gran cantidad de datos perdidos y cuando hay menos entidades que caracteres se recomienda un ACoP (Rohlf, 1972).

Un método de agrupamiento basado en la estadística bayesiana es el descrito por Pritchard et al. (2000) en la inferencia de la estructura de población utilizando datos genotípicos multilocus de marcadores no ligados. La fortaleza de este enfoque se basa en un análisis efectivo de la estructura poblacional, la agrupación y la asignación de individuos en poblaciones, poblaciones a grupos, etc. Esta metodología permite estimar la proporción del genoma de un individuo que contribuye a una subpoblación específica, denominada "matriz de origen genético" o "matriz de proporción de ancestro" ( $Q$ ). Thornsberry et al. (2001) proporcionó la primera demostración empírica de la utilidad de la inferencia de la estructura poblacional en estudios de asociación genética, específicamente en la identificación de un gen asociado con la variación del tiempo de floración en maíz.

#### **1.4 Caracterización genotípica de maíz**

Aunque las relaciones genéticas entre poblaciones de maíz pueden ser medidas por varios métodos, uno de los métodos tempranos fue la medida de características cromosómicas. Esas características cromosómicas, en este caso fueron la posición de nódulos en forma de nudo como componentes distintivos, visibles al microscopio, que pueden o no estar presentes en una posición definida en cada uno de los cromosomas del maíz, y que son caracteres fijos y heredables. Estas características permitieron medir las relaciones genéticas, lo que reflejó en el maíz una mezcla de germoplasmas muy distintos, que alguna vez estuvieron aislados geográficamente unos de otros (McClintock et al., 1981). Estudios más recientes utilizando isoenzimas (Sanchez et al., 2006) y microsatélites han verificado esas observaciones (Matsuoka et al., 2002a; Reif et al., 2006; Vigourox et al., 2008).

En la última década se han llevado a cabo estudios con microsatélites en maíz con diferentes finalidades, por ejemplo, Reif et al. (2006) caracterizaron razas mexicanas de maíz y estudiaron las relaciones asumidas por estudios morfológicos; Pressoir y Berthaud.

(2004) utilizaron microsatélites para estudiar la estructura genética de razas de maíz en el Valle de Oaxaca– México, y Dubreuil et al., (2006) analizaron las frecuencias alélicas en muestras conjuntas (*bulks*) de ADN (como el presente estudio) para esclarecer el origen de las razas de maíz Europeas. Más recientemente Warburton et al., (2008) utilizaron esta clase de marcadores para examinar los niveles de diversidad y estructura poblacional en razas de maíz mexicanas, variedades de polinización abierta y líneas mejoradas. También se han empleado este tipo de marcadores en muestras de maíz arqueológicas para evaluar su afiliación genética con las razas existentes en el sur de Sudamérica (Oliveira-Freitas et al., 2003; Lia et al., 2009; Bracco et al., 2009). Incluso, los microsatélites también han sido utilizados en estudios en relación al origen y domesticación del maíz (Matsuoka et al., 2002b; Vigourox et al., 2008).

Los microsatélites también han ayudado a dilucidar que el origen de la gran diversidad encontrada en el maíz tiene que ver con la incidencia de diferentes factores. En primera instancia, la mayoría de la evidencia apunta a un solo evento de domesticación del maíz a partir de un teocintle anual (*Zea mays* ssp. *parviglumis*) (Matsuoka et al., 2002b); a su vez, el teocintle presenta niveles más altos de diversidad genética que el maíz domesticado (Fukunaga et al., 2005; Warburton et al., 2011). Adicionalmente, factores involucrados en la evolución del maíz después del proceso de domesticación fueron probablemente un número alto de mutaciones que generó nuevos alelos, mientras que la recombinación generaría nuevas combinaciones de éstos (Reif et al., 2006). Presumiblemente, el flujo génico con el teocintle incrementó la base genética existente del maíz (Doebley, 2004), especialmente, en este campo, la evidencia actual sugiere que teocintles de las regiones Balsas y Chalco (ssp. *parviglumis* y ssp. *mexicana*) han contribuido en los últimos 200 años y siguen contribuyendo como fuentes importantes de alelos para el maíz domesticado (Warburton et al., 2011). Mientras que la deriva genética, la selección natural y la humana dirigida (realizada por los primeros agricultores) modelaría la variación genética existente y que eventualmente resultó en un gran número de razas adaptadas a condiciones ambientales específicas y con caracteres deseados por los agricultores tempranos (Reif et al., 2006). Suplementariamente, la acción de sistemas de transposones -elementos genéticos movibles- dentro del genoma (Buckler et al., 2006), así como también, diferentes modalidades y

prácticas en el intercambio de semilla efectuado por los pequeños agricultores ha hecho este proceso de generación de diversidad genética más activo y cambiante. Es así como durante siglos cada grupo étnico ha seleccionado y desarrollado nuevas variedades de maíz que se han adaptado a las diversas condiciones de clima, suelo, ecológicas y agronómicas, cubriendo los diversos requerimientos culturales y nutricionales. De esta manera, cada variedad local representa una fuente de diversidad genética. Estas variedades son poseedoras de características que son importantes para la adaptación local, la estabilidad económica y la sostenibilidad del agricultor. Asimismo, las variedades de toda una región pueden contener niveles de diversidad genética comparables a los de las colecciones depositadas en muchos bancos de germoplasma (Taba, 2005).

La caracterización genotípica de las especies de polinización abierta como es el caso del maíz, ha sido muy difícil en el pasado porque la variabilidad encontrada refleja más a la variación intrapoblacional que entre poblaciones, y estos niveles de variación pueden ser muy altos (Doebley et al., 1984; Dubreuil et al., 2006; Warburton et al., 2008). Esto significa que para describir bien toda la variación dentro de una población, y para poder diferenciar entre diferentes poblaciones, al menos 15 individuos deben ser caracterizados para representar suficientemente la diversidad alélica presente en una población de maíz (Dubreuil et al., 2006). Esto hace que los estudios existentes sean escasos y con frecuencia realizados con un número insuficiente de individuos, lo que dificulta sacar conclusiones fiables.

Un método de análisis con microsatélites en muestras conjuntas de individuos de la misma población de maíz fue desarrollado por INRA (el Instituto Nacional de Investigación Agrícola de Francia) y el CIMMYT (Dubreuil et al., 2006). Al menos cincuenta marcadores SSR fueron optimizados en el CIMMYT para ser usados en la estimación de la frecuencia alélica dentro de una muestra conjunta de 15 individuos por población, usando la altura máxima extraída de un secuenciador de ADN automático y adaptando el software para el análisis de estos datos. Este método es mucho más eficiente en cuestión de tiempo y costo que la caracterización genotípica de múltiples individuos, uno por uno, por población, y por supuesto mucho más preciso que la caracterización genotípica de sólo un individuo por población. El ahorro de tiempo, reactivos y dinero es significativo, estimándose el costo en

una décima parte de lo que sería necesario en un estudio genotípico analizando individualmente el mismo número de individuos por población. Por estas razones en esta Tesis Doctoral hemos utilizado esta misma metodología para caracterizar las poblaciones de maíces objeto de nuestro estudio.

### **1.5 Importancia de los estudios de diversidad genética de maíz**

Durante el siglo XX la pérdida de biodiversidad a nivel global fue un hecho, particularmente cultivos y razas tradicionales han estado bajo amenaza por el definitivo reemplazamiento por materiales mejorados (Bracco et al., 2009). Por esta razón el estudio de la diversidad presente en los materiales nativos de cada país ha despertado interés ya que es posible que estas razas estén en riesgo de extinción o erosión genética. El conocimiento genético de las razas de maíz es de interés no solo por el estudio de las fuerzas evolutivas en el proceso de domesticación y sus aplicaciones en el diseño de programas de mantenimiento y conservación de germoplasma de maíz, sino también en el aprovechamiento de la diversidad existente, pues solo una pequeña proporción de esta diversidad disponible en América es utilizada en los programas de mejoramiento (Bracco et al., 2009). La evaluación genotípica y fenotípica de estos materiales nativos podrían acelerar los programas de mejoramiento convencional e incrementar su eficiencia identificando material potencial para ser aprovechado desde el pre-mejoramiento detectando la presencia de alelos exóticos favorables y estimando las frecuencias alélicas con el fin de ampliar la base genética en los materiales mejorados. De otro lado, una comprensión de la diversidad presente en las poblaciones de maíz y las relaciones entre ellas es de interés no solo para tener una visión global de la variabilidad genética existente en Latinoamérica, sino también para tener un mejor entendimiento de la migración del maíz a partir de su centro de origen y las probables rutas que históricamente siguieron los grupos que han cultivado el maíz en Latinoamérica.

Lamentablemente los cambios ambientales asociados al cambio climático como el aumento de temperatura, pueden afectar negativamente la productividad del maíz en Latinoamérica y el mundo. El maíz es uno de los principales cultivos para los agricultores con bajos recursos especialmente de las zonas tropicales y subtropicales, y son los que más pueden sufrir las

consecuencias de las elevadas temperaturas ya que por lo general cultivan en condiciones de temporal o seco y carecen de otras alternativas de manejo como la posibilidad de aplicar riegos o rotaciones de cultivo o aplicación de abono. De otro lado, en los últimos años se ha aumentado la presión sobre los sistemas agrícolas para garantizar el abastecimiento de las múltiples demandas de maíz. El futuro cercano del maíz no podría estar más necesitado de avances en conocimientos integrados genéticos, fisiológicos, interacción genotipo-ambiente, que permitan lograr estrategias de mejoramiento y/o manejo de cultivo que garantice una adecuada productividad. Teniendo en cuenta la importancia de ampliar el conocimiento sobre diversidad genética de los maíces nativos en su contexto histórico e identificar fuentes de germoplasma potenciales para ser utilizados en la mejora de maíz, esta Tesis Doctoral aborda dos temas principales: un estudio de la variabilidad genética de maíz existente en Latinoamérica “Caracterización genotípica de poblaciones de maíz en Latinoamérica” y si esta variabilidad lleva aparejada adaptaciones a tolerancia de estreses abióticos “Caracterización fenotípica y genética de un panel de maíces tropicales precoces de interés por su tolerancia al a sequía”.

## 2. OBJETIVOS

En la presente Tesis Doctoral se ha estudiado la diversidad genética presente en poblaciones nativas procedentes de Latinoamérica, complejos genéticos de mejoramiento desarrollados por el Banco de Germoplasma y materiales mejorados por el Programa de Fisiología del CIMMYT. Los principales aspectos relacionados con este estudio son la estimación de la diversidad genética en los diversos materiales, el análisis de estructura poblacional, y el estudio del contexto histórico que ha estado relacionado con el cultivo del maíz, así como la evaluación en campo de poblaciones promisorias por la buena adaptabilidad y respuesta bajo condiciones de estrés abiótico.

Las diferentes investigaciones que forman parte de esta tesis han tratado de alcanzar una serie de objetivos específicos, en particular, en el primer tema: “Caracterización genotípica de poblaciones de maíz en Latinoamérica” se ha enfatizado en los temas vinculados con la diversidad genética presente en las razas de maíz típicas de los países latinoamericanos y sus relaciones interpoblacionales tratando de ligar la información histórica, dataciones arqueológicas, e información antropológica disponible. Sus objetivos específicos han sido:

- 1) Conocer la diversidad genética de 194 poblaciones de maíces latinoamericanos por medio de la caracterización genotípica utilizando marcadores moleculares tipo microsatélite (SSR) en muestras conjuntas de ADN.
- 2) Validar el conjunto de marcadores microsatélite seleccionado, tomando como referencia las clasificaciones de las accesiones mexicanas de las cuales se tiene una gran variedad de información disponible.
- 3) En base al estudio de marcadores moleculares y el análisis de estructura poblacional aportar nueva información y profundizar en el conocimiento acerca de la migración del maíz desde Mesoamérica hacia América del Sur teniendo en cuenta el contexto histórico y cultural que ha revestido el maíz hasta nuestros días.

De de otro lado, en la segunda parte de esta Tesis Doctoral, “Caracterización fenotípica y genética de un panel de maíces tropicales precoces de interés para estudios de asociación de

tolerancia a la sequía” se ha profundizado en la caracterización genotípica y su adaptación fenotípica a la sequía en los materiales mejorados en el CIMMYT junto con 56 poblaciones nativas provenientes de varios países latinoamericanos. En esta segunda parte de la tesis los objetivos específicos fueron:

- 1) Obtener una distribución fenotípica basada en el cálculo del Índice de Selección de un panel inicial de 70 accesiones de maíz tropical evaluadas bajo condiciones de estrés hídrico y riego normal.
- 2) Seleccionar un conjunto de poblaciones, que representen un 50% del panel inicial, con comportamiento fenotípico contrastado frente a la sequía (es decir que representen las mejores y peores accesiones con respuesta adaptativa bajo condiciones de estrés) en base a un Índice de Selección y la fenología (evaluada como floración masculina).
- 3) Evaluar la diversidad genética en el panel inicial y en el panel seleccionado de los extremos de la distribución fenotípica.
- 4) Determinar la subestructura poblacional del panel inicial y panel seleccionado.
- 5) Demostrar el potencial del panel inicial para un estudio de mapeo por asociación para la identificación de nuevos alelos favorables para fines agronómicos en diferentes regímenes hídricos.

### **3. CARACTERIZACIÓN GENOTÍPICA DE POBLACIONES DE MAÍZ EN LATINOAMÉRICA**

#### **3.1 Introducción**

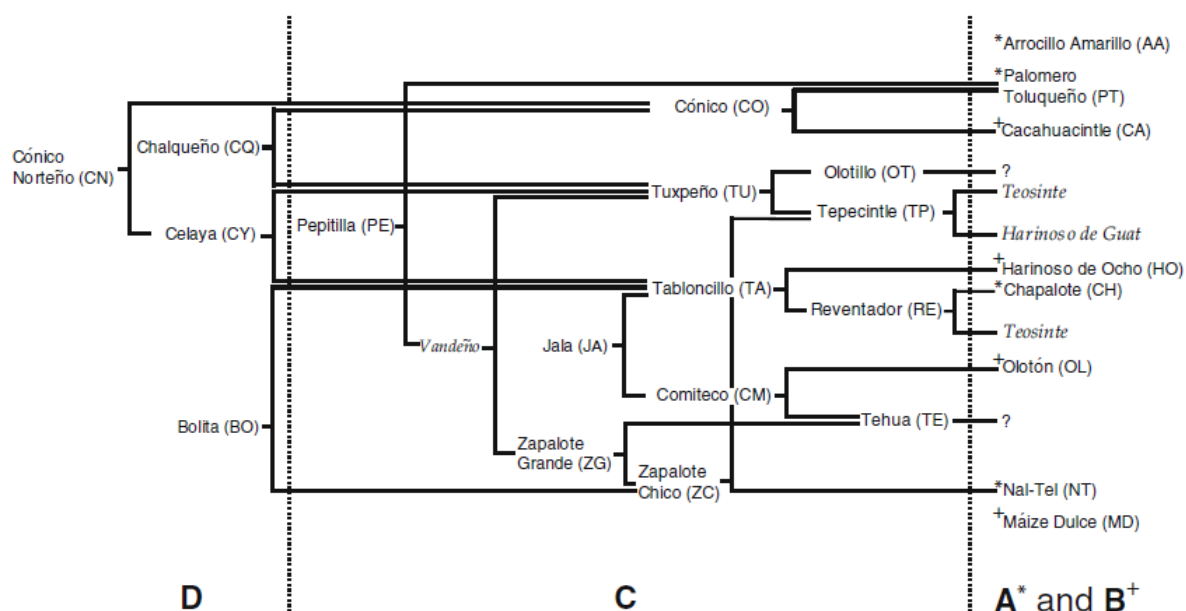
En los primeros ensayos para la clasificación del maíz se utilizó la definición de raza que propusieron Anderson y Cutler (1942): “Un grupo de individuos emparentados, con suficientes características en común para permitir su reconocimiento como grupo”. Esta definición fue la base para los primeros estudios de la diversidad del maíz en América con la que se clasifican los grupos o razas de las distintas regiones del continente. Los primeros trabajos de clasificación del maíz se fundamentaron en la descripción de estas razas sobre bases morfológicas, fisiológicas, genéticas, isoenzimáticas, agronómicas y características citogenéticas (nudos cromosómicos) que permitieron establecer patrones de relaciones genealógicas preliminares (Kato et al., 2009).

En forma particular, el trabajo que Wellhausen y colaboradores (1951) realizaron en México es considerado el primer estudio organizado y sistemático de la diversidad de maíz. En este trabajo se clasifican e ilustran 25 razas mexicanas en cuatro grupos, tres sub razas y siete tipos como no bien definidas. Los grupos definidos son: 1) Razas Indígenas Antiguas, consideradas como las descendientes de un maíz primitivo (Palomero Toluqueño, Arrocillo Amarillo, Chapalote y Nal-tel. 2) Razas Exóticas Precolombinas procedentes de Centro y Sudamérica (Cacahuacintle, Harinoso de Ocho, Olotón y Maíz Dulce. 3) Mestizas Prehistóricas que son el resultado del cruzamiento de las anteriores y su introgresión con el teocintle (Cónico, Reventador, Tabloncillo, Tehua, Tepecintle, Dzit-Bacal, Zapalote Chico, Zapalote Grande, Tuxpeño, Pepitilla, Comiteco, Jala, Olotillo y Vandéño. 4) Razas Modernas Incipientes, que se generarían en la época de la conquista (Bolita, Chalqueño, Celaya y Cónico Norteño). De igual forma que en México, en el resto de Latinoamérica se encuentran variedades y razas de maíz que provienen de diferentes épocas, procesos de selección y consecuentemente existe una gran diferencia genética entre las mismas que por cientos de años, desde épocas precolombinas y coloniales, han sido mantenidas por los agricultores/conservadores.



En estudios más recientes como el llevado a cabo por Reif y colaboradores (2006), se examinaron las relaciones asumidas previamente por la caracterización morfológica en el estudio de Wellhausen y colaboradores (1951) pero utilizando marcadores moleculares tipo microsatélite (como en el presente estudio) corroborando las relaciones propuestas para las razas Cónico Norteño, Chalqueño, Celaya, y Bolita ( Figura 3.1).

**Figura 3.1** Relaciones de las razas Indígenas Antiguas (A), Exóticas Precolombinas (B), Mestizas Prehistóricas (C) y Modernas Incipientes (D) sugeridos por Wellhausen et al. (1951). (progenie –lado izquierdo-, progenitores –lado derecho). Razas en *itálicas* no están incluidas en dicho estudio (Reif et al., 2006).



### 3.1.1 Maíz en Latinoamérica y culturas precolombinas

La máxima expresión de la domesticación del maíz ha sido su difusión y subsecuente diversificación en numerosas variedades adaptadas a las condiciones ambientales particulares (Vigouroux et al., 2008), empezando en Mesoamérica desde las zonas templadas de los altiplanos hasta las tierras bajas tropicales, alcanzando Norteamérica, el Caribe y Sudamérica (Kato et al., 2009; McClung, 1997). En este sentido, este trabajo de tesis hipotetiza que el maíz puede ser uno de los pocos cultivos en el mundo, cuya variación morfológica y genética reflejaría la sucesión de importantes acontecimientos históricos,

culturales, y la influencia determinante de factores geográficos. Por esta razón se hace a continuación una descripción de algunas zonas geográficas y culturas pre-colombinas que estarían ligadas a la dispersión temprana del maíz.

En la actualidad, se entiende a Mesoamérica como una vasta área cultural pre-hispánica que discurre entre la zona norte-centro de México hasta la costa del Océano Pacífico en Costa Rica que abarcó cerca de 5.000 años de historia (Figura 3.2, Figura 3.3). Puede definirse como una región cuyo ambiente estuvo constituido por una compleja multiplicidad de culturas regionales, donde los avances en la tecnología agrícola y artesanal, así como la expansión de ciertas redes de ideología y poder político, fueron los motores que, en esencia, caracterizaron el devenir común de todos sus pueblos.

Tomando como referencia las características geomorfológicas, étnicas y lingüísticas de las culturas Mesoamericanas, así como, materiales que las identificaron en la antigüedad se pueden definir diversas sub-áreas culturales:

- Centro de México, se trata de una extensa región de altiplano situada a una altura promedio de 2.300 msnm. Se halla integrada por diferentes valles o mesetas de clima templado, sub árido y árido, entre ellos, el Valle o Cuenca de México, el Valle de Toluca, el Valle de Tula, el Valle de Morelos y el Valle de Puebla-Tlaxcala. La existencia de cinco lagos navegables y de una agricultura intensiva asociada a éstos propició que el Valle de México fuese un área neurálgica en el pasado prehispánico del México Central, siendo el lugar donde se desarrollaron culturas como la Teotihuacana, Tolteca y Mexica.

- Sur de México y parte del área Maya. Es una extensa región que muestra una gran variabilidad medioambiental formada por zonas de clima sub árido, valles templados, cuencas subtropicales y una franja costera tropical paralela al Océano Pacífico; durante la época prehispánica, las culturas zapoteca y mixteca fueron las más sobresalientes en Oaxaca, y en países como Guatemala, Belice, Honduras y El Salvador, así como en los estados mexicanos de Campeche, Yucatán y Quintana Roo asistieron al florecimiento de la compleja cultura Maya.

- Golfo de México, región integrada por los actuales estados mexicanos de Tamaulipas, Veracruz y Tabasco; en esta región se desarrollaron las culturas como la Olmeca y Totonaca.
- Occidente de México, comprendiendo los estados de Guerrero, Michoacán, Sinaloa, Nayarit, culturas como la Tarasca y Mezcala son las más sobresalientes de la región.

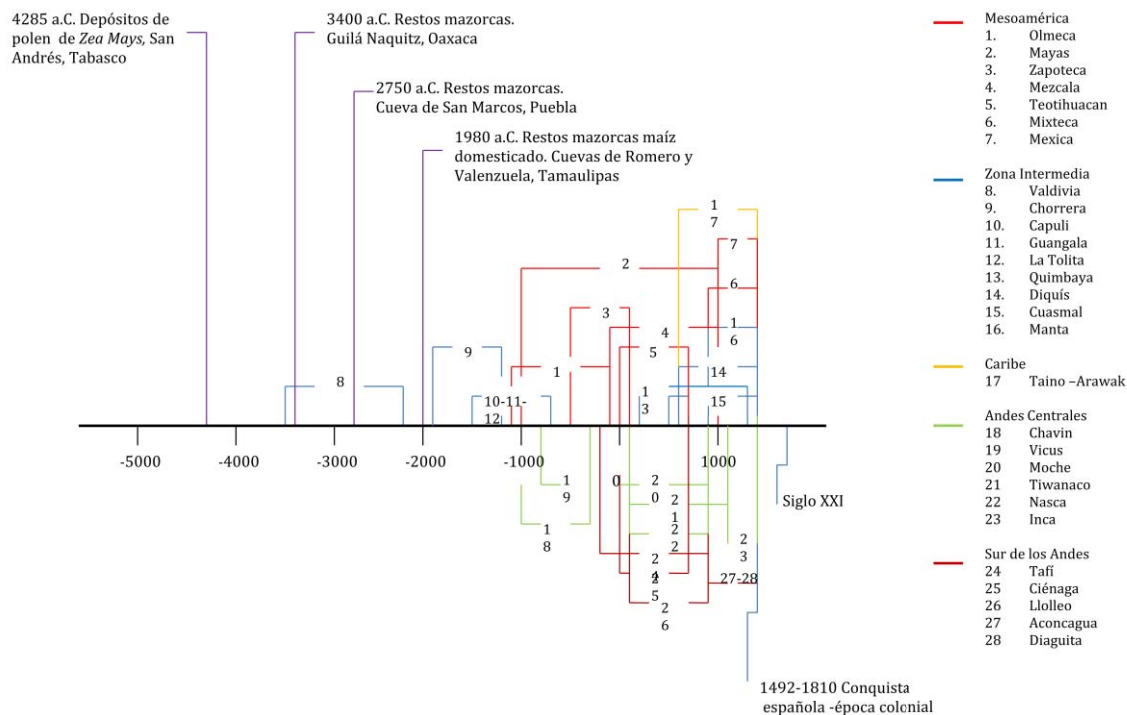
**Figura 3.2.** Regiones geográficas en América latina (Museo Chileno de Arte Precolombino, 2011)



Similarmente, desde Mesoamérica hacia el sur, pueden identificarse varias regiones de interés incluyendo las áreas de la Amazonia y del Chaco (Figura 3.2). Enfatizando en otras regiones, se reconoce una zona intermedia, entre Mesoamérica y el norte de Sudamérica que comprendería parte de Honduras y Costa Rica, Panamá, noroeste de Colombia, Venezuela y Ecuador, en esta zona se establecieron culturas como Valdivia y Chorrera

(Figura 3.3), ambas ligadas al cultivo temprano del maíz. El Caribe comprende las Antillas Mayores, Antillas Menores y norte de Venezuela, esta región está relacionada con la cultura Taino-Arawak. Los Andes centrales en la ladera occidental comprende una parte del sur de Ecuador, tierras altas y costa de Perú, y tierras altas de Bolivia, en este territorio pueden mencionarse culturas como Chavin, Vicus y Moche, entre otras. Sur de los Andes comprende territorios del sur de Bolivia, noroccidente de Argentina y la región austral de la cadena montañosa andina a lo largo de Chile, culturas como Aconcagua, Diaguita, Tafi pueden ser relacionadas en esta región (Figuras 3.2 y 3.3).

**Figura 3.3** Línea del tiempo de algunas culturas pre-colombinas por regiones geográficas ligadas al cultivo del maíz y contexto histórico (Basada en datos recopilados de: Blake, 2006; Vela, 2011; Museo Chileno de Arte Precolombino, 2011)



La dispersión del maíz a partir de su centro de origen en Mesoamérica siguió diferentes rutas, probablemente relacionadas a la migración del hombre americano arcaico en épocas tempranas, y posteriormente su dispersión estaría ligada a complejos sistemas de

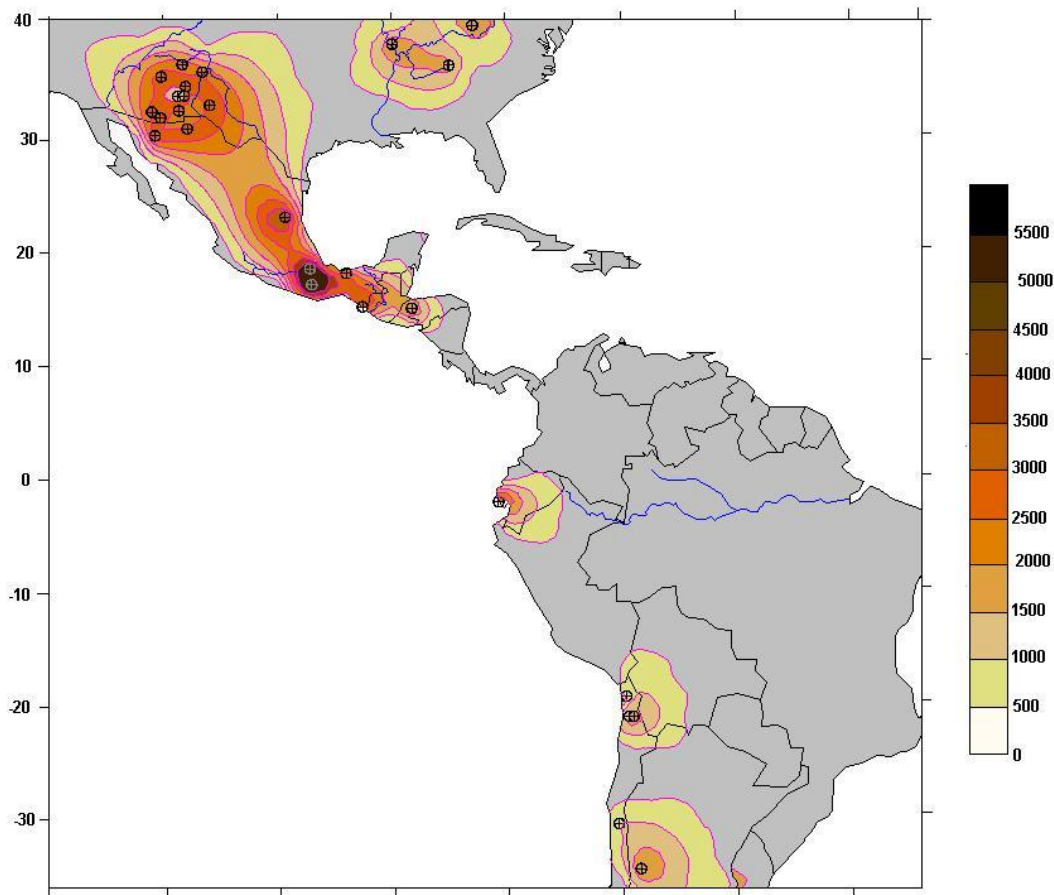
intercambio y comercio entre las culturas que se desarrollaron en la América precolombina. Los medios utilizados para comprender la difusión del maíz desde su centro de domesticación se han basado en datos históricos, genéticos, datación de restos macrobotánicos, fitolitos y muestras de polen recuperadas de sedimentos en lagos, pantanos, y yacimientos arqueológicos, así como también estudios etnolingüísticos. Dos avances fundamentales han ayudado a reformular el entendimiento de la domesticación y la dispersión temprana del maíz. El primero es la acumulación de evidencias genéticas de que el maíz ha surgido de un teocintle anual (*Zea mays ssp. parviglumis*), que se encuentra actualmente en la región del Río Balsas al oeste de México (Matsuoka et al., 2002b). El segundo, la datación por Carbono-14 (mediante espectrometría de masas por acelerador) de fragmentos de maíz minúsculos, lo que ha permitido formular una cronología suficientemente precisa para la aparición y posterior dispersión del maíz (Blake, 2006) (Figura 3.4).

Sin embargo, la evidencia arqueológica referente a la datación del maíz en Sudamérica que indicaría la antigüedad de la introducción del cultivo en el subcontinente es un área todavía de debate. En este sentido, existen dos propuestas. Mientras Piperno y Pearsall (1988) sostienen que el maíz ya estaba presente en el sur de Centroamérica y en el norte de Sudamérica entre 7.700 y 6.000 años antes del presente (AP), la primera introducción en la parte occidental de Sudamérica dataría de aproximadamente 7.000 años (AP), esto es cuestionado por Staller y Thompson (2002) quienes sostienen que el maíz tuvo una introducción más tardía entre el 4.000 y 3.500 años antes del presente.

Existe otro punto de controversia relacionado con la dispersión temprana del maíz, si su éxito fue debido no solamente a su valor nutricional o si también tuvo que ver con su empleo como materia prima, especialmente tallo y semillas, para la producción de bebidas alcohólicas y como fuente de azúcares que podrían ser extraídos mediante la masticación directa de estos tallos y pequeñas semillas (Smalley y Blake 2003). Con la independencia del uso original que se le pudo haber dado al maíz, la emergencia de la agricultura basada en el maíz (seguida por la calabaza, frijol y chile) en Centroamérica es importante porque está ligado al cambio en la economía de los grupos recolectores, pasando de un nomadismo

estacionario a un sedentarismo propio de las culturas agrícolas, surgiendo así las primeras aldeas, procesos vinculados con un incremento en los niveles de población y el desarrollo cultural (Castillo, 2009).

**Figura 3.4** Datos cronológicos antes del presente (AP) de muestras macro botánicas de maíz, mapa adaptado de Blake. 2006\*



\* Basados en la datación de Carbono-14 para los restos macrobotánicos de maíz registrados como los más antiguos recuperados en diferentes yacimientos arqueológicos. Se muestra la distribución de 30 datos obtenidos directamente usando espectrometría de masas por acelerador. Las líneas de contorno de edad están en intervalos de 500 años y muestran las tendencias de la dispersión temprana basadas solamente en los datos de los especímenes datados.

Debido a que México es considerado como centro de origen y domesticación del maíz, el maíz mexicano ha sido objeto de numerosas investigaciones, a lo largo de más de 70 años (Wellhausen et al., 1951; McClintock et al., 1981; Hernandez, 1987; Goodman y Brown, 1988; Matsuoka et al., 2002b; Sanchez et al., 2000). Particularmente, Goodman y Brown.

(1988) describen las razas de maíces mexicanos y su distribución en una escala regional, desarrollando el concepto de grupos raciales y elaborando un estudio comprensivo de investigaciones anteriores (Figura 3.5). Estos autores agruparon las razas mexicanas en tres grupos: El primero contiene las razas que tienen mazorcas largas y angostas que se encuentran en noroeste de México y Estados Unidos; otro grupo que contiene a las razas con mazorcas cónicas encontradas en México a elevadas altitudes y el último grupo con mazorcas largas y pequeñas que se encuentran en las tierras bajas del sur, ya sea a lo largo de la costa Atlántica o Pacífica.

Comparando la riqueza de los datos disponibles para las razas de maíces mexicanos, los materiales nativos para el resto de la América latina, en general, han sido escasamente caracterizados. En Sudamérica, el germoplasma nativo de maíz de la región andina, es el que ha recibido mayor interés, especialmente materiales de Perú y Bolivia, y pocos estudios con un número limitado de individuos o accesiones han incluido material de las tierras bajas de Sudamérica (Bracco et al., 2009).

Este trabajo doctoral ha tenido como propósito la caracterización genotípica de materiales nativos de maíz latinoamericanos, con el fin de explorar y dar una idea general de la diversidad presente y entender la variabilidad genética existente, las relaciones entre estos materiales y cómo están geográficamente distribuidos utilizando marcadores genéticos tipo microsatélite.

Con el fin de elucidar la migración del maíz desde su centro de origen en Mesoamérica hacia Sudamérica y el Caribe se usaron dataciones arqueológicas publicadas de restos macrobotánicos de maíz, estudios antropológicos y etnográficos disponibles para comparar las relaciones encontradas y tener un conocimiento comprensivo del contexto histórico que ha acompañado el maíz hasta nuestros días. Primero se estudiaron las razas de maíz mexicano de manera independiente para la validación del conjunto de marcadores SSRs utilizados y después se extendió el análisis en toda la colección de germoplasma seleccionado esperando clarificar algunas lagunas y controversias concernientes a la dispersión temprana del maíz.

**Tabla 3.5** Relaciones raciales de maíz en México\* (Adaptado de Goodman y Brown, 1988). \*Las razas encerradas en recuadros son las que están más relacionadas entre sí.

**Grupo 1** Maíces con mazorcas largas y delgadas típicas del Noroccidente de México.

Chapalote    Reventador		Dulcillo del Noroeste	Elotes Occidentales
Apachito Cristalino de Chihuahua	Harinoso de Ocho Maíz Blanco de Sonora Onaveño Bofo Gordo Azul		Tabloncillo Tabloncillo Perla Tablilla de Ocho

**Grupo 2** Maíces con endospermo que varían desde dentados harinosos hasta cristalinos tipo “pop” encontrados en Sur y Suroeste de México

Olotillo Olotón Jala Zamorano Amarillo Comiteco Tehua	Tuxpeño Vandeño Celaya Tuxpeño Norteño Ratón	Tepecintle	Zapalote Grande Zapalote Chico Bolita Ancho	Nal-Tel
---	--	------------	--	---------

**Grupo 3** Mazorcas cónicas con granos puntiagudos frecuentemente encontrados en los Valles Altos del Centro de México

Mushito Pepitilla Cacahuacintle	Cónico Cónico Norteño Elotes Cónicos	Chalqueño Arrocillo Amarillo Palomero Toluqueño Palomero de Chihuahua	Maíz Dulce
---------------------------------------	--	--	------------

### 3.2 Materiales y métodos

En este estudio se seleccionaron sobre bases eco-geográficas 194 poblaciones de maíz Latinoamericanas, representando 131 razas procedentes del Banco de Germoplasma del CIMMYT (Tabla 3.1, Figura 3.6, Anexo 7.1).



**Tabla 3.1** Lista por países de origen de las 194 poblaciones de maíz

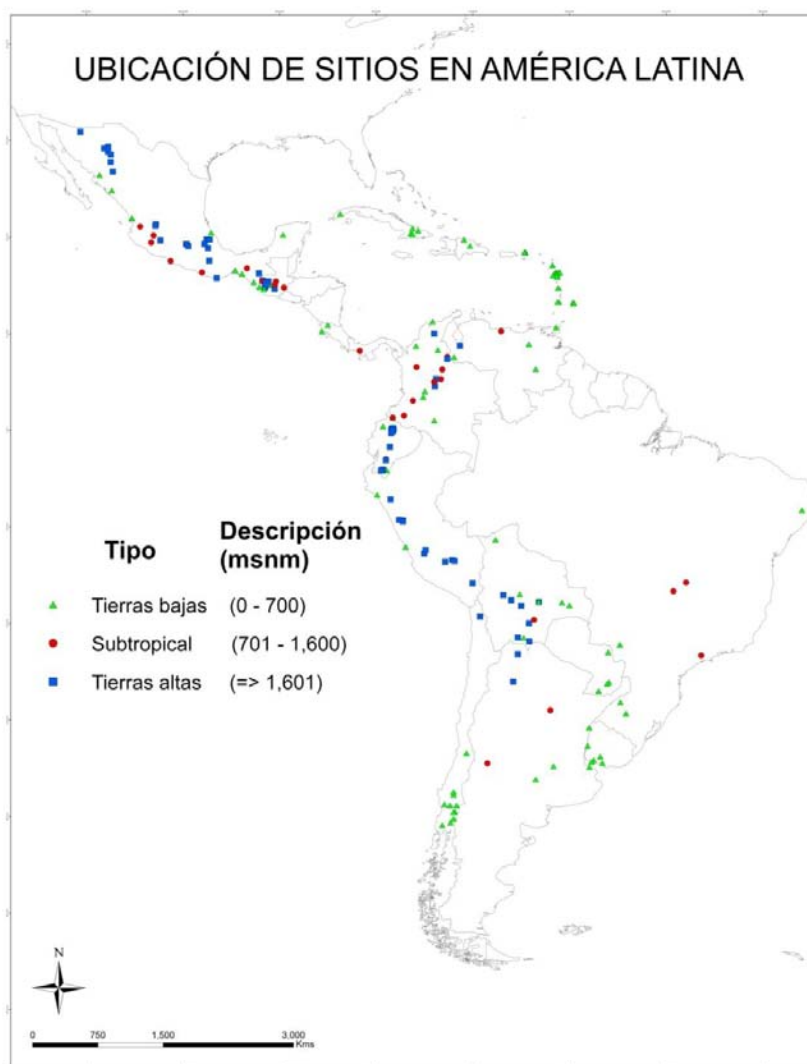
<b>Región</b>	<b>Países de origen</b>	<b>Número de poblaciones</b>
<b>México</b>	México	38
<b>Centroamérica</b>	Guatemala	16
	Costa Rica	2
	Panamá	1
<b>Caribe</b>	Antigua	1
	Barbados	3
	Islas Vírgenes Británicas	4
	Cuba	5
	República Dominicana	2
	Archipiélago de Guadalupe	9
	Martinica	1
	Isla San Vicente	2
	Trinidad y Tobago	1
<b>Sudamérica</b>	Argentina	9
	Bolivia	17
	Brasil	7
	Chile	12
	Colombia	20
	Ecuador	16
	Paraguay	4
	Perú	14
	Uruguay	5
	Venezuela	5

Se sembraron en invernaderos de CIMMYT (30 semillas por cada población) cosechando para cada población un fragmento de hoja de 10-15 cm por individuo formando una muestra compuesta de 15 individuos por población que se analizó como muestra conjunta de tejido (*bulk*) representada por estos 15 individuos. A partir del tejido liofilizado se extrajo el ADN y de acuerdo al protocolo del CIMMYT, disponible en [http://www.cimmyt.org/english/docs/manual/protocols/abc\\_amgl.pdf](http://www.cimmyt.org/english/docs/manual/protocols/abc_amgl.pdf) (ver también Anexo 7.2), y a continuación se cuantificó el ADN genómico, utilizando el espectrofotómetro NanoDrop (Thermo Scientific, Wilmington DE).

En este estudio se hace referencia al término de “población” para referirse a lo que se conoce como variedades criollas o nativas que son mantenidas por el agricultor. Cada una de estas poblaciones se registra como una accesión en los bancos de germoplasma. A veces

la clasificación trae confusión pues en un número considerable de colectas las poblaciones no son puras para una raza particular, sino que se observa gran influencia de otras razas, inclusive, en algunos casos se registran diferentes proporciones de dos o más razas y formas intermedias. Lo anterior puede explicarse en parte por la naturaleza reproductiva del maíz, el intercambio de semilla, el aislamiento geográfico y cultural de algunos grupos de agricultores, al dinamismo de las poblaciones nativas del maíz de los agricultores tradicionales que continúan en evolución, recombinándose con otros maíces y que son seleccionados por los agricultores.

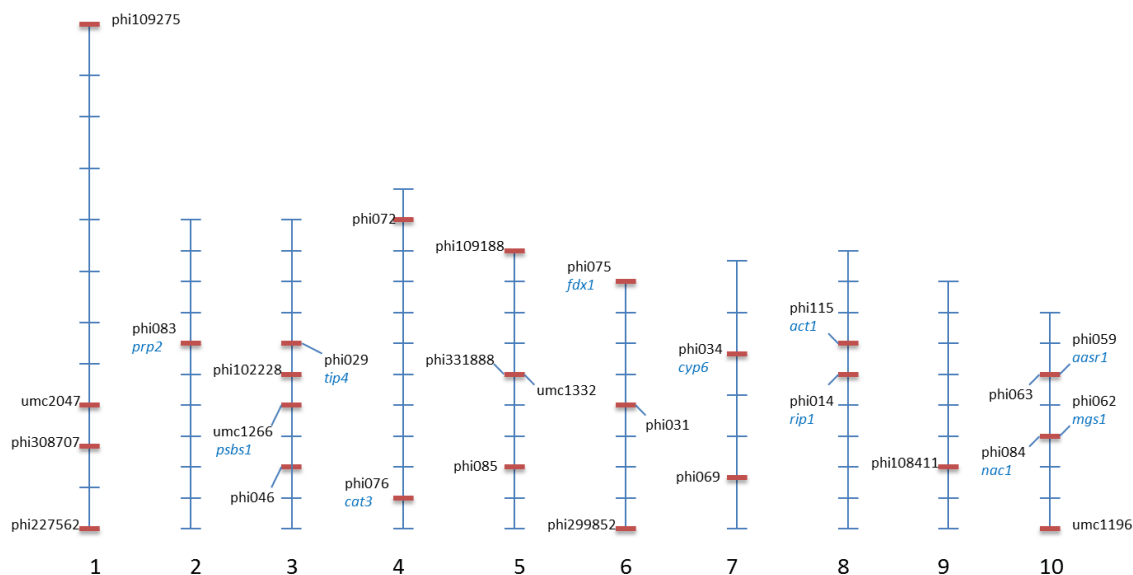
**Figura 3.6** Ubicación geográfica de los lugares de colecta de las 194 poblaciones de maíz basada en los datos de pasaporte (Anexo 7.1)



### 3.2.1 Caracterización Genotípica

Para el análisis genético se usó como primera medida 28 microsatélites de la colección general de SSRs marcados con fluorescencia que posee el laboratorio de diversidad genética de maíz del CIMMYT (Figura 3.7, Anexo 7.3), y que provienen de los marcadores disponibles en la base de datos genéticos y genómicos en línea disponible para maíz (Maize Genetics and Genomics Database; <http://www.maizegdb.org/>). Estos 28 marcadores fueron seleccionados teniendo en cuenta: i) que cubrieran los 10 cromosomas del maíz; ii) que estos marcadores hubieran presentado polimorfismos en trabajos anteriores para diferentes clases de germoplasma de maíz; iii) que el patrón de amplificación del marcador fuera claro y los alelos fácilmente identificables (respetando la unidad de repetición del marcador), en los controles rutinariamente usados en el laboratorio. Mientras que esta información fue obtenida a partir de la base de datos disponible en nuestro laboratorio generada con trabajos previos, algunos de estos marcadores se optimizaron para el trabajo en muestras conjuntas de ADN.

**Figura 3.7** Esquema de los microsatélites utilizados a través de los cromosomas, se muestran los genes de los que se han derivado algunos de ellos (<http://www.maizegdb.org/>).



El trabajo de optimización de los marcadores consistió en el afinamiento de la reacción de PCR para identificar los posibles alelos con una mayor exactitud, fiabilidad y

reproducibilidad. En concreto se optimizó la reacción de PCR de acuerdo al protocolo estandarizado por CIMMYT y disponible en: [http://www.cimmyt.org/english/docs/manual/protocols/abc\\_amgl.pdf](http://www.cimmyt.org/english/docs/manual/protocols/abc_amgl.pdf) (página 48), (Anexo 7.4). Esta reacción de PCR estándar propuesta, se usa especialmente cuando se va a utilizar un secuenciador automático de ADN para el análisis de fragmentos.

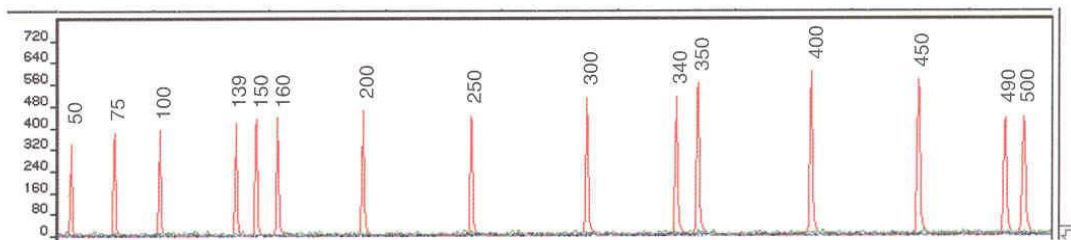
Tomando en cuenta que la técnica de la PCR consta de tres etapas (Rodríguez y Arencibia, 2002), desnaturalización del ADN a 95°C; acoplamiento del marcador a la cadena complementaria (lo cual se realiza a una temperatura específica para cada marcador con el objeto de lograr la máxima especificidad), y extensión de la nueva cadena mediante la enzima Taq ADN polimerasa; el proceso de optimización de un marcador se lleva a cabo: (i) variando como primera medida la concentración de algunos reactivos en la PCR, (ii) incrementando o disminuyendo el número de ciclos en la reacción y/o (iii) variando la temperatura de alineamiento, en el cual se acopla el marcador a la cadena complementaria. La optimización de marcadores se realizó de manera independiente para cada microsatélite en cuestión.

Sin embargo, para la utilización de los marcadores en muestras conjuntas de ADN, los marcadores a utilizar debieron cumplir un requisito adicional. Este requisito consistió en que las frecuencias alélicas presentes a través de 15 muestras individuales de ADN tuvieran una alta correlación con las frecuencias alélicas cuantificadas para estos mismos 15 individuos pero evaluados como una muestra conjunta de ADN.

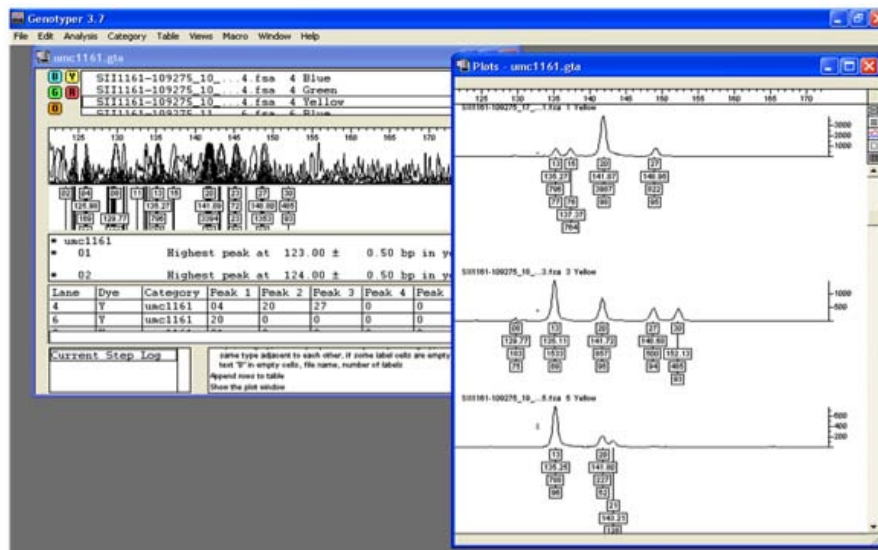
Una vez generados los fragmentos con las técnicas de la PCR, utilizando los marcadores seleccionados y/o optimizados, se llevó a cabo la electroforesis capilar de los productos de la PCR usando un secuenciador automático de capilares ABI 3100 (Applied Biosystems, Foster City, CA, USA) para asegurar la máxima resolución de los productos. La electroforesis capilar fue utilizada para identificar los fragmentos encontrados por tamaño en pares de bases (alelos) y realizar el análisis de los datos tanto para la optimización de marcadores (en los casos necesarios) como para el trabajo general de amplificación en todas las muestras conjuntas de ADN. Las extracciones de ADN y la electroforesis capilar fueron llevadas a cabo en los laboratorios de CIMMYT, México.

En el secuenciador automático de capilares, el tamaño en pares de bases de los fragmentos encontrados en los productos de la reacción de PCR, fue calculado de manera automática usando el programa GeneScan® 3.1 (PerkinElmer/Applied Biosystems, Foster City, CA, USA) cuando se compara con los fragmentos del estándar interno GS500 ROX (Figura 3.8). Los fragmentos encontrados con GeneScan se consideraron como posibles alelos usando la función categoría del programa Genotyper® 2.1 (PerkinElmer/Applied Biosystems, Foster City, CA, USA). Los fragmentos encontrados se identificaron con la información de su tamaño en pares de bases, altura de pico, nombre de la categoría (rango de tamaño), y un valor asignado a la calidad del pico (Figura 3.9).

**Figura 3.8** Patrón de fragmentos del estándar interno GS500 ROX bajo condiciones de desnaturalización.



**Figura 3.9** Identificación de fragmentos utilizando programa Genotyper® 2.1 (Applied Biosystems).



Para el cálculo de las frecuencias alélicas de cada fragmento presente en cada muestra de ADN conjunta se utilizó el programa Freqs-R (Franco et al., 2005); este programa fue creado por CIMMYT con el objetivo de la caracterización genotípica de poblaciones heterogéneas de maíz y utiliza el software R (<http://cran.r-project.org>). Este programa se basa, principalmente en el hecho, de que la frecuencia de un alelo en una muestra conjunta de ADN es proporcional a la altura de pico que tuvo este alelo en la muestra conjunta, esto es, para los casos extremos, el alelo con la mayor altura de pico representa el alelo más frecuente en la muestra conjunta y de manera análoga, el alelo con la altura más baja representa al alelo menos frecuente. En general, para muestras conjuntas de ADN, la frecuencia de un alelo está determinada por la fracción entre la altura de alelo en cuestión y la suma de todas las alturas de los alelos encontrados en la muestra conjunta de ADN. Los datos de las frecuencias alélicas son extraídos en formato Excel y se integrarán en una base de datos pública y en el Registro Central del GCP (Generation Challenge Programme, <http://www.generationcp.org>).

Este programa, Freqs-R, tiene la ventaja que puede remover las señales o ruido de fondo y artefactos generados por las bandas *stutter*, causadas por el llamado “tartamudeo de la polimerasa” que sucede cuando al menos un ciclo de realineamiento incorrecto y amplificación proporciona una fracción de productos de extensión incompleta. Una vez la frecuencias alélicas han sido calculadas con el programa Freqs-R, el programa FtoL-R (Franco et al., 2007) fue usado para simular 15 individuos dialélicos, expresando los alelos como longitudes en pares de bases para cada uno de ellos, de forma tal que el conjunto de 15 individuos simulados satisficieron las frecuencias alélicas y la heterocigosidad esperada bajo condiciones de equilibrio “Hardy-Weinberg” para todo el conjunto (muestra conjunta de ADN).

La transformación de las frecuencias alélicas calculadas con Freqs-R (Franco et al., 2005) en individuos dialélicos con el programa FtoL-R (Franco et al., 2007), es necesaria debido a que diferentes programas para el análisis de diversidad no aceptan frecuencias alélicas provenientes de análisis de muestras conjuntas de ADN. Estas muestras conjuntas de ADN pueden presentar múltiples alelos por *locus*, es decir, en el caso de un individuo heterocigótico se identificarían 2 alelos, pero al tener 15 individuos juntos se pueden

encontrar múltiples alelos basados en las frecuencias alélicas de éstos. El porcentaje de datos perdidos en este estudio para los 28 microsátélites y a través de las 194 accesiones de maíz fue de 5,8%.

### 3.2.2 Cálculos estadísticos

La diversidad genética o heterocigosidad esperada es la estimación de la variabilidad genética en la población o unidad en estudio, y está definida como la probabilidad de que, en un locus único, cualquier par de alelos, escogidos al azar de la población, sean diferentes entre sí, es decir, no vienen de un ancestro común.

Heterocigosidad por *locus j* con *i* alelos:

$$h_j = 1 - \sum_{i=1}^{n_a} p_{ij}^2$$

Heterocigosidad promedio para *L loci*:

$$H = 1/L \sum_{j=1}^L h_j$$

Donde:

$h_j$  = la heterocigosidad esperada para el *i*-ésimo *locus*

$p_{ij}$  = frecuencias del *i*-ésimo alelo en el *locus j*

$H$  = la heterocigosidad promedio para varios *L loci*

$L$  = el número total de *loci*

El número promedio de alelos es la suma de todos los alelos diferentes detectados en todos los *loci*, dividido por el número total de *loci*:

$$n = \frac{1}{L} \sum_{i=1}^l n_i$$

Donde:

$n_i$  = Número de alelos detectados en el *i*-ésimo *locus*

$L$  = Número de *L loci*

Una de las medidas útiles para la división en subpoblaciones de una población global es el Índice de Fijación ( $F_{ST}$ ) desarrollado por Wright (1965), que mide la reducción de la heterocigosidad esperada por la subdivisión en subpoblaciones de una población global y la deriva genética, y está definido como:

$$F_{ST} = 1 - H_s/H_T$$

Donde:

$H_s$  = Promedio de las heterocigosidades esperadas en las subpoblaciones

$H_T$  = Heterocigosidad esperada en la población global

$F_{ST}$  tiene un valor mínimo de 0 y un máximo teórico de 1 que indica la fijación de alelos alternativos en las poblaciones o subpoblaciones;  $F_{ST}$  suele utilizarse como medida de diferenciación genética entre subpoblaciones o entre poblaciones. Para interpretar los valores de  $F_{ST}$ , Wrigth sugirió lo siguiente:

0 - 0,05 = muy poca diferenciación

0,05 - 0,15 = poca a moderada diferenciación

0,15 - 0,25 = alta diferenciación

> 0,25 = diferenciación muy importante.

Los parámetros de diversidad genética como número de alelos, diversidad genética (heterocigosidad esperada), así como el método de análisis de varianza molecular (AMOVA), que determina la variación molecular dentro material diverso o de una especie, Índice de Fijación ( $F_{ST}$ ) y distancias genéticas entre poblaciones usando la proporción de alelos compartidos, fueron calculadas utilizando el programa PowerMarker (Liu y Muse, 2005) para cada subgrupo de poblaciones. El programa estadístico Darwin 5.0 (Perrier y Jacquemoud-Collet, 2006) fue usado para crear los análisis de agrupamiento por el método de Neighbor-Joining y el Análisis de Coordenadas Principales (ACoP) basado en la matriz de distancias genéticas obtenida con PowerMarker.

El programa Structure 2.2 (Pritchard et al., 2000) fue utilizado para analizar los datos generados con los microsatélites, en primera instancia en las 38 poblaciones mexicanas



solamente y posteriormente en el conjunto de las 194 poblaciones para tener un análisis de estructura poblacional. El programa implementa un modelo de agrupamiento basado en un modelo estadístico Bayesiano usando datos genotípicos de marcadores no ligados. Una de las aplicaciones consiste en demostrar la presencia de una estructura poblacional, identificando distintas subpoblaciones genéticas y/o asignando individuos a subpoblaciones. Para este propósito, se asume un modelo en el cual existen  $K$  poblaciones/grupos (donde  $K$  es usualmente desconocido), cada uno de los cuales es caracterizado por un set de frecuencias alélicas en cada *locus*. Los individuos/poblaciones son asignados probabilísticamente en la población/grupo a la cual presentan mayor probabilidad de pertenencia buscando (como función objetivo) que dentro de los grupos generados los *loci* estén en equilibrio Hardy-Weinberg y equilibrio de ligamiento. Usando en el análisis de estructura el modelo de mezcla de ancestros (*admixture model*), en el cual los individuos/poblaciones podrían tener un conjunto de ancestros comunes, se puede asumir que la población  $i$  ha heredado una fracción de su genoma del grupo  $K$ . De esta manera el programa calcula para cada individuo/población la proporción de ancestro ( $Q_{i-k}$ ) para cada uno de los grupos asumidos. A las poblaciones no se les asignó *a priori* la pertenencia a ninguno de los grupos, a cada individuo simulado de cada población se le permitió variar y agrupar con los individuos que estuvieran más relacionados genéticamente basados en los datos de microsatélites.

Para un análisis preliminar, el número  $K$  fue explorado para las 38 poblaciones mexicanas y para el total de las 194 poblaciones (incluyendo las mexicanas) variando desde  $K=1$  hasta  $K=20$ . En este sentido, y de acuerdo con la documentación de STRUCTURE el valor de  $K$  que representaría (definiría) la estructura poblacional correspondería al valor máximo de los valores medios de la probabilidad de ocurrencia de los datos observados dado el modelo y el número de grupos en cuestión  $\ln[\text{prob}(\text{dato})]$ . Un total de 100.000 iteraciones fueron utilizadas en cada proceso; se utilizó una fase inicial de optimización y filtrado (*burn-in period*) de 50.000 iteraciones y dos corridas independientes para cada  $K$ , excepto para  $K=1$  (asume que no hay estructura). Para estos primeros parámetros de análisis la determinación del número de grupos,  $K$ , no fue determinante porque el valor  $\ln[\text{prob}(\text{dato})]$  contra el número de  $K$  va incrementándose a medida que  $K$  aumenta, sin lograr una estabilidad o

cambio de inflexión (Anexo 7.5). Evanno et al. (2005) plantea otro procedimiento para tratar de determinar el número  $K$ , a partir de los datos obtenidos del STRUCTURE, que consiste en calcular un delta  $K$  ( $\Delta K$ ) basado en la razón de cambio en el logaritmo de la probabilidad de los datos entre valores sucesivos de  $K$  (Anexo 7.6).

Basados en resultados previos de estructura poblacional en México y en América (Vigourox et al., 2008; Reif et al., 2006) y con el fin de optimizar el uso de STRUCTURE y aumentar el número de corridas independientes, ya que en el cálculo de  $\Delta K$  es determinante el valor de la desviación estándar, en este caso, las poblaciones mexicanas se asumieron valores de  $K=1$  hasta  $K=8$ , y para el total de las poblaciones latinoamericanas,  $K=1$  variando hasta  $K=15$ . El proceso se realizó con un total de replicaciones de 1.000.000 que podría mejorar la estimación de las proporciones de ancestro de los individuos y también el promedio sobre los 15 individuos de cada población, *burn-in period* de 500.000 y 5 corridas independientes para cada valor de  $K$ . En este caso, el valor más alto de  $\Delta K$  representaría el número apropiado de grupos que definen el nivel de la estructura. Para asignar una población a un grupo se tomó el valor arbitrario de proporción de ancestro  $Q_{jk}$  mayor al 51,0%.

### **3.3 Resultados**

#### **3.3.1 Validación de los marcadores seleccionados con las poblaciones de maíz mexicanas**

En la primera parte de este estudio se enfatiza en la clasificación de los maíces mexicanos para confirmar la relevancia para estudios de diversidad de 28 marcadores moleculares en la colección entera de germoplasma latinoamericano seleccionado para este estudio. Para la validación del set de marcadores se usó como referencia la clasificación según complejos raciales propuesta por Goodman y Brown, (1988) por ser un estudio que recoge la información generada por trabajos anteriores (Tabla 3.5). Tratando de abarcar los diferentes complejos raciales y geográficos descritos en dicho estudio, se investigaron 27 razas típicas de las regiones: Valles Centrales, Sierra Madre Occidental y Sur de México representadas por 38 poblaciones. Las razas mexicanas aquí representadas fueron: razas de la Mesa Central Mexicana (Cacahuacintle, Chalqueño, Cónico, Cónico Norteño, Palomero

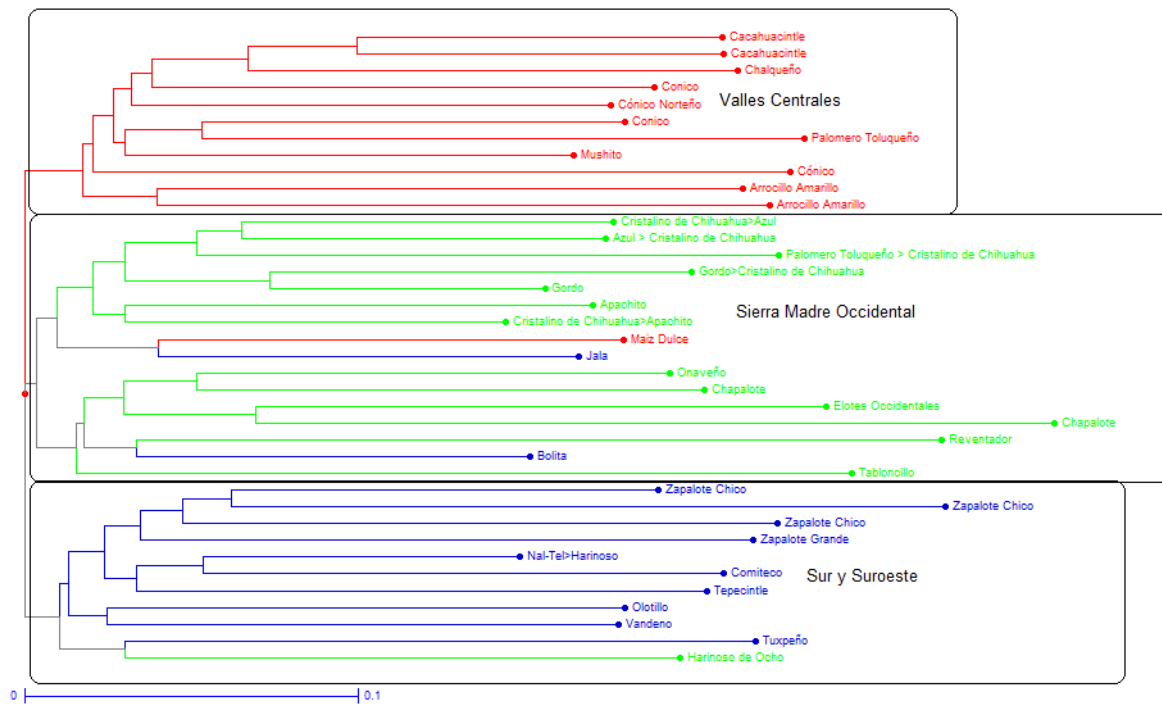
Toluqueño, Mushito y Arrocillo Amarillo), región geográfica del la Sierra Madre Occidental (Azul, Cristalino de Chihuahua, Gordo, Apachito, Elotes Occidentales, Chapalote, Tabloncillo, Reventador, Onaveño y Harinoso de Ocho) y la región del Sur y Suroeste de México (Jala, Bolita, Olotillo, Vandeño, Tuxpeño, Zapalote Chico, Zapalote Grande, Nal-Tel, Comiteco y Tepecintle).

El primer paso en la validación de los marcadores fue el análisis de agrupamiento (Figura 3.9). Se etiquetó con el color verde (Grupo 1) a los maíces con mazorcas largas y delgadas típicas del Noroccidente de México. Marcado con color azul (Grupo 2) se encuentran maíces con una gran variabilidad en longitud de mazorca del Sur y Suroeste de México, por ejemplo mazorcas de 30 cm o más como es el caso de la raza Jala hasta los maíces con longitud de mazorca de aproximadamente 8 cm de la raza Nal-Tel, 10 cm de la raza Zapalote Chico y 11 cm de la raza Bolita (Wellhausen et al., 1951). A este grupo también pertenecen los maíces ampliamente usados de las razas Tuxpeño y Vandeño, así como también el complejo Zapalote. Con color rojo (Grupo 3), se incluyen razas con mazorcas cónicas con granos puntiagudos frecuentemente encontrados en los Valles Altos del Centro de México, y en la descripción morfológica en el tipo de planta, se encuentran hojas pubescentes y tallos gruesos con un sistema radicular débil (Goodman y Brown, 1988).

Observando el análisis de agrupamiento (Figura 3.10), la mayoría de las razas agruparon de acuerdo con la clasificación de Goodman y Brown (1988), con pocas excepciones. Algunas discrepancias con estudios previos fueron para las razas Bolita, Maíz Dulce, Jala y Harinoso de Ocho. Para el subgrupo denominado como región Geográfica de la Mesa Central comprendió todas las poblaciones exceptuando la identificada como JALI78 de la raza Maíz Dulce, catalogada en el grupo de maíces provenientes de las tierras altas del Centro de México por Goodman y Brown (1988). Esta raza agrupó en el Complejo de la Sierra Madre Occidental, consistente con la clasificación reportada por Reif et al. (2006) en el grupo de accesiones pertenecientes al Noroccidente de México y catalogada en el grupo de los maíces Cónicos del Occidente de México por Ron et al. (2006). Además en el presente estudio la raza Maíz Dulce presenta relación con la población NAYA 337 de la raza Jala. Esto puede tener explicación por la cercanía geográfica que presentan las dos poblaciones, indicando un posible flujo génico entre ellas, aunque la raza Maíz Dulce se

encuentra generalmente en alturas intermedias (1000-1500 msm) (Hernandez, 1987) mientras que la raza Jala se encuentra muy localizada en la parte sur de Nayarit, a unos 1000 metros sobre el nivel del mar.

**Figura 3.10** Análisis de agrupamiento para 38 poblaciones de maíz de México, generado con el método Neighbour-Joining basado en la distancia genética de proporción de alelos compartidos para 28 SSRs. Las poblaciones están etiquetadas según la raza a la que pertenecen y están coloreadas de acuerdo a la clasificación Goodman y Brown (1988), rojo corresponde a las Valles Centrales; verde a la Sierra Madre Occidental; y azul al Sur y Suroeste.

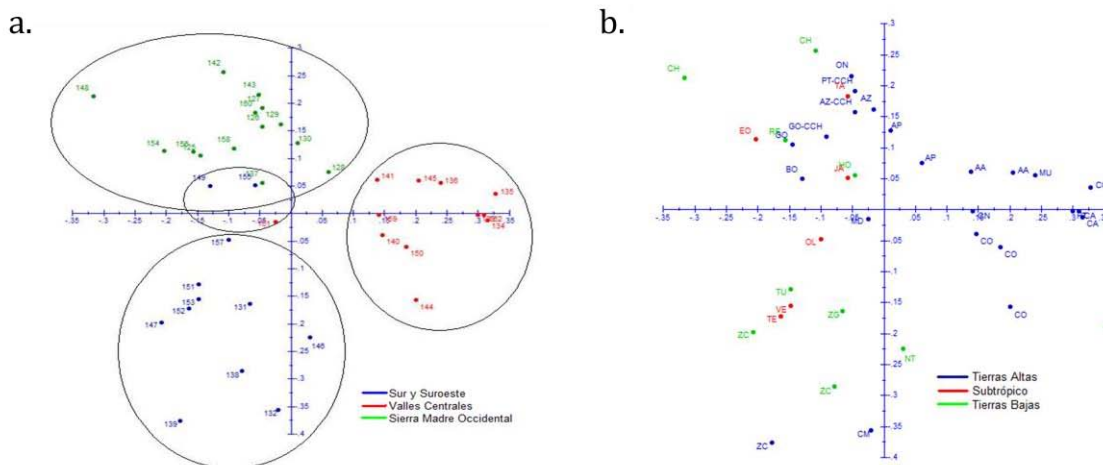


En el subgrupo observado para las poblaciones que están en su mayoría en la zona geográfica de la Sierra Madre Occidental se encuentra la raza Bolita (OAXA 223) que ha sido clasificada en el grupo de maíces del complejo Sur (Goodman y Brown, 1988; Reif et al., 2006). Esto podría explicarse en base a sus orígenes y parentescos pues su origen probable fue a partir de la hibridación entre Zapalote Chico y Tabloncillo (Reif et al., 2006; Wellhausen et al., 1951). A su vez la raza Tabloncillo es consistente con la clasificación de Goodman y Brown (1988), pero no para la clasificación de Reif et al. (2006), pues se confirma su relación con las razas Reventador y Chapalote. De otro lado, la raza Harinoso

de Ocho agrupa en este estudio con las razas del Sur de México y no con el grupo del Noroccidente como sugiere Goodman y Brown (1988) y Reif et al. (2006), pero es en alguna forma consistente con los estudios cromosómicos de McClintock et al. (1981), donde Harinoso de Ocho se encuentra relacionada como raza costera del Pacífico con el grupo de Zapalotes, Nal-Tel, Tepecintle y Comiteco, razas del Suroeste de México.

El segundo paso en el proceso de validación de los marcadores fue analizar las asociaciones entre las 38 poblaciones de maíz mexicanas basadas en el Análisis de Coordenadas Principales que están representadas en la Figuras 3.11a y 3.11b.

**Figura 3.11** Análisis de Coordenadas Principales para las 38 poblaciones de maíz mexicanas basado en la distancia genética de proporción de alelos compartidos para 28 SSRs. En la figura 3.11a las poblaciones están identificadas con número ID (ver Anexo 7.1) y coloreadas según la clasificación de Goodman y Brown (1988). En la Figura 3.11b se encuentran identificadas por código de raza (Anexo 7.7) y por altitud del área donde se recolectaron.



Los círculos de la derecha, superior e inferior de la Figura 3.11a apoyan claramente la estructura del análisis de agrupamiento pues congregan las poblaciones según los grupos definidos anteriormente (Figura 3.10a). El Componente Principal (CP) 1 explica el 7.11% del total de la varianza y el CP 2 el 6.44%. Siendo de particular interés en este análisis (Figura 3.11a), el ovalo central pequeño donde se encuentran asociadas las accesiones

número 137 de la raza Harinoso de Ocho, 149 de la raza Bolita, 155 de la raza Jala, y 161 de la raza Maíz Dulce que han representado discrepancias con estudios anteriores.

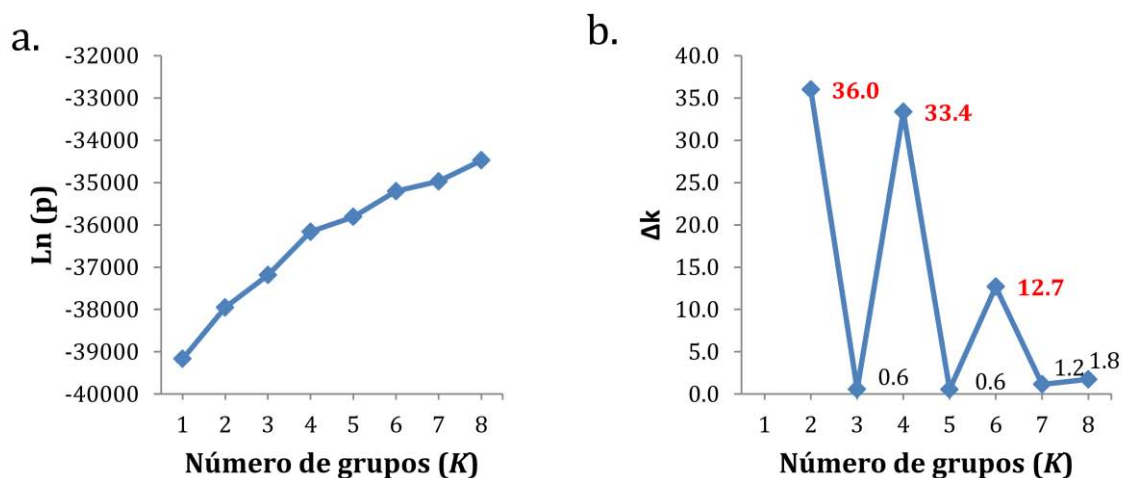
En la Figura 3.10b se encuentran identificadas las poblaciones por código de raza (Anexo 7.7) y por altitud del área donde se recolectaron. De esta manera se observa que los grupos formados no son específicos para la altura de colecta, a saber: trópico, subtropical y valles altos centrales, sino más bien representa la clasificación según regiones geográficas mencionadas anteriormente.

El tercer paso para la validación del conjunto de marcadores utilizados fue realizar un análisis de estructura poblacional. Del mismo modo que en los análisis preliminares de estructura para los maíces mexicanos no fue posible determinar directamente el valor de  $K$  (Figura 3.12a). Utilizando el procedimiento reportado por Evanno et al. (2005) en el análisis de estructura para las 38 poblaciones mexicanas el valor más alto de  $\Delta K$  fue encontrado en  $K=2$  y los siguientes valores más altos son para  $K=4$  y  $K=6$  (Figura 3.12b).

En este análisis, según los resultados obtenidos para los diferentes valores de  $K$ , se encontraron importantes consistencias con los análisis de agrupamiento (Figura 3.10) y Componentes Principales (Figuras 3.11). Se identificaron grupos específicos según la geografía mexicana y en el caso más específico complejos raciales particularmente para  $K=4$  y  $K=6$ . Los nombres de los 6 grupos identificados serían (Tabla 3.2): Valles Altos Centrales (VAC), Sierra Madre Occidental (SMO), Sur de México (S), Chapalote y afines (CH-a), Ladera Oriental de la Sierra Madre Occidental (LOri), Ladera Occidental de la Sierra Madre Occidental (LOcc) y Complejo Zapalote (CZ).

Para  $K=2$ , el 100,0% de las poblaciones pertenecieron a los dos grupos, que se definieron como VAC incluyendo la población con JALI 102 (ID 160) de la raza Tabloncillo que en análisis previos estaba catalogada como perteneciente al complejo de la Sierra Madre Occidental; el siguiente grupo comprende todas las poblaciones del Sur de México y del sistema montañoso occidental (Tabla 3.2).

**Figura 3.12** Gráficas del Ln de la probabilidad (a), y (b)  $\Delta K$ , del análisis de estructura poblacional de 38 poblaciones de maíz mexicanas. Se graficaron los valores promedio de  $\ln[\text{prob}(\text{dato})]$  de 5 corridas independientes para cada valor de K, de la misma manera estos valores fueron utilizados para el cálculo de los valores  $\Delta K$ . Con color rojo están marcados los más altos valores encontrados para  $\Delta K$ .



Si se asume  $K=4$ , el 95,0% de las poblaciones fueron asignadas a un grupo, exceptuando las poblaciones identificadas con ID 149 e ID 161 de las razas Bolita y Maíz Dulce respectivamente (Tabla 3.2). El supuesto de esta estructura apoya análisis previos en los maíces mexicanos, donde se pueden relacionar los maíces etiquetados como VAC con el grupo de maíces cónicos; S con el grupo definido como maíces de las partes altas e intermedias del sur de México (Kato et al., 2009); SMO se relaciona con el grupo de maíces de alturas intermedias de temporal y costas semiáridas de riego (Kato et al., 2009) o según Ron et al. (2006) con las razas típicas del occidente y noroeste de México.

En este estudio se definió de igual forma que Kato et al. (2009) al grupo de Cahapalote y afines (CH-a) comprendiendo la raza Chapalote (ID 148, SINA 2) y Reventador (ID 154, NAYA 39) pero incluyendo además la raza Elotes Occidentales (ID 156, JALI 71), sin embargo y sorprendentemente la población con ID 142, SINA6 de la raza también Chapalote no se incluye en este grupo sino en el grupo SMO. Admitiendo el valor  $K=6$ , todas las poblaciones fueron adjudicadas a algún grupo. El grupo más consistente a través de los tres análisis realizados fue el grupo de los maíces pertenecientes a los Valles Altos del centro de México (Figura 3.13, Tabla 3.2).

**Tabla 3.2 Asignación de las poblaciones mexicanas a cada grupo.** \*Se especifica el porcentaje de ancestro para las poblaciones que no pertenecieron a ningún grupo, tomando un tope mínimo de ancestro arbitrario del 51,0%.

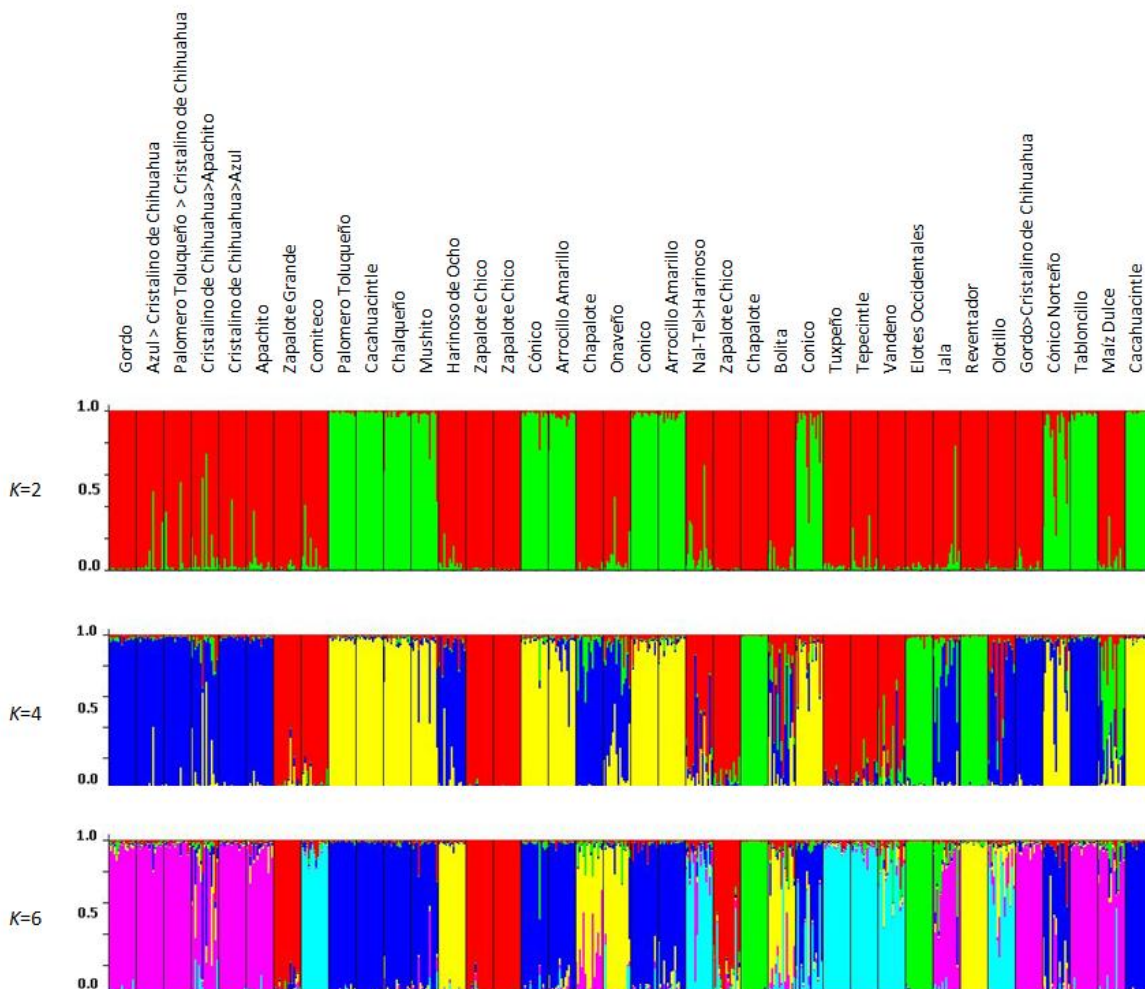
<b>ID</b>	<b>K=2</b>	<b>K=4</b>	<b>K=6</b>
125	SMO y S	SMO	LOri
126	SMO y S	SMO	LOri
127	SMO y S	SMO	LOri
128	SMO y S	SMO	LOri
129	SMO y S	SMO	LOri
130	SMO y S	SMO	LOri
131	SMO y S	S	CZ
132	SMO y S	S	S
133	VAC	VAC	VAC
134	VAC	VAC	VAC
135	VAC	VAC	VAC
136	VAC	VAC	VAC
137	SMO y S	SMO	LOcc
138	SMO y S	S	CZ
139	SMO y S	S	CZ
140	VAC	VAC	VAC
141	VAC	VAC	VAC
142	SMO y S	SMO	LOcc
143	SMO y S	SMO	LOcc
144	VAC	VAC	VAC
145	VAC	VAC	VAC
146	SMO y S	S	S
147	SMO y S	S	CZ
148	SMO y S	CH-a	CH-a
149	SMO y S	30,7(S) 37,8(SMO) 16,0(CH-a) 15,5 (VAC)*	LOcc
150	VAC	VAC	VAC
151	SMO y S	S	S
152	SMO y S	S	S
153	SMO y S	S	S
154	SMO y S	CH-a	CH-a
155	SMO y S	SMO	LOri
156	SMO y S	CH-a	LOcc
157	SMO y S	SMO	S
158	SMO y S	SMO	LOri
159	VAC	VAC	VAC
160	VAC	SMO	LOri
161	SMO y S	12,3(S) 26,0(SMO) 50,6(CH-a) 11,0(VAC)*	LOri
162	VAC	VAC	VAC

El grupo formado por los maíces del Sur en la asunción de  $K=4$  fue subdividido para formar un grupo independiente al que se denominó Complejo Zapalote, formado por las 4 poblaciones de las razas Zapalote Chico y Zapalote Grande. El grupo SMO fue fraccionado de igual manera para encontrar grupos de maíces que están distribuidos geográficamente en



las laderas Oriental y Occidental (LOri y LOcc) del complejo montañoso de la Sierra Madre Occidental que atraviesa el país desde el sur hasta el norte. Se mantiene el grupo CH-a, pero en este caso, la raza Reventador se encuentra en el grupo de LOcc.

**Figura 3.13** Estructura de población estimada para 38 poblaciones de maíz mexicanas. Cada población está identificada con el número ID correspondiente y representada por un segmento de línea vertical, que puede estar dividida en segmentos coloreados que representan la pertenencia de la población a cada agrupamiento  $K$ . El valor más alto de  $[\ln[\text{prob}(\text{dato})]]$  de las 5 corridas independientes fue escogido para la representación gráfica de cada agrupamiento  $K$ .



En la Tabla 3.3, se puede observar detalladamente la clasificación de las poblaciones de este estudio comparadas con estudios previos de la siguiente manera: la columna A

representa los resultados del presente análisis para cuatro grupos asumidos en el análisis de estructura poblacional del programa Structure y B para tres grupos asumidos en el análisis de agrupamiento; columna C, clasificación utilizada como referencia de Goodman y Brown (1988) basada en la combinación de datos morfológicos, citológicos e isoenzimáticos; D, clasificación en tres grupos basada en datos obtenidos con microsatélites por Reif et al. (2006); E, clasificación basada en la combinación de datos morfológicos e isoenzimáticos asumiendo tres grupos principales y reportada por Sanchez et al. (2000); columna F, clasificación basada en áreas ecológicas que tienden a ocupar las razas definidas y características morfológicas de mazorca de Ortega-Paczka (2003) asumiendo 4 grupos; finalmente en la columna G, clasificación en cuatro grupos por Kato et al. (2009) basada en características de mazorca.

En particular, las razas pertenecientes al complejo de la Sierra Madre Occidental como se ha llamado en este estudio, pero que se cultivan a alturas mayores de 1600 metros como: Apachito, Apachito/Cristalino de Chihuahua, Azul/Cristalino de Chihuahua, Gordo, Gordo/Cristalino de Chihuahua y Palomero Toluqueño /Cristalino de Chihuahua, y Maíz Dulce agrupan con las razas típicas del los Valles Altos Centrales cuando la clasificación es basada también en datos morfológicos (Columnas E y F) (Tabla 3.3). De esta manera, se sugiere nuevamente, como se indicó en el análisis de Coordenadas Principales (Figura 3.11b), que este conjunto de marcadores no es especialmente discriminatorio para clasificar las poblaciones según altura de colecta, pero si para regiones naturales definidas, subrayando de esta manera la importancia del flujo génico que se ha podido llevar a cabo en estas regiones.

Separadamente, cuando se asumen cuatro grupos por análisis de estructura poblacional fue posible identificar el grupo Chapalote y afines (4) reportado por Sanchez et al. (2000) y Ortega-Paczka. (2003) que incluye la raza Chapalote, Reventador y Dulcillo del Noroeste, esta última no presente en este estudio; pero incluyendo la raza Elotes Occidentales reportada por los estudios mencionados en grupo de maíces de ocho hileras del occidente de México.

**Tabla 3.3** Razas, accesiones, alturas de colecta, clasificación de acuerdo a cinco estudios previos; ng: raza no agrupada en un grupo específico, np: raza no presente en el estudio. <sup>a</sup>Resultados del presente estudio basados en resultados de estructura poblacional; <sup>b</sup>Resultados del presente estudio basados análisis agrupamiento; <sup>c</sup>Goodman y Brown (1988); <sup>d</sup>Reif et al. (2006); <sup>e</sup>Sanchez et al. (2000); <sup>f</sup>Ortega-Paczka (2003) y <sup>g</sup> Kato et al. (2009).

Raza	Accesión	Altura (m)	A <sup>a</sup>	B <sup>b</sup>	C <sup>c</sup>	D <sup>d</sup>	E <sup>e</sup>	F <sup>f</sup>	G <sup>g</sup>
Apachito	CHIH 38	1800	1	1	1	np	3	3	np
Arrocillo Amarillo	PUEB 91	2060	3	3	3	3	3	3	3
	VERA 359	2200	3	3					
Azul/Cristalino de Chihuahua	CHIH 133	2095	1	1	1	np	3	3	np
Bolita	OAXA 223	200	ng	1	2	2	1	1	1
Cacahuacintle	MEXI 7	2562	3	3	3	3	3	3	3
	MEXI 212	2800	3	3	3				
Chalqueño	MEXI 726	2700	3	3	3	3	3	3	3
Chapalote	SINA 6	75	1	1	1	1	2	4	4
	SINA 2	61	4	1	1				
Comiteco	CHIS 94	1800	2	2	2	2	2	2	2
Cónico	PUEB 70	2469	3	3	3	3	3	3	3
	TLAX 151	2463	3	3	3				
	MEXI 3	2652	3	3	3				
Cónico Norteño	ZACA 12	1950	3	3	3	3	3	3	3
Cristalino de Chihuahua/Apachito	CHIH 207	2510	1	1	1	np	3	3	np
Cristalino de Chihuahua/Azul	CHIH 218	1964	1	1	1	np	3	3	np
Elotes Occidentales	JALI 71	1555	4	1	np	np	1	1	np
Gordo	CHIH 131	2095	1	1	1	np	3	3	np
Gordo/Cristalino de Chihuahua	CHIH 160	2000	1	1	1	np	3	3	np
Harinoso de Ocho	NAYA 24	100	1	2	1	1	1	np	np
Jala	NAYA 337	1080	1	1	2	2	1	1	1
Maíz Dulce	JALI 78	1890	ng	1	3	3	3	3	3
Mushito	MICH 362	2340	3	3	np	np	2	np	np
Nal-Tel	YUCA GP2	30	2	2	2	2	2	2	2
Olotillo	OAXA GP1	120	1	2	2	2	2	2	2
Onaveño	SONO 24	1640	1	1	1	np	1	1	np
Palomero Toluqueño/Cristalino de Chihuahua	CHIH150	2140	1	1	1	np	3	3	np
Palomero Toluqueño	MEXI 5	2652	3	3	3	3	3	3	3
Reventador	NAYA 39	100	4	1	1	1	2	4	4
Tabloncillo	JALI 102	1280	1	1	1	2	1	1	1
Tepecintle	CHIS 63	760	2	2	2	2	2	2	2
Tuxpeño	VERA 39	106	2	2	2	2	2	2	2
Vandño	GUER GP25	1000	2	2	2	np	2	2	2
Zapalote Chico	OAXA 50	100	2	2	2	2	2	2	2
	PI 217413	100	2	2	2				
	CHIS 662	100	2	2	2				
Zapalote Grande	CHIS 104	100	2	2	2	2	2	2	2

La diversidad genética presente en los maíces mexicanos también queda reflejada en con el análisis de varianza molecular (AMOVA) que se realizó para los grupos de acuerdo al análisis de agrupamiento (Figura 3.9), y para los complejos raciales descritos por Goodman

y Brown, 1988, porque el mayor porcentaje de variación en ambos casos fue debido a la gran variabilidad dentro de cada grupo más que entre los grupos (Tabla 3.4), hallándose el valor más alto en el grupo denominado Sierra Madre Occidental (38,57 y 33,36%).

**Tabla 3.4** Análisis de Varianza Molecular (AMOVA) de los 28 SSRs en las 38 poblaciones Mexicanas según los grupos formados en el análisis de agrupamiento y clasificación de acuerdo a complejos raciales de Goodman y Brown (1988). \*Las fuentes de variación son significantes a un nivel de  $P=0.05$

Fuente de variación	% de variación basado en análisis de agrupamiento	% de variación basado en complejos raciales
Entre grupos	19,31*	18,34*
Sierra Madre Occidental	38,56*	33,36*
Sur	15,85*	20,38*
Valles Centrales	26,28*	27,91*

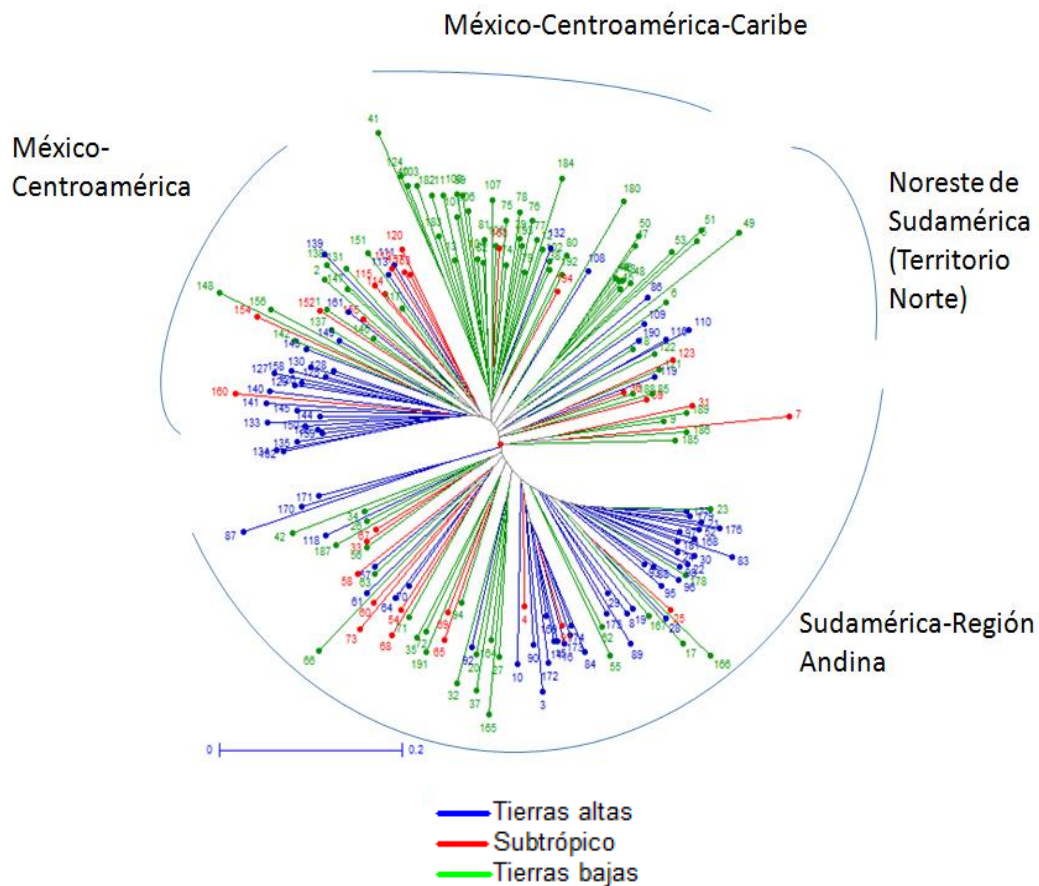
En general, para tener una comprensión de la distribución de las razas de maíz en México es necesario tener en cuenta las condiciones topográficas de las distintas regiones. Se pueden encontrar entonces regiones naturales bien definidas como la Sierra Madre Occidental, la Mesa Central y regiones al Sur de México como las llanuras costeras del Golfo, y del Pacífico, valles altos y planicies en Oaxaca y Chiapas, y esto ha quedado evidenciado con los análisis presentados en la validación de los marcadores en el sentido que diferenciaron los complejos raciales y los grupos de razas nativas según regiones geográficas específicas de México, generando un nivel de confianza adicional para emprender el análisis de la colección entera de nativos de toda Latinoamérica.

### 3.3.2 Asociaciones entre las 194 poblaciones de maíz latinoamericanas

Para tener un mejor entendimiento de la diversidad global en Latinoamérica, sus relaciones interpopulacionales y su potencial como fuente de alelos favorables para mejoramiento, es importante ampliar el número y la representación de diferentes fuentes de maíz en los análisis. Para lograr este objetivo se realizó un estudio a nivel Latinoamérica con 194 poblaciones diferentes. El análisis de agrupamiento (Figura 3.14) identificó los siguientes grupos: México-Centroamérica, México-Centroamérica-Caribe, Noreste de Sudamérica o Territorio Norte, y Región Andina.

En el ACoP (Figura 3.15), el Primer Componente representa el 8,4% de la varianza y el CP2 el 4,0%. La Figura 3.15a está coloreada según los grupos encontrados en el análisis de agrupamiento (Figura 3.14) mientras que la Figura 3.15b se encuentra coloreada según la elevación a la que se encuentra la zona de recolección, para evidenciar que existen asociaciones entre los maíces que provienen de tierras altas, subtrópico y trópico.

**Figura 3.14** Análisis de agrupamiento para 194 poblaciones de maíz identificadas con número ID, provenientes de México, Centroamérica, Caribe y Sudamérica generado con el método Neighbor-Joining y basado en la distancia genética de proporción de alelos compartidos para 28 SSRs. Los colores denotan las distintas altitudes de las zonas de recolección de germoplasma.



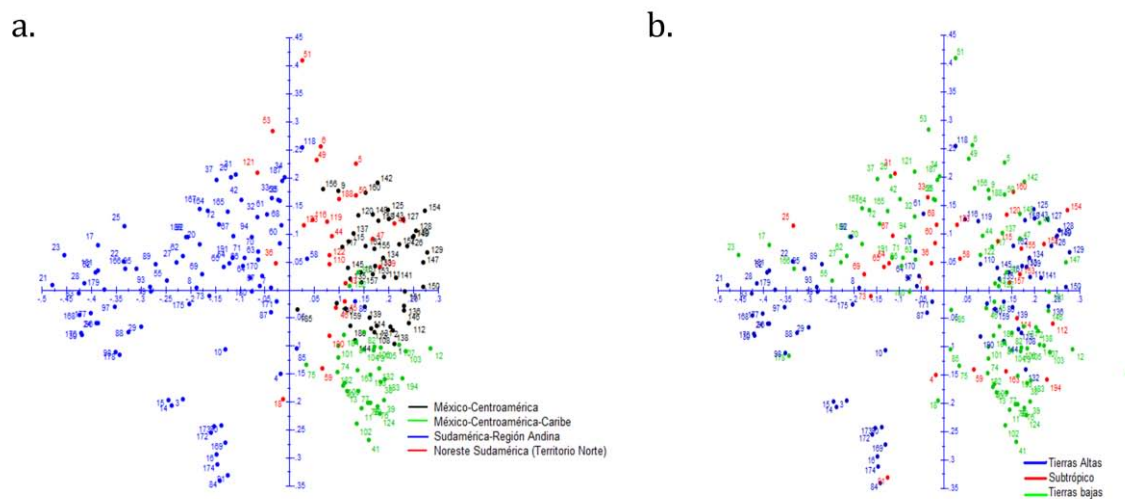
El grupo observado en la parte superior derecha (Figura 3.14), marcado como Noreste de Sudamérica, se asemeja en su constitución al grupo denominado por McClintock et al. (1981) “Territorio Norte”, al encontrarse poblaciones de razas diversas que albergan

características cromosómicas comunes según el estudio mencionado, estos materiales tienen una distribución amplia en el Caribe, México y Centroamérica. Entre estas poblaciones encontramos ECUA500 (ID 86) de la raza Canguil, también se encuentran relaciones con la raza chilena Pisankalla; otro ejemplo es la población BOLI711 de la raza Perola (ID 18) que según esta misma definición de grupo contiene sus componentes cromosómicos característicos. Adicionalmente estas poblaciones se encuentran relacionadas con las poblaciones de la costa oriental sudamericana encontrándose el grupo de Catetos provenientes de Brasil, Argentina y Uruguay (el término “cateto” es aplicado a variedades de maíz que presentan semejanzas con Cristalino Costeño Tropical del área del Caribe) y finalmente las poblaciones de las tierras bajas Chilenas (razas no determinadas) que podrían tener relación con algunas razas del norte de Argentina.

En el ACoP (Figura 3.15a) también se pueden ver estas relaciones entre las poblaciones de México-Centroamérica-Caribe con la costa oriental sudamericana marcadas con color rojo. Un gran grupo formado por 79 poblaciones (parte izquierda de la Figura 3.15) marcada como Región Andina y coloreada con azul en el ACoP de la Figura 3.15a, presenta 78 poblaciones Sudamericanas y una sola Centroamericana de la raza Negro de Tierra Fría (GUATEM 390). Este grupo está integrado por todas las poblaciones colombianas excepto por una de las dos poblaciones de la raza Amagaceño. Siguiendo con este grupo se encuentran en su totalidad poblaciones provenientes de Perú, Paraguay, y de Ecuador (exceptuando la población ECUA500 de la raza Canguil), las poblaciones Bolivianas (excepto BOLI711 de la raza Perola) y poblaciones de Brasil, Uruguay, Chile, Argentina y una de las cuatro Venezolanas (VENEZU 391 de la raza Cariaco).

El ACoP muestra que las razas consideradas como indígenas antiguas en México (Wellhausen et al., 1951) como Palomero Toluqueño, Arrocillo Amarillo, Chapalote y Nal-Tel) (Figuras 3.15), no se encuentran asociadas directamente con otras razas Sudamericanas consideradas como primitivas, como por ejemplo las razas Confite Morocho, Kully y Confite Puneño de Perú (Manrique, 1997), Pollo y Pira de Colombia (Roberts et al., 1957) o Avati’ Moroti y Avati’ Pichinga en Paraguay (Salhuana y Machado, 1999).

**Figura 3.15** Análisis de Coordenadas Principales para 194 poblaciones de maíz Latinoamericanas, identificadas con número ID, basado en la distancia genética de proporción de alelos compartidos para 28 SSRs. La Figura 3.15a está marcada según región geográfica basado en el Análisis de agrupamiento de la Figura 3.13. La Figura 3.15.b está marcada según la altitud de la zona de recolección.



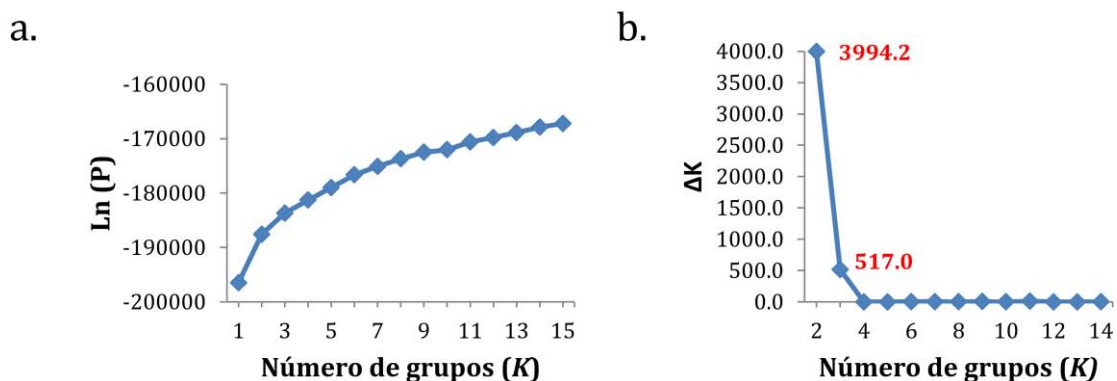
### 3.3.3 Análisis de estructura poblacional en Latinoamérica

Para investigar la estructura de población a nivel Latinoamérica se utilizó nuevamente el procedimiento reportado por Evanno et al. (2005), el valor más alto de  $\Delta K$  fue encontrado para un  $K=2$  (3.394,2), siendo aproximadamente 8 veces el siguiente valor más alto de  $\Delta K$  para  $K=3$  (517.0) (Figura 3.16). Para Latinoamérica se analizó en detalle los grupos cuando se asume un valor de  $K=3$  con el fin de estudiar la subestructura poblacional, corriendo de nuevo con los parámetros anteriormente definidos y de manera independiente el programa STRUCTURE, variando  $K$  desde 1 hasta 8 y tres corridas independientes para cada valor de  $K$ .

En el análisis de subestructura el valor más alto de  $K$  fue 2 en todos los casos (como se detallará para cada caso más adelante, (Figuras 3.16b, 3.20b, 3.22b y 3.24b), incluyendo los análisis anteriores cuando se analizaron las poblaciones mexicanas y todo el conjunto de germoplasma (Figura 3.12b y Figura 3.16b). Notablemente la metodología de Evanno et al. (2005) favorece siempre  $K=2$ . El máximo valor de  $\Delta K$  en  $K=2$  parece ser un artefacto que

resulta de la baja probabilidad para  $K=1$  en todos los casos. En el trabajo realizado por Vigourox et al. (2008), utilizando la metodología de Evanno et al. (2005) también se encontraron los valores más altos para  $\Delta K$  en  $K=2$  pero atribuyendo estos resultados a un artefacto causado debido al gran poder asociado al hecho de un elevado número de marcadores empleados (83). En el presente estudio, los marcadores fueron aproximadamente una tercera parte, pero tomando 15 individuos por población sugiriendo que el número de individuos estudiados provee potencia en el análisis para rechazar la hipótesis de que no existe estructura poblacional entre cualquiera de los conjuntos de germoplasma estudiados. De otro lado, el máximo valor de  $\Delta K$  en  $K=2$  en todos los casos, basado en la experiencia, siempre es notable porque indicaría el paso de no agrupamiento a un primer nivel de estructura (Jorge Franco, comunicación personal). Como el interés de este estudio es analizar los grupos de maíces en Latinoamérica y sus relaciones se analizaron en detalle los siguientes valores más altos encontrados para  $\Delta K$ .

**Figura 3.16** Gráficas del Ln de la probabilidad (a), y (b)  $\Delta K$ , del análisis de estructura poblacional de 194 poblaciones de maíz. Se graficaron los valores promedios de  $\ln[\text{prob}(\text{dato})]$  de 5 corridas independientes para cada valor de  $K$ , de la misma manera estos valores fueron utilizados para el cálculo de los valores  $\Delta K$ . Con color rojo están marcados los más altos valores encontrados para  $\Delta K$ .



Explorando los niveles de estructura para todo el germoplasma, y primero analizando el valor de  $K=2$ , el 98,0% de las poblaciones fueron asignadas a uno de los dos grupos. Tomando este valor, las poblaciones se podrían definir en dos grandes grupos, el primero como un grupo mixto que indicaría las relaciones básicas entre todos los maíces de México,

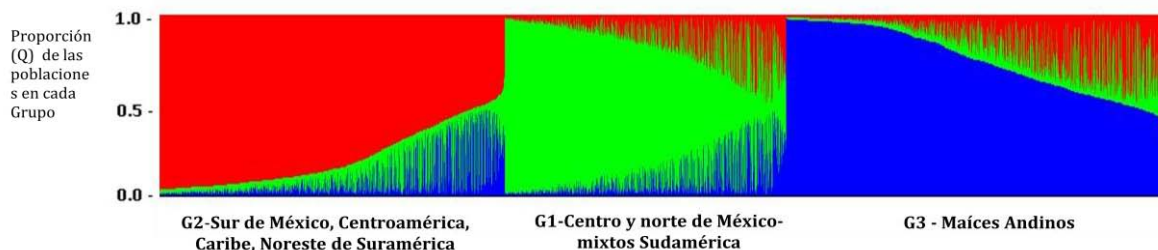


Caribe y Centroamérica con algunas razas sudamericanas como los catetos de Argentina y Uruguay, Tusones de Venezuela, y los maíces de las tierras bajas del sur de Chile. La representación de los maíces sudamericanos en este grupo es de un 31,0% de las 104 poblaciones en total provenientes de Sudamérica y un 27,0% del total de las poblaciones en este grupo. El segundo grupo se podría exponer como maíces andinos exceptuando por 2 poblaciones provenientes de Guatemala de las razas Salpor y Nal-tel.

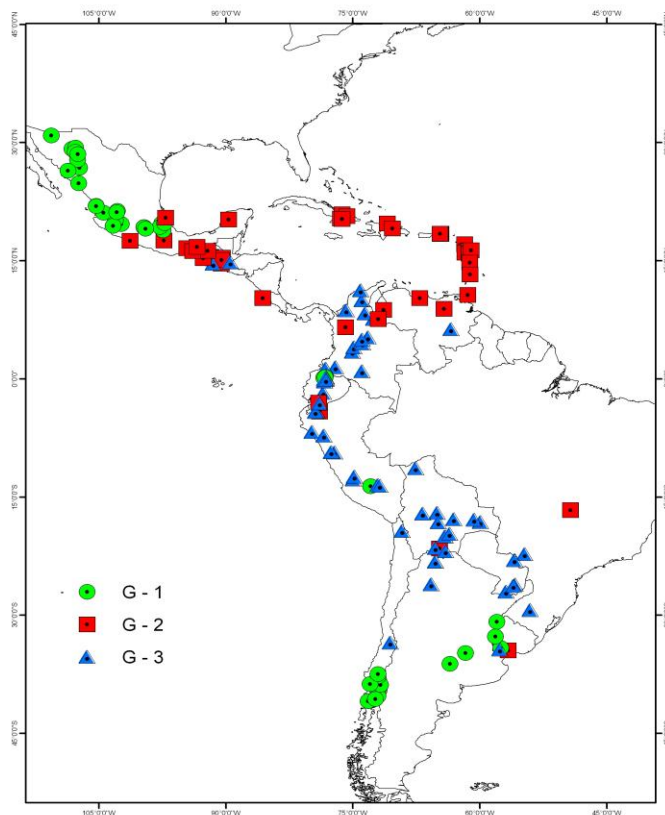
Posteriormente analizando la estructura poblacional admitiendo el valor de  $K=3$  (Figura 3.17), (denominados en adelante como G1, G2, y G3). En este caso el 87,0% de las poblaciones fueron asignadas a alguno de los tres grupos, las poblaciones restantes que no agruparon en alguno de los tres grupos anteriores se definió como G4. Con este nivel de estructura se pone en evidencia la particularidad y posible origen de los maíces caribeños porque en el grupo mixto mencionado anteriormente cuando se definió  $K=2$ , surge una división entre las poblaciones del centro y norte de México, tierras bajas del sur de Chile y Argentina, población de la raza Canguil de Ecuador, Confite Puneño y San Jerónimo de Perú y las razas Cateto Sulino y Dente Branco de Uruguay (G1) con las poblaciones del Sur de México, Guatemala, Corta Rica y Panamá, todas la poblaciones caribeñas, norte de Venezuela, las poblaciones de la raza Tusilla de Ecuador y la población de la raza Cateto Nortista de Brasil (G2). El tercer grupo aquí definido como maíces andinos sigue conservando en su gran mayoría el grupo que se había definido como maíces sudamericanos en el análisis anterior ( $K=2$ ) y las dos poblaciones de Guatemala (G3). De manera general, G1 representa a los maíces del centro y Norte de México y un subgrupo de mixtos sudamericanos; G2, Sur de México, Centroamérica, Caribe, Noreste de Suramérica y G3 a los maíces andinos (Figuras 3.17 y 3.18), Las proporciones de ancestro están definidas en la Tabla 3.5.

Las poblaciones no agrupadas (G4), poblaciones con ancestro diverso, es decir, poblaciones con proporciones por debajo del mínimo del 51,0% aceptado, son de procedencia diversa: Argentina, Bolivia, Brasil, Colombia, Ecuador, Guatemala, México, Perú y Uruguay. En este caso se determinó la proporción de ancestro respecto a G1, G2 y G3 en los países que tuvieron 3 o más poblaciones (Tabla 3.6).

**Figura 3.17** Representación gráfica de la estructura de población estimada para 194 poblaciones de maíz Latinoamericanas (G1, G2 y G3). El valor más alto de  $\ln[\text{prob}(\text{dato})]$  de las corridas independientes fue escogido para la representación gráfica de cada agrupamiento  $K$ . Proporciones ( $Q$ ) en un rango de 0-1 de las poblaciones en los grupos y de los subgrupos en los grupos.



**Figura 3.18** Distribución geográfica de las poblaciones de nativas de maíz ( $Q > 51,0\%$ ) basada en el análisis de estructura de población donde fueron identificados tres grupos principales: Centro y norte de México – mixtos sudamericanos (G1); Sur de México, Centroamérica, Caribe, noreste de Sudamérica (G2) y maíces Andinos (G3).



**Tabla 3.5.** Media de las proporciones de ancestro para uno de los grupos encontrados en la estructura poblacional ( $K=3$ ). (Porcentaje mínimo aceptado de ancestro en todos los casos 51,0%). Están remarcados los valores mayores de ancestro dentro de cada grupo.

Proporción de ancestro ( $Q$ )			
	G1	G2	G3
G1	<b>0,79</b>	0,12	0,14
G2	0,12	<b>0,81</b>	0,09
G3	0,08	0,07	<b>0,76</b>

**Tabla 3.6** Media de las proporciones de ancestro por país para poblaciones no agrupadas (G4)

G4				
ancestro ( $Q$ )	Argentina	Brasil	Colombia	Guatemala
G1	0,36	0,27	0,31	0,36
G2	0,37	0,27	0,32	0,37
G3	0,27	0,46	0,36	0,26

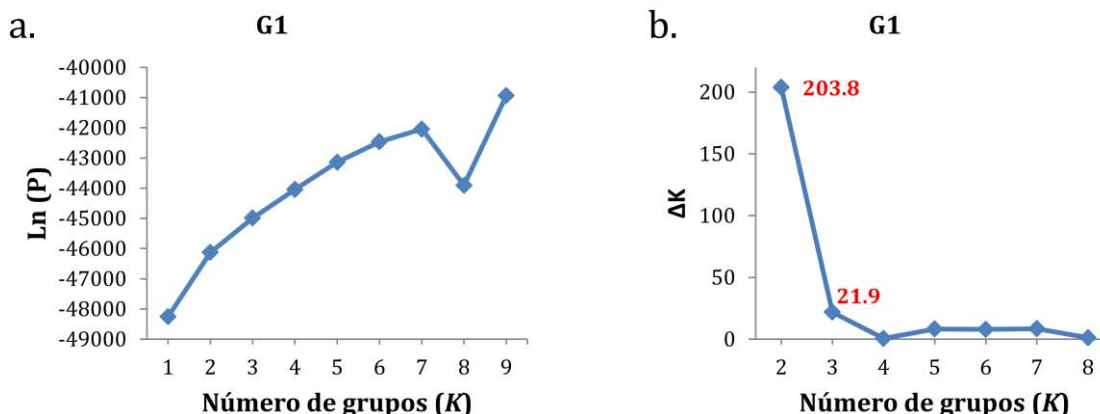
### 3.3.4 Análisis de estructura subpoblacional en Latinoamérica

Para obtener una mejor visión de la diversidad a nivel regional, un nivel necesario para entender las introgresiones y la evolución a nivel de las diferentes poblaciones debido a la intervención humana se analizó la estructura de las poblaciones dentro de cada grupo definido anteriormente (G1, G2 y G3). De esta manera, y utilizando el mismo procedimiento (Evanno et al., 2005), se definió la estructura subpoblacional en cada uno de los grupos. A continuación se describe en detalle para G1, G2 y G3 la definición del número de subgrupos ( $K$ ), la composición de cada subgrupo y la distribución geográfica de las poblaciones.

#### G1- Centro y norte de México- mixtos Sudamérica

Este grupo posee una media de proporción de ancestro de 0,79, 0,12 de ancestro proveniente G2 y 0,08 de G3 (Tabla 3.5). En el análisis de subestructura (Figura 3.19a y 3.19b) se asumió un  $K=3$ , esto significa 3 subgrupos. La distribución geográfica para estos subgrupos se encuentra representada en la Figura 3.20.

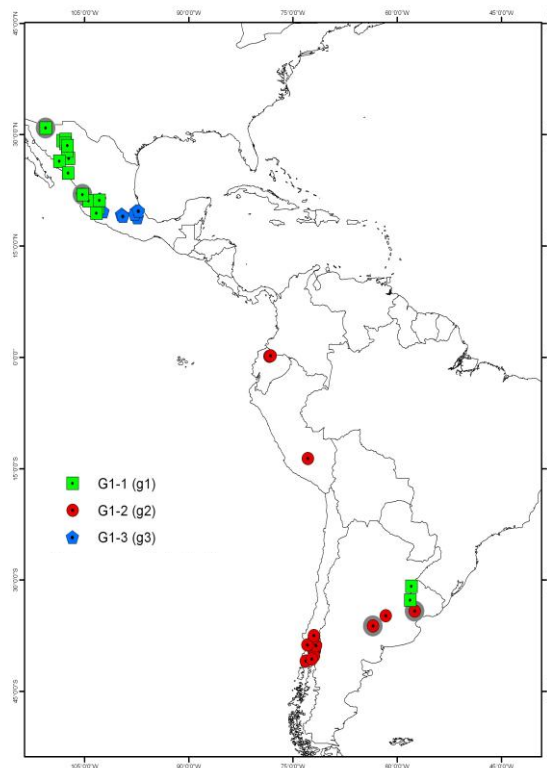
**Figura 3.19** Gráficas del Logaritmo de la probabilidad (a) y  $\Delta K$  (b) para el análisis de subestructura poblacional de Latinoamérica para G1. Se graficaron los valores promedios de  $\ln[\text{prob}(\text{dato})]$  de 3 corridas independientes para cada valor de  $K$ , de la misma manera estos valores fueron utilizados para el cálculo de los valores  $\Delta K$ . Con color rojo están marcados los valores más altos encontrados para  $\Delta K$ .



El subgrupo g1 está especialmente constituido por maíces del norte de México identificados como “Sierra de Chihuahua” en estudios previos (Sánchez et al., 2000) y que comprende las razas Azul, Cristalino de Chihuahua y Gordo. También, se encuentran las razas restantes definidas en el complejo de la Sierra Madre Occidental como Onaveño, Elotes Occidentales, Reventador, Tabloncillo y Jala. Adicionalmente conforman este subgrupo dos poblaciones sudamericanas de Argentina y Uruguay de la raza Camelia y Dente Branco respectivamente. Se presume que la raza Dente Branco es una introducción reciente de alrededor 100-150 años atrás desde Estados Unidos de América a territorios como Brasil y Uruguay y que ha sido mantenida localmente con pequeñas introgresiones de los maíces nativos (Paterniani y Goddman 1977), la presencia de esta raza en este subgrupo respalda de manera indirecta las relaciones entre los maíces del Norte de México y Estados Unidos (Matsuoka et al., 2002b; Vigourox et al., 2008) no presentes en este estudio.

El subgrupo g2 está formado por 17 poblaciones sudamericanas. A simple vista este grupo podría ser impreciso o de difícil definición, sin embargo, podría representar dos diferentes clases de eventos de introducción. Ocho de las once poblaciones chilenas presentes en este subgrupo son de las tierras bajas; lamentablemente son poblaciones sin clasificación en razas; también se encuentran Canario de Ocho de Argentina, Canguil de Ecuador, Confite Puneño y San Jerónimo de Perú y Cateto Sulino de Uruguay.

**Figura 3.20** Distribución geográfica de las poblaciones del grupo G1 de acuerdo al análisis de subestructura poblacional. Las poblaciones sombreadas con gris son las poblaciones menos representativas del subgrupo ( $51,0\% < Q > \%71,0$ )



Los eventos de introducción mencionados serían, primero, referente a la introducción de germoplasma proveniente de Estados Unidos a Chile, que a diferencia del resto de los países sudamericanos posee una gran cantidad de germoplasma de los Estados Unidos, como por ejemplo, la raza Araucano (presente en este estudio) que tiene similitud a los maíces de maduración temprana harinosos y cristalinos de ese país (Timothy et al., 1961). La raza Canario de Ocho de Argentina sigue presentado la misma dificultad que en estudios previos para esclarecer sus orígenes, pues se presume de un lado que es resultado de la introducción desde Estados Unidos y por otro lado podría ser el resultado de la selección a partir de Cateto Sulino (Paterniani y Goodman 1977). Cateto Sulino a su vez, presente en este mismo subgrupo, tendría relación con razas antiguas que presumiblemente han crecido y han sido mantenidas por habitantes que vivieron en el Sur de la costa Atlántica Sudamericana.

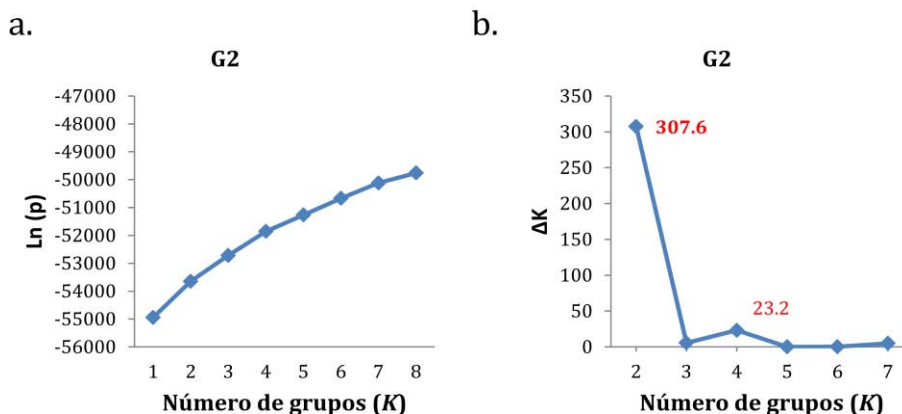
Como aspecto particular y una segunda característica de este grupo es la presencia de poblaciones de maíces primitivos reventadores o palomeros de las tierras altas, como las razas Canguil de Ecuador, Confite Puneño de Perú que podrían referir las antiquísimas relaciones con los maíces palomeros de México si se analiza en el contexto de grupo (G1) y no del subgrupo particular al que pertenecieron (g2).

El subgrupo g3 está constituido por los maíces del centro de México y representado por las razas, Palomero Toluqueño, Cacahuacintle, Cónico, Cónico Norteño, Mushito, Arrocillo Amarillo, mientras que la raza Harinoso de Ocho perteneciente al complejo del Occidente en análisis anteriores, no está representada.

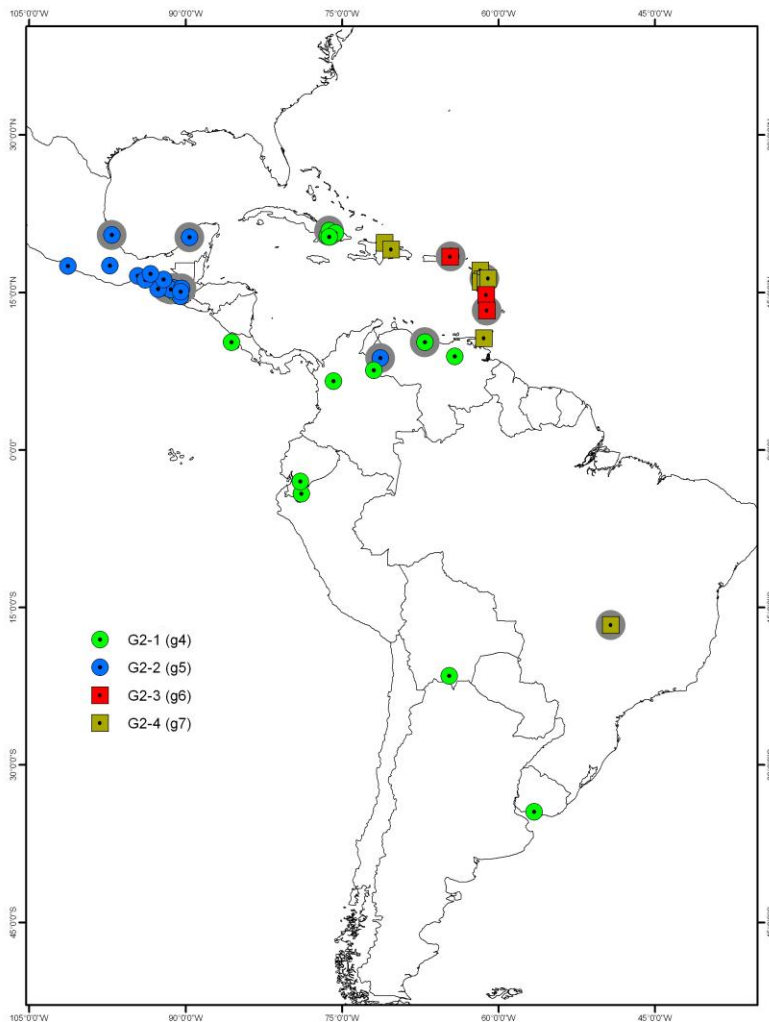
### G2- Sur de México, Centroamérica, Caribe y Territorio Norte

El G2 está compuesto por maíces del sur de México, Centroamérica, Caribe y poblaciones del grupo determinado anteriormente como Territorio Norte. Cuenta con una media de proporción de ancestro de 0,81, a ncestro de 0,12 de G1 y 0,07 de G3 (Tabla 3.5). Se encontraron 4 subgrupos: g4, g5, g6, g7, y se definió como g8 a las poblaciones que mostraron ancestro diverso. (Figura 3.21a 3.21b), La distribución geográfica de estos subgrupos está detallada en la Figura 3.22.

**Figura 3.21** Gráficas del Logaritmo de la probabilidad (a) y  $\Delta K$  (b) para el análisis de subestructura poblacional de Latinoamérica para G2. Se graficaron los valores promedios de  $\ln[\text{prob}(\text{dato})]$  de 3 corridas independientes para cada valor de  $K$ , de la misma manera estos valores fueron utilizados para el cálculo de los valores  $\Delta K$ . Con color rojo están marcados los valores más altos encontrados para  $\Delta K$ .



**Figura 3.22** Distribución geográfica de las poblaciones del grupo G2 de acuerdo al análisis de subestructura poblacional ( $Q > 51,0\%$ ), las poblaciones sombreadas con gris son las poblaciones menos representativas de los subgrupos ( $51,0\% < Q < 71,0\%$ )



El subgrupo g4 está formado por razas de alturas bajas e intermedias de Sudamérica como Perola de Bolivia, Amagaceño de Colombia, Tusilla de Ecuador, Cristalino Cubano y Tusón de Venezuela, Cateto Sulino de Uruguay; también constituyen este subgrupo poblaciones de las razas Cristalino Cubano de Islas Vírgenes y Cuba; también de Cuba, las razas Chandelle y Canilla; del archipiélago de Guadalupe la raza Caribeño Precoz, proveniente de Centroamérica Salvadoreño de Costa Rica y Quicheño Precoz de Guatemala.

El subgrupo g5 ilustra las relaciones entre los maíces del sur de México y los maíces guatemaltecos explicadas y sugeridas por los análisis anteriores (Figuras 3.14 y 3.15), además de las razas Costeño Tropical de las Islas Vírgenes y la raza Sabanero de Venezuela. El subgrupo g6 está constituido por cuatro poblaciones de las razas Caribeño Precoz de las Islas Vírgenes y Martinica, además de la raza Cristalino Costeño Tropical de la Isla de San Vicente.

El subgrupo g7 está integrado por las razas Cristalino Costeño Tropical de Antigua y República Dominicana, poblaciones de la raza Chandelle de República Dominicana y Guadalupe, Tusón de Trinidad y Tobago de la parte continental por la raza Cateto Nortista de Brasil. En el subgrupo g8 se encuentran algunas poblaciones de las razas como Cristalino Costeño Tropical, Tusón, Cristalino Cubano, Caribeño Precoz, y Salvadoreño-Clavillo de Barbados, Cuba, Guadalupe, Panamá y San Vicente, que se agruparon en alguno o en diferentes subgrupos arriba mencionados

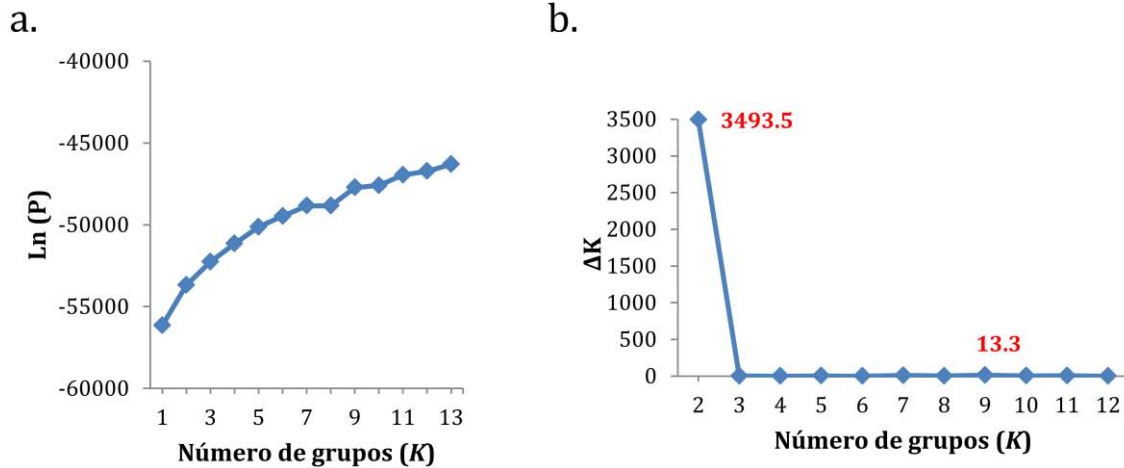
### **G3 - Maíces andinos**

El grupo de maíces andinos posee una media de proporción de ancestro de 0,76, un 0,14 proveniente de G1 y 0.09 de G2 (Tabla 3.5). En el análisis de subestructura para este grupo se encontraron 9 subgrupos (Figuras 3.23) y la distribución geográfica del análisis de subestructura se encuentra representada en la Figura 3.24.

El primero de ellos, g9, corresponde a maíces típicos de los Andes Centrales, compuesto por razas de tierras altas de Bolivia, Perú y Norte de Chile (Chechi, Chuspiilla, Uchuquilla, Kulli, Hualtaco Colorado- exceptuando por la raza Hualtaco de las tierras bajas bolivianas, Cuzco Cristalino Amarillo, Chulpi, Kculli de Perú y la raza Chulpi del norte Chile. El subgrupo g10 solo está conformado por poblaciones colombianas, de las razas Pira y Pira Naranja, Yucatán, Guira, Andaqui, Sabanero, Negrito, Cariaco y Cacao. El subgrupo g11 contiene razas del trópico y subtropical de Brasil (Cateto Assis), Colombia (Pollo, Puya, Puya Grande) y de Uruguay la raza Cateto Sulino. El subgrupo g12 está compuesto por razas de las tierras altas de Argentina, Bolivia, Ecuador y Perú (Cateto Amarillo, Capia Blanca, Altiplano, Perola, Morocho, San Jerónimo-Huancavelicano y Confite Morocho.



**Figura 3.23.** Gráficas del Logaritmo de la probabilidad (a) y  $\Delta K$  (b) para el análisis de subestructura poblacional de Latinoamérica para G3. Se graficaron los valores promedio de  $\ln[\text{prob}(\text{dato})]$  de 3 corridas independientes para cada valor de  $K$ , de la misma manera estos valores fueron utilizados para el cálculo de los valores  $\Delta K$ . Con color rojo están marcados los valores más altos encontrados para  $\Delta K$ .

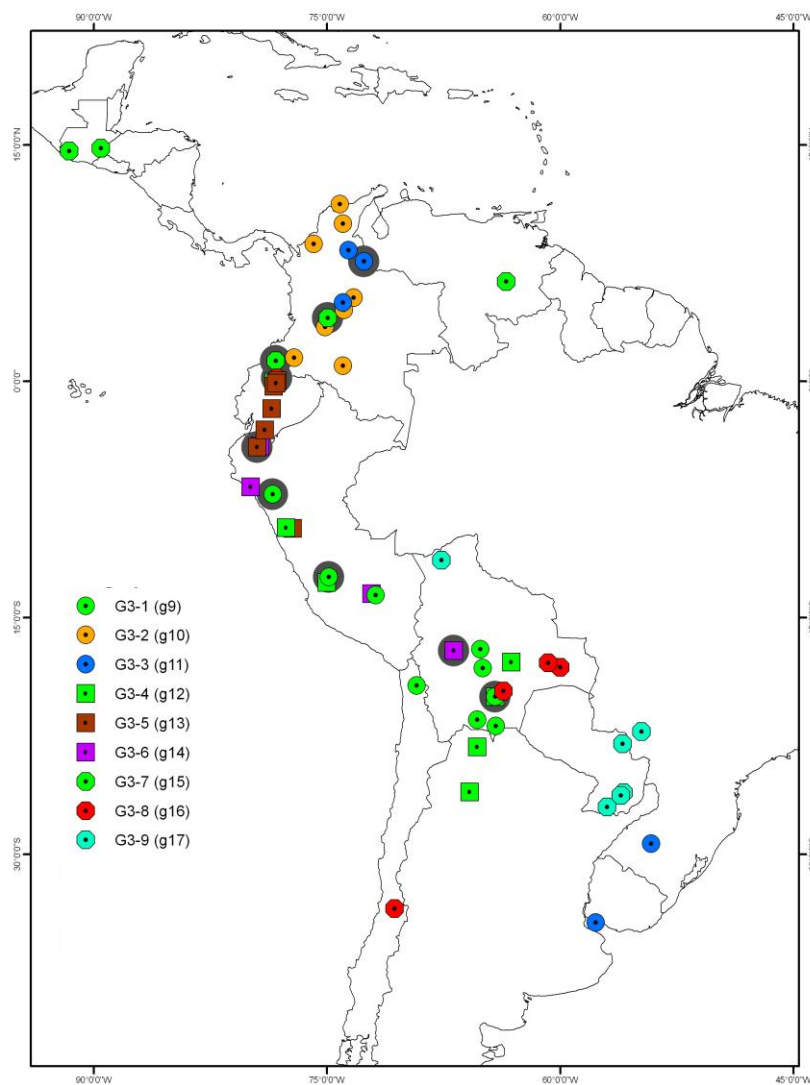


Los maíces pertenecientes al subgrupo g13 son especialmente de las tierras altas del norte del Complejo Andino, Ecuador (Racimo de Uva, Chillo, Uchima, Chulpi, Morocho, Shima, Mishca) y la población de la raza Cuzco de Perú. El subgrupo g14 corresponde a 5 poblaciones de los Andes centrales de Argentina, Bolivia, Ecuador y Perú representado a razas típicas de las tierras altas (Culli, Paru, Kcello, Cuzco) con la excepción de la raza Mochero del trópico peruano.

El subgrupo g15 está compuesto por poblaciones de las razas Pira, Chococeño, Montaña de Colombia; Salpor y Nal-tel Amarillo de las tierras bajas de Guatemala y Cariaco de Venezuela. La presencia de las dos poblaciones provenientes de Guatemala en este subgrupo llama particularmente la atención por varias razones, si se piensa en el origen de los maíces andinos o en las relaciones o contrapartes de algunos maíces centroamericanos y sudamericanos. La primera razón, porque apoya las relaciones previamente sugeridas entre las razas Nal-tel y las Sudamericanas, aunque la raza Nal-tel se ha diversificado tanto que se encuentran identificadas 5 subrazas; la relación de Nal-tel en especial con la raza Pira de Colombia, ambas consideradas como primitivas, y ésta a su vez puede relacionarse con razas de Perú como la primitiva Confite Morocho de las tierras altas (Wellhausen et al.,

1957; Roberts et al., 1957). La segunda es que se puede inferir una relación indirecta entre la otra raza de Guatemala, Salpor, con la raza Montaña de Colombia sugiriendo la introducción de diferentes maíces sudamericanos a Centroamérica. La presencia de Chococeño en este subgrupo podría sustentar la influencia de esta raza en algunas de las razas colombianas (Roberts et al., 1957).

**Figura 3.24** Distribución geográfica de las poblaciones del grupo G3 de acuerdo al análisis de subestructura poblacional ( $Q > 51,0\%$ ), las poblaciones sombreadas con gris son las poblaciones menos representativas de los subgrupos ( $51,0\% < Q < 71,0\%$ )



El subgrupo g16 contiene a razas de las tierras bajas de Bolivia, como: Blando Amazónico, Pojoso Chico, Acre Entrelazado y Camelia de Chile; la excepción en este grupo es la raza Montaña de las tierras altas de Ecuador. El subgrupo g17 contiene razas de la tierras bajas de la parte este del centro de Sudamérica, como Coroico, Acre Entrelazado de Bolivia, Cateto Grande de Brasil, Avati' Moroti' Ti' y Avati' Pichinga, Avati' Moroti' y Avati' de Paraguay. Las razas Moroti' son consideradas primitivas y producto de selección elaborada por los aborígenes guaraníes (Salhuana y Machado, 1999), se presume además que la raza Avati' Moroti' pudo haber dado origen a la raza Coroico presente en este subgrupo. Se puede particularizar que este subgrupo es una conjunción formada por razas consideradas de la Cuenca Amazónica (Goodman y Brown., 1988) y Catetos de Brasil siendo el grupo de Catetos uno de los más difundidos por su considerable grado de adaptación a aéreas específicas (Paterniani y Goodman, 1977). Finalmente, el subgrupo g18 contiene tres poblaciones que no se agruparon en ninguno de los subgrupos anteriores: Aysuna y Duro Amazónico de Bolivia y Ancashino de Perú.

#### **G4 – poblaciones mixtas no agrupadas**

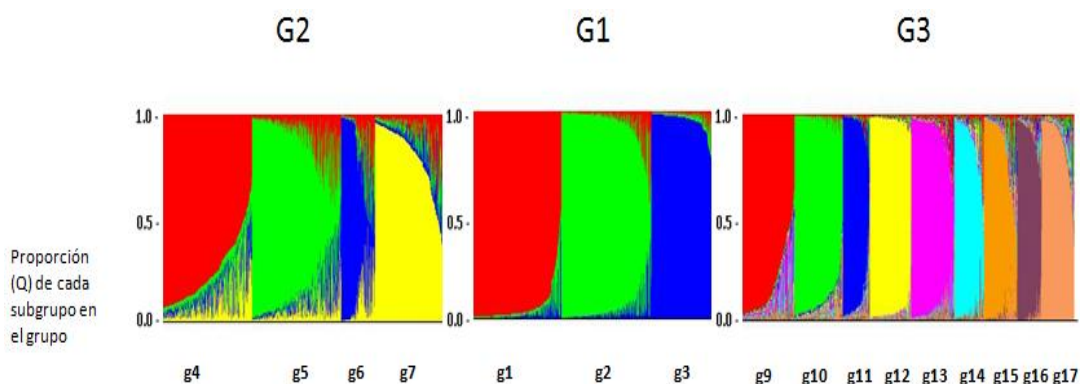
Esta conjunción de poblaciones mixtas no agrupadas da indicios sobre la complejidad de las relaciones entre los maíces sudamericanos y las culturas que pudieron estar vinculadas a los movimientos migratorios. Las proporciones de ancestro para las razas de Guatemala indican la influencia de los maíces del centro y norte de México (G1) y los maíces aquí catalogados como andinos (G3). Para Colombia, que por su especial situación geográfica pudo haber recibido introducciones simultáneas o en diferentes tiempos por las costas Pacífica y Atlántica, además de posibles rutas continentales por Panamá, presentan ancestro diverso y en similares proporciones de G1, G2 y G3 (Tabla 3.6).

Las razas dentadas de Brasil así como la raza Cateto Paulista (Tabla 3.7) muestran ancestro diverso, los maíces dentados en Brasil son considerados introducciones probablemente de Estados Unidos y que han ganado gran aceptación con sus altos rendimientos cuando se comparan con los Catetos (Paterniani y Goodman, 1977). En Argentina las razas Cuarento Cateto, Cristal Sulino y Cateto muestran también ancestro diverso pero en menor proporción proveniente de los maíces Andinos (G3). En este grupo también se encuentran

las razas Bolita y Olotillo del Sur de México, Uchuquilla de Bolivia y Cuarento- Cateto de Uruguay.

Con el objetivo de tener una síntesis de la información generada en los análisis de estructura y subestructura se expone la representación grafica de la subestructura de población para G1, G2 y G3 (Figura 3.25), y en la Tabla 3.7 se describen las razas de maíz presentes en cada subgrupo y países de procedencia según estos resultados para todo el germoplasma estudiado.

**Figura 3.25** Representación gráfica de la subestructura poblacional (g1, g2, g3, g4, g5, g6, g7, g9, g10, g11, g12, g13, g14, g15, g16 y g17) estimada para cada uno de los grupos G1, G2 y G3. El valor más alto de  $[\ln[\text{prob}(\text{dato})]]$  de las corridas independientes fue escogido para la representación gráfica de cada agrupamiento  $K$ . Proporciones ( $Q$ ) en un rango de 0-1 de las poblaciones en los grupos y de los subgrupos en los grupos.



### 3.3.5 Comparaciones genéticas

Tomando ventaja de que las distancias genéticas y la diversidad pueden ser analizadas a diferentes niveles jerárquicos, en esta sección se hace un análisis de las distancias genéticas a nivel de las poblaciones nativas mexicanas y a nivel Latinoamérica; similarmente se hace un disección de la diversidad total encontrada teniendo en cuenta los grupos de poblaciones definidos por el análisis de estructura poblacional.

**Tabla 3.7** Composición de subgrupos basados en análisis de subestructura poblacional. Se presentan los nombres de las razas, entre paréntesis se muestra si hay más de una accesión y las iniciales del país de procedencia (An-Antigua, Ar-Argentina, Bar-Barbados, Bol-Bolivia, Br-Brasil, IsVir-Islas Vírgenes, Ch-Chile, Col-Colombia, CR-Costa Rica, Cu-Cuba, RDom-Republica Dominicana, Ec-Ecuador, Guad-Guadalupe, Guat-Guatemala, Mex-México, Pan-Panamá, Py-Paraguay, Pe-Perú, SVic-San Vicente, Tri-Trinidad y Tobago, Ur-Uruguay, Ven-Venezuela).

### G1- Centro y norte de México- mixtos Sudamérica

- 
- g1** Camelia(Ar); Gordo,Palomero Toluqueño > Cristalino de Chihuahua, Azul(2) > Cristalino de Chihuahua, Cristalino de Chihuahua>Apachito(2), Chapalote(2), Onaveño, Jala, Elotes Occidentales, Reventador, Gordo> Cristalino de Chihuahua, Tabloncillo, Maíz Dulce(Mex); Dente Branco(Ur)
- g2** SC(1), Canario de Ocho (Ar); SC(8), Pisankalla, Ocho Corridas, Araucano (Ch); Canguil(2) (Ec); Confite Puneño, San Jerónimo (Pe); Cateto Sulino (Ur)
- g3** Palomero Toluqueño, Cacahuacintle(2), Chalqueño, Mushito, Harinoso de Ocho, Cónico(3), Arrocillo Amarillo(2), Cónico Norteño (Mx)

### G2 - Sur de México, Centroamérica, Caribe, noreste de Suramérica

- 
- g4** Perola (Bol); Cristalino Cubano (IsVir); Amagaceño (Col); Salvadoreño(2) (CR); Cristalino Cubano (2), Canilla, Chandelle (Cu); Tusilla (2)(Ec); Caribeño Precoz (Guad); Quicheño precoz (Guat); Cateto Sulino (Ur); Cristalino Cubano, Tusón(2) (Ve)
- g5** Cristalino Costeño Tropical(IsVir); San Marceño, Olotón, Quicheño Precoz, Comitico, Olotón > Nal-Tel, Negro de Chimaltenango, Nal-Tel Blanco Tierra Baja (Guat); Zapalote Grande, Zapalote Chico (3), Comitico, Nal-Tel>Harinoso, Tuxpeño, Tepecintle, Vandeño (Mex); Sabanero (Ve)
- g6** Caribeño Precoz (2) (IsVir); Caribeño Precoz (Mar); Cristalino Costeño Tropical (SVic)
- g7** Cristalino Costeño Tropical(An); Cateto Nortista (Br), Chandelle (2) (RDom); Chandelle (2), Cristalino Costeño Tropical(2) (Guad); Tusón (Tri)
- g8\*** Cristalino Costeño Tropical (2), Tusón (Bar); Cristalino Cubano (Cu), Caribeño Precoz(2) (Guad); Salvadoreño-Clavillo (Pan); Cristalino Costeño Tropical (SnVic)

### G3 - Maíces Andinos

- 
- g9** Chechi, Hualtaco, Chuspilla- Chuspilla, Uchuquilla, Kulli, Hualtaco Colorado (Bol); Chulpi (Ch); Cuzco Cristalino Amarillo, Chulpi, Kculli (Pe)
- g10** Pira Naranja, Yucatán, Guira, Pira, Andaqui, Sabanero, Negrito, Cariaco, Cacao (Col)
- g11** Cateto Assis (Br); Pollo, Puya Grande, Puya (Col); Cateto Sulino (Ur)
- g12** Cateto Amarillo, Capia Blanca (Ar); Altiplano, Perola (Bol); Morocho (2) (Ec); San Jeronimo-Huancavelicano, Confite Morocho (Pe)
- g13** Racimo de Uva, Chillo, Uchima, Chulpi, Morocho, Shima, Mishca (Ec); Cuzco (Pe)
- g14** Culli (Ar); Paru (Bol); Kcello (Ec); Cuzcri, Mochero, (Pe)
- g15** Pira, Chococeño, Montaña (Col), Salpor, Nal-Tel Amarillo de Tierra Baja (Guat); Cariaco (Ve)
- g16** Blando Amazónico, Pojoso Chico, Acre Inter (Bol); Camelia (Ch); Montaña (Ec)
- g17** Coroico(Bol); Cateto Grande (Br) ; Avati' Moroti' Ti', Avati' Pichinga, Avati' Moroti', Avati (Py)
- g18\*** Aysuna (Bol); Duro Amazónico, (Bol); Ancashino, (Pe);

### G4 (poblaciones no agrupadas)

---

Cuartento Cateto, Cristal Sulino, Cateto (Ar); Uchuquilla (Bol), Cateto Paulista Grosso, Dente Paulista, Dente Riograndense, Dentado (Br); Amagaceño, Pira (2), Cabuya (Col); Chococeño (Ec); Negro de Chimaltenango, Nal-Tel Blanco Tierra Alta, Olotón, Quicheño Precoz, Negro de Altura-Negro de Tierra Fria, San Marceño (Guat); Bolita, Olotillo (Mex); Uchuquilla, Huancavelicano, Perla (Pe); Cuarentón-Cateto Colorado (Ur)

---

### **3.3.5.1 Análisis distancias genéticas**

La máxima distancia genética basada en la proporción de alelos compartidos, calculada con el programa PowerMarker entre las 38 poblaciones mexicanas (Anexo 7.8), se encontró entre las poblaciones PUEB91 y SINA2 de las razas Arrocillo Amarillo y Chapalote respectivamente, con un valor de 0,58. Las siguientes distancias genéticas entre poblaciones mexicanas en orden decreciente son: 0,57 entre las poblaciones de MEX5 (la ID133) de raza Palomero Toluqueño y la población PI217413 (ID139) de la raza Zapalote Chico; la distancia de 0,56 entre la población SINA2 de la raza Chapalote (ID 148) y la población PI217413 (ID139).

Examinado la matriz de distancias genéticas para las 194 poblaciones totales (resultados no mostrados) el valor máximo de distancia genética (0,73) se encontró entre las poblaciones BRV1142 de las Islas Vírgenes de la raza Caribeño Precoz y CHZM08060 de la raza Pisankalla de Chile. La primera es típica del Caribe y la segunda de las tierras bajas chilenas, pertenecientes de acuerdo al análisis de estructura a los grupos G1 y G2 respectivamente. La mínima distancia genética (0,18) fue encontrada para las poblaciones BOLI905 de la raza Altiplano de Bolivia, y para ANC333 de la raza San Jerónimo-Huancavelicano de Perú, ambas típicas de las tierras altas de los Andes y pertenecientes según el análisis de estructura y subestructura poblacional al grupo de maíces andinos (G3) y subgrupo compuesto por maíces de las tierras altas de Ecuador, Perú, Bolivia y norte de Argentina (g12).

### **3.3.5.2 Análisis de diversidad genética**

Aunque la distancia genética es una buena estimación de la diversidad genética, otro tipo de análisis es necesario para completar el análisis y dar una visión más completa de esta diversidad. Por esta razón se estimaron los índices de fijación  $F_{ST}$  por parejas para los grupos de poblaciones definidas de acuerdo al análisis de estructura poblacional (Tabla 3.8). De acuerdo con los intervalos para la interpretación de los valores de  $F_{ST}$  sugeridos por Wright (1965), descritos anteriormente, se encontraron dos niveles de diferenciación genética entre grupos. Entre el grupo de centro y norte de México – mixtos Sudamérica (G1) y sur de México, Centroamérica, Caribe, noreste de Sudamérica (G2) se encontró un

nivel bajo de diferenciación genética (0,04) en un rango de 0,0 a 0,05 (Tabla 3.8). La diferenciación genética entre G1 y G3, G2 y G3 fue moderada (valores en el rango 0.05-0.15).

**Tabla 3.8** Valores de Índices de Fijación  $F_{ST}$  calculados por parejas entre los grupos de poblaciones clasificados de acuerdo al análisis de estructura poblacional.

	<b>G1</b>	<b>G2</b>
<b>G2</b>	0,04	
<b>G3</b>	0,08	0,09

Para las 194 poblaciones, con un total de 2.910 genotipos simulados, se calcularon los valores de diversidad genética, media de alelos por locus, alelos únicos entre grupos y dentro de subgrupos determinados en los análisis de estructura y subestructura poblacional (Tabla 3.9). El promedio de la diversidad genética a través de todas las 194 accesiones fue de 0,62. El número total de alelos detectados para los 28 SSRs en las 194 poblaciones fue 291 con un rango de alelos por locus de 2 hasta 21, con un promedio de 10,39, todos los 28 marcadores fueron polimórficos. El mayor índice de diversidad dentro de cada grupo fue encontrado para G1- Centro y norte de México- mixtos Sudamérica (0,63) y el menor para G3 - maíces Andinos (0,57). El mayor número de alelos únicos por grupo se obtuvo en los maíces Andinos (24) seguido por G1 con 15 y G2 con 14 alelos.

La variabilidad genética, presente a través de todo el germoplasma evaluado, también puede describirse como la distribución de esta variabilidad entre grupos y dentro de cada grupo de acuerdo al análisis de estructura. Con el análisis de varianza molecular detallado en la Tabla 3.10, se encontró que la variación dentro de los grupos G1 y G2 fue mayor a la variabilidad entre grupos. El grupo de los maíces del centro y norte de México y mixtos sudamericanos (G1) tiene la mayor variabilidad genética (30,56%) y los maíces Andinos (G3) mostraron la menor variabilidad intragrupal (18,81%).

En general, la distribución de la diversidad genética de las razas de maíz está principalmente dividida dentro de grupos poblacionales que representan complejos

raciales y menos entre grupos (Warburton et al., 2008) quedó comprobada también en este estudio. Para los maíces andinos, aunque pudiera parecer contradictorio: menor índice de diversidad genética pero contiene el mayor número de alelos únicos encontrados, este hecho puede respaldarse con la información generada con el AMOVA (Tabla 3.10), pues los maíces andinos poseen la menor variabilidad genética comparada intra e intergrupalmente.

**Tabla 3.9** Alelos encontrados y diversidad genética fueron calculados como el promedio para cada grupo y subgrupo según análisis de estructura poblacional. Número de alelos únicos \* comparados entre grupos, \*\* comparados entre subgrupos de un mismo grupo.

Grupos en estructura	Grupos en subestructura	Número de poblaciones	Alelos encontrados	Número de alelos únicos	Diversidad genética
<b>G1</b>		46	227	15*	0,63
	g1	17	192	24**	0,60
	g2	17	173	15**	0,61
	g3	12	176	11**	0,61
<b>G2</b>		58	221	14*	0,59
	g4	17	172	10**	0,56
	g5	18	191	21**	0,58
	g6	4	112	5**	0,47
	g7	11	150	4**	0,56
	g8-mixtas no agrupadas	8	145	1**	0,57
<b>G3</b>		65	226	24*	0,57
	g9	10	127	4**	0,46
	g10	9	143	6**	0,56
	g11	5	115	3**	0,55
	g12	8	127	8**	0,48
	g13	8	117	5**	0,47
	g14	5	120	3**	0,52
	g15	6	119	3**	0,53
	g16	5	116	8**	0,48
	g17	6	114	2**	0,49
	g18-mixtas no agrupadas	3	101	1**	0,45
<b>G4</b> Poblaciones mixtas no agrupadas		25	202	10*	0,61
<b>Total</b>		194	291	no aplica	0,62



**Tabla 3.10** Análisis de Varianza Molecular (AMOVA) de los 28 SSRs de acuerdo al análisis de estructura poblacional para los grupos compuestos por los maíces del centro y norte de México y mixtos sudamericanos (G1), Sur de México, Centroamérica, Caribe, Noreste de Suramérica (G2) y los maíces Andinos (G3). \*Las fuentes de variación son significantes a un nivel de  $P = 0.05$

<b>Fuente de variación</b>	<b>% de variación</b>
Entre grupos	24,82*
G1	30,56*
G2	25,80*
G3	18,81*

### **3.4. Discusión**

Este estudio describe la diversidad genética y estructura en un conjunto representativo de poblaciones de razas nativas de maíz latinoamericanas, donde parte de este germoplasma ha existido desde tiempos precolombinos. La comparecencia del maíz fuera de su zona de origen y domesticación, sin duda, ha estado relacionada con movimientos migratorios humanos pre y poscolombinos. Las interacciones a lo largo de la historia que el cultivo ha establecido con los diferentes grupos humanos han jugado un papel primordial en el establecimiento de las razas de maíz existente hoy día en Latinoamérica. El fin de este estudio ha sido realizar un análisis exhaustivo de la variabilidad del maíz en Latinoamérica, donde las relaciones genéticas encontradas pueden sustentarse con el contexto histórico que ha envuelto el maíz y que ha sido revelado hasta ahora por la historia, la arqueología y la antropología.

#### **3.4.1 Variabilidad de maíz en México**

Se analizaron las disimilitudes entre los materiales, tomando los valores generados en las matrices de distancias genéticas, y se utilizaron como base en los análisis de agrupamiento, Coordenadas Principales y estructura poblacional. Las razas mexicanas entre las que se encontró la máxima distancia genética fueron Arrocillo Amarillo y Chapapote. Esta distancia indicaría las diferencias genéticas entre estas dos razas como consecuencia de su desarrollo en diferentes localidades y ambientes durante mucho tiempo. Aunque estas razas han sido catalogadas como indígenas antiguas, las cuales se cree que se originaron de un maíz primitivo (Wellhausen et al., 1951), la raza Arrocillo

Amarillo solo ha sido encontrada en los valles altos a más de 2000 metros sobre el nivel del mar, mientras que Chapalote ha sido encontrada en regiones tropicales aproximadas de 100 metros sobre el nivel del mar (Hernández, 1987). Esta información podría sustentar también la existencia de dos sistemas distintos e independientes del cultivo del maíz en épocas prehispánicas (Oliveira-Freitas et al., 2003), entre cuales el flujo génico aparentemente ha sido poco.

Los siguientes valores más altos de distancias genéticas en las razas mexicanas fueron las encontradas entre Palomero Toluqueño, Zapalote Chico y Chapalote. Estas razas son representativas de cada uno de los complejos geográficos mencionados: Valles Centrales, Sur - Suroeste de México, y Sierra Occidental, respectivamente. La raza Zapalote Chico es considerada como raza Mestiza Prehistórica, es decir, una hibridación entre las Razas Exóticas (de Sudamérica) y las Razas Indígenas Antiguas Mexicanas y de ambas con el teocintle (Wellhausen et al., 1951; Hernandez, 1987). Estos valores de distancias genéticas apoyan el origen y parentescos propuestos por Wellhausen et al., (1951) en el sentido que Zapalote Chico no tiene relación con Chapalote ni Palomero Toluqueño. Esta afirmación también se encuentra sustentada por el Análisis de Coordenadas Principales (Figura 3.11a y 3.11b) donde estas razas no se encuentran relacionadas y sugieren adaptaciones a complejos geográficos muy diferentes.

En la exploración de distancias genéticas entre accesiones de una misma raza en las poblaciones mexicanas (Anexo 7.8), se observó un incremento en la distancia, que probablemente está relacionado por las diferentes procedencias de cada accesión y también causado por la deriva genética que causa diferencias subpoblacionales cuando el tamaño de muestra colectado es pequeño. Así por ejemplo, para la raza Cónico las accesiones proceden de tres estados de la República mexicana (Puebla, Tlaxcala y México), mientras que en el caso de la raza Zapalote Chico provienen de Oaxaca y Chiapas. Estos resultados con microsatélites evidencian la gran variabilidad intrapoblacional e interracial a pesar de tener algo de homogeneidad morfológica por algunos caracteres, al estar catalogadas como pertenecientes a una misma raza.

La variabilidad de maíz en México ha sido caracterizada en diversos estudios para describir las razas existentes y las relaciones entre ellas. Comparando con previos

estudios, la presente clasificación destaca algunas diferencias en las clasificaciones realizadas, así como en las estructuras de los grupos. Esto puede ser atribuible a la clase de estudio realizado; si el estudio ha sido basado en datos morfológicos, eco-geográficos, enzimáticos, moleculares (incluyendo el tipo de marcador empleado), o la combinación entre estos datos de distinta índole. Estas discrepancias podría también explicarse por la deriva genética, o una considerable cantidad de flujo génico entre poblaciones como en los valles centrales de Oaxaca en el sur de México (Pressoir y Berthaud, 2004). Estos hechos causarían cambios poblacionales a través del tiempo, si se toma en cuenta el tiempo transcurrido en el que se realizaron los estudios. Sin embargo, los resultados presentados, indican que en base a marcadores moleculares tipo microsatélite y utilizando muestras conjuntas de ADN, las accesiones mexicanas pueden ser agrupadas en distintos complejos raciales, y estos a su vez son consistentes con distribuciones geográficas específicas, que en el pasado, algunas de ellas, fueron ocupadas por culturas precolombinas quienes desarrollaron y cultivaron el maíz. La información obtenida de frecuencias alélicas para esta colección de germoplasma representa un nuevo conjunto de datos de valiosa utilidad, como primer punto, refleja la diversidad de cada accesión representada al menos por 15 individuos, que en estudios anteriores un solo individuo representaba una población. Como segundo punto, la riqueza alélica y la especificidad encontrada en los grupos raciales de los maíces mexicanos podría ser asociada con la distribución geográfica y con algunas características de interés.

### **3.4.2 Variabilidad de maíz en Latinoamérica**

A nivel de Latinoamérica, el valor máximo de distancia genética (0,73) fue encontrado entre las razas Caribeño Precoz y Pisankalla, ambos maíces típicos de tierras bajas, Caribe y tierras bajas de los Andes respectivamente. Esta diferencia podría sugerir migraciones e introducciones independientes de diferentes tipos de maíces de las tierras bajas desde Mesoamérica; uno concentrado a través del Caribe y una ruta continental o por mar vía Pacífico de Sudamérica en épocas tempranas que generarían procesos de adaptación y selección independientes. La distancia genética entre las razas Altiplano de Bolivia, y la raza San Jerónimo-Huancavelicano de Perú (0,18), fue la mínima encontrada en todo el conjunto de germoplasma, y a su vez menor que la mínima distancia

encontrada para las razas mexicanas (0,21), indicando por medio de las distancias genéticas la gran variabilidad presente en los maíces mexicanos del complejo de Sierra Madre Occidental y de otra mano, las estrechas relaciones encontradas dentro de los maíces andinos. Este hecho también sugiere que no hay mucha diversidad que se ha desplazado desde México a Perú y en general a los Andes, sugiriendo un efecto cuello de botella. Estas premisas son también apoyadas por la partición de la variabilidad genética encontrada con los análisis de varianza molecular para los maíces mexicanos y para los grupos definidos por el análisis de estructura de población para todas las accesiones.

El valor de la diversidad genética para el total de germoplasma estudiado tuvo un valor alto (0,62) y solamente menor que en otros estudios cuando se compara con la diversidad para todos los taxones de teocintle (en un rango de 0,65 - 0,89) reportado por Fukunaga et al. (2005), un rango de 0,55 - 0,71 reportado por Warburton et al. (2011). La comparación de los índices de diversidad con los estudios mencionados u otros estudios debe ser cuidadosa e indirecta. En este caso, por la diferencia en taxón, sin embargo, el valor de la diversidad genética en el presente estudio fue consistente con los resultados presentados por Warburton et al. (2008) para razas y poblaciones de maíz (0,61) y con el valor de 0,64 para líneas mejoradas y desarrolladas en CIMMYT (CMLs), que en su base genética podrían tener material exótico de diferentes regiones geográficas no presentes en este estudio.

El total de alelos encontrados para los 28 SSRs y a través de las 194 poblaciones fue 291, pero, si se quiere comparar el número de alelos totales encontrados con otros estudios se debe recordar que el número de alelos presentes es consecuencia de la naturaleza de los marcadores, especialmente si es de tipo dinucleótida, y el número de estos marcadores utilizados probablemente determinarán las diferencias de estos valores entre estudios debido a su alta tasa de mutación (Vigourox et al., 2002). Este hecho es debido a que los SSRs de repetición trinucleótida son más largos, y también porque es el tipo de secuencia repetitiva más común dentro de secuencias de regiones codificadas de un gen activo; los tres pares de bases en la repetición no interrumpen el patrón de tres bases por cada aminoácido que constituye la proteína. Sin embargo, la adición de más o menos aminoácidos pueden truncar la función de la proteína a la cual corresponde el gen, y por

eso, se encuentra menos variación en el número de repetición de SSR dentro de genes que dentro de la cadena de ADN que no codifica genes; sin embargo, los SSR trinucleótidos son más confiables en cuestión de reproducibilidad de datos que los dinucleótidos.

A pesar de lo anteriormente mencionado, el valor promedio de alelos por *locus* encontrado en este estudio (10,39) es mayor que los valores reportados por Warbuton et al. (2008) utilizando 85 SSRs para 261 líneas de maíz y 497 individuos (que representaban 23 razas mexicanas), con valores de 4,9 y 6,3 respectivamente. Esta diferencia proviene del tamaño de muestra en la determinación de los alelos encontrados (15 individuos por población) y la diferencia en la diversidad del material considerado, la cual es más amplia en este estudio con 194 poblaciones. Este trabajo con marcadores moleculares que comprende material tan amplio y diverso es el primero que considera un mayor número de muestras por población pudiéndose identificar alelos raros, que probablemente algunos de ellos pueden tener un gran potencial para mejorar algunos parámetros agronómicos importantes, ya que los alelos que poseen frecuencias alélicas altas tienen más posibilidad de estar ya introgresados en material mejorado.

El valor de diversidad genética evaluado para los grupos definidos por el análisis de estructura en G1 con 46 poblaciones (0,63), confirma que este valor no se ve influenciado por el tamaño de muestra poblacional sino por la naturaleza misma de las poblaciones si se compara este valor con los valores para G2 (0,59) con 58 poblaciones, G3 (0,57) con 65 poblaciones y G4 (0,61) con 25. La diferencia en diversidad genética dentro de cada grupo podría ser el resultado de la deriva genética, sucesivos cuellos de botella y sucesivos eventos de introducción de la agricultura del maíz desde su centro de origen en Mesoamérica hasta alcanzar el Caribe y Sudamérica. Así, consistente con lo sugerido por Rebourg et al., (2003), existiría una tendencia a disminuir los polimorfismos en maíz a medida que se incrementa la distancia geográfica del centro de domesticación, en buena medida porque una parte de la variación original no se ha introducido a otras regiones. Los resultados presentados sugieren que nueva variación alélica resultante de la dinámica de flujo génico en México no ha migrado a nivel macro regional, pero aún más, nueva diversidad genética es generada constantemente en México por el flujo génico con el teocintle (Warburton et al., 2011), eventos que no pueden llevarse a cabo en regiones

fuera de México y de algunos sitios de Guatemala y Nicaragua por la ausencia del teocintle en el Caribe y Sudamérica (Bedoya y Chávez, 2010).

La observación de que la mayor diversidad genética analizada a nivel de grupos y subgrupos (especialmente g1, g2 y g3) se encuentran en G1 es consistente con lo propuesto a cerca de el origen del maíz en la región que hoy es México (Matsuoka et al., 2002b). Entre los subgrupos que conforman cada grupo, el número de alelos únicos confirma la gran diversidad presente en las razas nativas de maíz, por ejemplo, el subgrupo g1 posee 24 alelos únicos en un número total de 192 alelos si se compara con los subgrupos (g2 y g3), que poseen 15 y 11 alelos en un total de 173 y 176 alelos respectivamente, que conforman el conjunto de los maíces mexicanos y mixtos de Sudamérica.

En el análisis de agrupamiento y estructura, algunas accesiones con la misma clasificación racial fueron puestas en más de un subgrupo. Por ejemplo, en el análisis de agrupamiento, una de las accesiones de la raza Amagaceño de Colombia agrupó con el grupo de Territorio Norte y la otra con los maíces Andinos; en el análisis de subestructura una de las accesiones se encuentra en G2 y la otra en G4 (poblaciones no agrupadas). En adición, la raza Perola de Bolivia cuenta con dos accesiones, una agrupó con el G2 y la otra los maíces Andinos (G3). Similarmente, la raza Cateto Sulino de Uruguay se agrupaba en G1 (centro y norte de México- mixtos Sudamérica) y G2. Estas divergencias en algunos casos podrían indicar que la asignación de una población a un subgrupo no refleja los nombres de las razas en que han sido clasificadas, reflejando una limitante en la clasificación basada en parámetros morfológicos. Además, las comparaciones entre regiones están a menudo dificultadas por una falta de homogeneidad en los nombres de las razas (es decir, una misma raza recibe diferentes nombres según la zona de cultivo) y por las inconsistencias en la definición conceptual de raza, especialmente entre los investigadores de diferentes países (Goodman y Bird 1977; Vigourox et al., 2008; Lia et al., 2009).

Lo anterior también es aplicable para los maíces Caribeños, pues también se encontraron estas discrepancias. La raza Cristalino Cubano de las Islas Vírgenes, Venezuela; dos accesiones de la misma raza de Cuba pertenecen al mismo subgrupo, mientras que una

tercera proveniente también de Cuba se encuentra no agrupada. La raza Cristalino Costeño Tropical de Islas Vírgenes, Antigua y Guadalupe se encuentran en un mismo subgrupo (g7) y las accesiones para esta raza de Barbados y San Vicente no están agrupadas. De otro lado, la raza Chandelle de Cuba se encuentra en g4 y las provenientes de República Dominicana y Guadalupe en g7, pero en este caso se identifican las dos formas de Chandelle (cristalino y dentado) reportadas previamente (Goodman y Brown. 1988). Aunque, de otro lado, la clasificación de los maíces caribeños en grupos claramente definidos no es fácil debido a las pocas barreras geográficas y la cercanía entre las Antillas Mayores (Cuba y República Dominicana, entre otras) y las Antillas Menores (Antigua, Barbados, Trinidad y Tobago, entre otras) que han facilitado la libre migración tanto humana como por supuesto de variedades de maíz.

### **3.4.3 Maíz: historia, arqueología, antropología y asociaciones genéticas en México**

En el periodo definido como Preclásico Temprano, 2.500-1.200 antes de Cristo (a.C.), surgen en Mesoamérica las primeras aldeas agrícolas y se comienza a usar la cerámica que, entre otras cosas, sirve para preparar alimentos que necesitan cocción. Sin embargo se tienen dataciones de restos de mazorcas que aún no presentan las características propias de una especie completamente domesticada de mil años atrás (Vela, 2011). El florecimiento de las grandes civilizaciones e imperios Americanos, como por ejemplo, las civilizaciones Maya, Mexica e Inca se sitúa aproximadamente entre el 1.500 a.C. y 1.500 después de Cristo (d.C.), mientras que, y aproximadamente entre el 2.000 a.C., fecha anterior al tiempo de establecimiento de estas culturas ya se encontraba maíz domesticado según las dataciones de las cuevas de Romero y Valenzuela en el estado de Tamaulipas, México (Blake, 2006). Lo anterior indica que estas civilizaciones recibieron el legado del cultivo del maíz de sus culturas progenitoras y/o antecesoras. En este sentido, vale la pena preguntarse en el periodo anterior al Preclásico Temprano en Mesoamérica que culturas coexistían y de igual manera que culturas coexistían en Sudamérica en un periodo anterior al Imperio Inca vinculadas al cultivo del maíz.

Si se toma en cuenta, en primera instancia: i) que estudios de restos carbonizados de maíz encontrados en contexto arqueológico en Teotihuacán-México, y fechados alrededor de 100 d.C, sugieren que, en algunos casos las razas de maíz existentes en el primer siglo de

nuestra era son similares a razas actuales encontradas en el Centro de México (Benz, 1997), y ii) que entre las razas más primitivas de México basados en estudios genéticos, arqueológicos y morfológicos se han citado, Arrocillo Amarillo, Tabloncillo, Pepitilla (no presente en este estudio), Palomero Toluqueño, Nal-tel y Chapalote (Benz, 1997; Wellhausen et al., 1951); así pues, y con las debidas precauciones se podría extender la relación entre maíces considerandos como primitivos, su zona geográfica de distribución y culturas precolombinas establecidas en estas regiones con la clasificación en complejos raciales de maíces mexicanos descrita en este estudio.

Con respecto a lo planteado, culturas como la Tolteca, Teotihuacana y Mexica establecidas en los Valles centrales de México estarían relacionados con Pepitilla (Muñoz, 2003), Palomero Toluqueño, Cónico y Chalqueño; estas dos últimas razas citadas como mestizos prehistóricos por Wellhausen et al. (1951). La zona occidental de México ha sido la menos explorada, a pesar de esto, la cultura Tarasca es representativa de la región y estaría asociada con la raza Chapalote. Al sur, y Golfo de México, se establecieron en el actual estado de Oaxaca, culturas como la Zapoteca y Mixteca, y podrían vincularse con el Complejo Zapalote y la raza Bolita, razas típicas de la región, y que además se han encontrado en contexto arqueológico (Reyna y Alvarez, 2009). Arrocillo Amarillo y Tuxpeño han sido asociados con la cultura Olmeca y Totonaca que se desarrollaron el Golfo (Muñoz, 2003), Finalmente la cultura Maya estaría asociada con las razas Nal-tel y Olotón en sur de México.

El flujo génico también se ha puesto en evidencia al sur de México en contexto arqueológico (Reyna y Alvarez, 2009), en donde se han encontrado, razas típicas de la región y otras razas foráneas como Cacahuacintle, Chalqueño de los Valles Centrales; Elotes Occidentales de la Sierra Madre Occidental y Olotón de una zona más al sur. Estos hallazgos podrían sugerir un fuerte intercambio de semilla con grupos culturales de las diferentes regiones y que alternativamente podrían estar experimentando con la adaptación de razas atípicas de esta zona geográfica.

#### **3.4.4 Dispersión del maíz por el litoral, tierra y mar**



A partir de su centro de origen en México y mediante un modelo no esclarecido de difusión, algunas poblaciones de maíz se difundieron hacia otras regiones de Latinoamérica donde se han adaptado a través de un proceso milenario de domesticación a las diferentes condiciones ambientales y estreses locales propios de la geografía latinoamericana. Sin embargo, estudios recientes de estructura de población y diversidad genética apoyan la difusión del maíz hacia el norte a partir de su centro de origen en la Cuenca del Río Balsas, llegando al suroeste mexicano y posteriormente al norte de Estados Unidos, y finalmente a Canadá. También los resultados de filogenia mostraron claramente que las razas del norte estadounidenses son derivadas de aquellas del suroeste de Estados Unidos, indicando que las razas de maíz del suroeste son intermedias entre las del norte estadounidense y las razas provenientes de las tierras altas mexicanas, sugiriendo una relación directa entre las dos regiones, proponiendo una dispersión del maíz por tierra hacia los Estados Unidos y no a través del Golfo de México o de las Antillas Mayores (Matsuoka et al., 2002a; Vigouroux et al., 2008).

En Mesoamérica, los investigadores en arqueología y antropología coinciden que después del 2.500 a.C. y antes del 150 d.C. la agricultura se convierte en la forma dominante de adquisición de alimentos y en el modo de subsistencia (González, 2009); sin embargo, en la zona intermedia de América, culturas como Valdivia y Chorrera, se desarrollaron en una de las zonas más áridas al sur de la costa ecuatoriana con evidencia temprana del cultivo del maíz (Museo Chileno de Arte Precolombino, 2011). La información de esta dispersión temprana y las culturas ligadas al cultivo del maíz puede haberse perdido en el tiempo y su esclarecimiento estaría sujeto a nuevos descubrimientos arqueológicos. De hecho, la cultura Chorrera ha sido relacionada con la cultura temprana Ocos en Chiapas-México por su similitud en la cerámica (Pye y Demarest, 1993); estas culturas podrían también haber compartido conocimientos relacionados con el cultivo del maíz. Esto no sería un hecho aislado, de la misma manera se han encontrado influencias culturales de Mesoamérica en Sudamérica y viceversa, en este sentido, el cultivo del maíz no podría estar ausente de una red de intercambio en la América precolombina.

De acuerdo con los patrones de diversidad genética de las poblaciones humanas en América se sugiere que las rutas de migración de los nativos americanos siguiendo el

litoral fueron más fáciles para la migración en la travesía de norte a sur que a través del interior (Wang et al., 2007). Esta premisa podría ser aplicada a los cultivadores tempranos del maíz que en su andar migratorio seguirían rutas tempranas por el Pacífico y Caribe. Basados en la teoría de dos sistemas de agricultura concentrados en las tierras bajas y altas (Oliveira-Freitas et al. 2003) y junto con los resultados presentados aquí, por las similitudes encontradas por McClintock et al., (1981) para el “Territorio Norte” se apoyarían la hipótesis de rutas tempranas a través de la costa caribeña que se extiende a lo largo de la costa del noreste de Sudamérica, y que penetraría en el continente por los sistemas fluviales que desembocan en el Océano Atlántico (Figura 3.26).

Trazar rutas por las tierras altas de Centroamérica, en la práctica, se torna algo complicado por las barreras geográficas extremas, montañas escarpadas especialmente al sur de Costa Rica y Panamá, selvas y montañas en el Darién y Choco en Colombia, y nuestros resultados no proporcionan evidencia directa para trazar estas rutas. Sin embargo, esta posible vía se ve sustentada por los estudios llevados a cabo por Oliveira-Freitas y colaboradores (2003), que muestra como el sistema agrícola de tierras altas estaría centrado en las cadenas montañosas que se extienden desde Centroamérica hasta las tierras altas de Panamá y continuaría en las regiones Andinas sobre el lado occidental de Sudamérica.

Los antiguos indígenas americanos asociaban el maíz con el mar y los ríos debido a que este complejo simbolizaba las cualidades o funciones del agua y la fertilidad. En Sudamérica, la chicha (bebida hecha de maíz fermentado) era ofrecida a la tierra o *Pachamama* (Melgar, 2003). En general, el escenario de las culturas pre-colombinas fue diverso y dinámico. En esta tónica, la navegación prehispánica estuvo lo suficientemente desarrollada para llevar a cabo viajes a través de las costas. La evidencia sugiere intercambio comercial y cultural entre los Andes y Mesoamérica, conectando El Soconusco y Tehuantepec en Mesoamérica con Huanchaco y Chíncha en Sudamérica, además, las tecnologías marítimas que fueron desarrolladas por culturas como la Maya, Chíncha y Mochica tuvieron la viabilidad de realizar viajes entre estas dos áreas (Melgar, 2003). Una vez el maíz establecido en Sudamérica, con migraciones desde la

costa hacia el interior y del maíz de altura hacia las tierras bajas tendría lugar las introgresiones de las variedades dando lugar a nuevos fenotipos.

**Figura 3.26** Rutas sugeridas de la migración del maíz desde su centro de origen en Mesoamérica hacia el Caribe y Sudamérica basada en la evidencia arqueológica presentada, estudios antropológicos e históricos y las relaciones genéticas encontradas en este estudio.



### 3.4.5 Relaciones genéticas y contexto histórico del maíz en Latinoamérica

Los análisis llevados a cabo para Latinoamérica evidencian que las poblaciones centroamericanas se agrupan entremezcladas con el resto de grupos. Esto probablemente refleja patrones de migración a través de América Central en diferentes tiempos, desde México hacia América Central, y de Guatemala hacia el Caribe y Sudamérica atravesando los actuales países Centroamericanos de Costa Rica y Panamá (de las que se encuentran algunas poblaciones en este estudio). Aunque en Colombia ciertas razas de maíz en muchos aspectos son parecidas a razas mexicanas que han sido consideradas

como introducciones precolombinas (Wellhausen et al., 1951) en este análisis no se encuentran relacionadas directamente con México ni Centroamérica. Esto indica que probablemente los maíces primitivos de las dos regiones han estado separados por mucho tiempo. Un resultado similar fue el reportado por Vigouroux et al. (2008) cuando estudiando la raza Montaña no agrupó con su contraparte centroamericana Olotón, no encontrando de esta manera evidencia para sugerir que el maíz de los Andes fue descendiente del maíz de las tierras altas de Guatemala. Sin embargo, para llegar a esta conclusión tan taxativa se deberían estudiar más razas de las tierras altas de Centroamérica y Colombia, sobre todo teniendo en cuenta que existen más de 300 razas que descienden de los maíces que originalmente cultivaron los indígenas latinoamericanos (Bracco et al., 2009).

El análisis de Coordenadas Principales para Latinoamérica reflejó la unicidad de los maíces de las tierras altas de los Andes. Estos maíces se separan muy bien a media que se alejan de las poblaciones del oriente de Sudamérica y de las poblaciones de maíces de tropicales y del subtrópico del norte de Sudamérica de acuerdo con otros estudios (Matsuoka et al., 2002b, Vigourox et al., 2008). Esto podría sugerir introducciones muy antiguas de las tierras altas andinas que se han conservado, también indica que el maíz de los Andes no ha migrado a nivel macro regional. Aunque de otro lado, este mismo análisis sugiere que parte del germoplasma Andino también podría derivarse de la posible mezcla del germoplasma proveniente de las laderas Occidentales de los Andes con el maíz proveniente de la costa Oriental Sudamericana (Territorio Norte).

En el caso de Chile algunas de sus poblaciones se encuentran relacionadas con el grupo de noreste de Sudamérica, y con poblaciones del Complejo Andino que tal vez podría apoyar la teoría de que el germoplasma chileno se deriva del maíz de los Andes, aunque como lo sugiere Vigourox et al. (2008) posiblemente de dos eventos diferentes.

Referente al origen del maíz del Caribe, los análisis aquí presentados sugieren que los maíces caribeños pudieron ser introducidos desde la zona continental latinoamericana a partir de diferentes fuentes de germoplasma, apoyando así la teoría propuesta por McClintock et al. (1981), pero refutando la idea de una sola fuente de germoplasma que

sería introducida desde la zona continental de Sudamérica vía Trinidad y Tobago como se sugirió en el estudio llevado a cabo por Vigourox et al. (2008). Estas diferencias en la deducción de la introducción del maíz en el Caribe podrían ser consecuencia de la disparidad en el número de muestras representativas empleadas en cada estudio, en este aspecto, el maíz del Caribe en el estudio de Vigoroux y colaboradores. (2008) no estaría suficientemente representado.

La introducción del maíz al Caribe, desde dos fuentes diferentes sugerida en el presente estudio está indicada específicamente, por las relaciones encontradas entre las razas Sabanero y Tusón de Venezuela, con las razas Tusón y Cristalino Costeño Tropical de Trinidad y Tobago, y San Vicente respectivamente, los resultados presentados apuntan a que islas al sur y cerca de la zona continental de Sudamérica experimentarían introducciones desde el territorio que hoy es Venezuela. Las islas más al norte recibirían introducciones de germoplasma centroamericano desde Panamá, Costa Rica y sur de México. Estas introducciones de maíces desde diferentes puntos geográficos hacia las islas del Caribe, podrían estar sustentadas por dos migraciones humanas desde tierra firme a las islas. Una migración temprana, aproximadamente de 5.500 a.C., de nativos sudamericanos quienes probablemente poblaron el Caribe vía Trinidad y Tobago; la segunda alrededor de 3.000 – 4.000 a.C. de nativos centroamericanos específicamente de Yucatán en México (Newson, 2006).

La complejidad de la clasificación de los maíces caribeños encontrada aquí podría explicarse mediante componentes geográficos e históricos. Las pocas barreras geográficas y la proximidad de las Antillas Mayores y las Antillas Menores podrían facilitar la libre migración de pueblos entre ellas. A su vez, estas migraciones pudieron haber estado acompañadas por el transporte de variedades maíz, que de alguna manera enmascararon las diferencias raciales entre ellas (Goodman y Brown, 1988). El componente histórico no es menos complejo. Especialmente, la cultura Taino parece estar relacionada con el cultivo del maíz y con la llegada de Cristóbal Colón a América. En 1492, fueron los primeros nativos en tener contacto con los españoles, liderados por Cristóbal Colón (Torres, 1988). Después de la llegada de Colón, fueron iniciados otras clases de movimientos migratorios, no solamente de personas desde la Península Ibérica,

movimientos también asociados a colonias inglesas, orientales y el tráfico de esclavos desde África, quienes indudablemente tuvieron una influencia en la agricultura, fomentado el intercambio entre las islas y tierra firme en el Nuevo Mundo; así como, la introducción y rápida dispersión del maíz a Europa (Angehiera, 1907, 1st complete ed. 1530; Dubriel et al., 2006).

Las relaciones encontradas entre los maíces caribeños y la zona continental de Sudamérica se pueden sustentar históricamente por la migración de los pueblos del norte, especialmente del área del Caribe, a través de las tierras bajas de lo que hoy es Perú, Brasil y Bolivia; estos pueblos, como los Arawak y Guaraníes, se movieron a lo largo de los ríos hasta alcanzar las estribaciones de los Andes (Ramirez et al., 1961) y contribuiría al grupo de maíces denominados aquí como Territorio Norte.

Si se tiene en cuenta que el sistema montañoso de los Andes bordea la costa Pacífica por más de 7.500 Km, existe una gran posibilidad de múltiples contactos entre las culturas precolombinas. La composición del grupo formado por los maíces andinos, en los análisis de estructura, algunas veces están superpuestos geográficamente, sugiriendo eventos tales como introducciones tempranas, flujo génico intenso y que probablemente en la actualidad empleando únicamente el germoplasma de maíz existente es difícil seguir con absoluta certeza las huellas de los orígenes de estos maíces y de sus relaciones ancestrales con las poblaciones tempranas Mesoamericanas.

Especialmente, los Andes Centrales representaron una zona de contacto, intercambio comercial y cultural que podría haberse dado de diferentes clases y en diferentes direcciones, migraciones en dirección sur-norte y viceversa, incluyendo también movimientos desde las laderas este y oeste de toda la cadena Andina. Este flujo e intercambio de germoplasma no sería sorprendente porque la evidencia arqueológica también ha sugerido que las culturas precolombinas de las tierras altas de los Andes han sido conquistadas o han comerciado con las culturas de las tierras bajas (Sanchez et al., 2006).

Los avanzados sistemas de irrigación diseñados por algunas culturas precolombinas, como los Aymara y Quechua (Ramirez et al., 1961) favorecerían la adaptación de razas

provenientes de las tierras altas y que podrían haberse mezclado con razas locales de la costa. Este intercambio también podría verse apoyado por algunas condiciones climatológicas especiales, como, en la costa de Perú y la costa sur de Ecuador, donde el clima a lo largo de la costa es subtropical árido debido a la corriente de Humboldt, y que equivaldría por temperatura a 2.000 metros de elevación sobre el nivel del mar. Estas condiciones climáticas particulares incrementarían la adaptabilidad de razas provenientes de tierras altas en la costa (Sánchez et al., 2006).

El maíz se convirtió en un cultivo importante alrededor del año 1.200 a.C. A lo largo de la costa del Pacífico que bordea los Andes, y durante el periodo 600-1.000 d.C. (Horizonte Medio), la evidencia arqueológica de maíz disponible refleja una producción y distribución de maíz muy dinámica; así como, la existencia de diferentes variedades de maíz, morfológicamente distintos y correlacionadas con su distribución geográfica en el centro-sur de los Andes (Hastorf et al., 2006). En la sierra, el cultivo del maíz tuvo una importancia religiosa, social y económica; su cultivo se distribuyó ampliamente a lo largo de los Andes por la influencia de la cultura Inca (Hastorf et al., 2006). Durante el período del Imperio Incaico, que se estableció como el estado más grande de América precolombina, cubriendo territorio desde el sur de Colombia a través de Ecuador, Perú, Bolivia, Chile y el noroeste de Argentina, los incas llegaron a un alto grado de conocimiento de maíz, al extremo de distinguir diferentes tipos de maíz por su calidad nutricional y su distribución geográfica se realizaba de acuerdo a las condiciones agroecológicas del cultivo (Manrique, 1997). La primera tarea de los Incas al conquistar nuevos territorios fue siempre la de hacer trabajos de irrigación con el propósito de favorecer el cultivo del maíz. En los valles más calientes, se hicieron terraplenes y canales de riego, terrazas intensamente regadas y fertilizadas; en las tierras altas hicieron posible el cultivo de maíces de alto rendimiento, aun en condiciones climáticas que imponían un largo periodo vegetativo y corta época de lluvia (Ramírez et al., 1961).

En el análisis de Coordenadas Principales para todo el germoplasma, los maíces andinos se encuentran separados del resto del germoplasma por una “pobre” zona intermedia, ubicada entre lo que se denominó México-Centroamérica, México-Centroamérica-Caribe y Territorio Norte. Podría ser entonces, que la razón por la que no se encuentran

asociaciones clave para esclarecer el origen del maíz de los Andes es la posible desaparición de algunas razas de maíz, e históricamente este razonamiento nos dirige a la época del descubrimiento de América. Aunque, otra posibilidad es que estas razas todavía están allí, existen, y son sembradas en zonas indígenas de difícil acceso y no han sido colectadas.

Con la llegada de los españoles se añaden nuevos ingredientes a la historia del maíz, no solamente en el Caribe, la primera zona de contacto entre las culturas de los dos continentes, sino también en tierra firme. La colonización, no solo interfiere en los sistemas productivos existentes en la época prehispánica, también, la introducción de cultivos de alto valor comercial como la caña de azúcar, el uso de tierras para el cultivo de trigo y la disminución de los niveles poblacionales en el Nuevo Mundo. La disminución drástica de la población nativa, aunque no se tiene un dato preciso al respecto (Cavalli-Sforza et al., 1994), conduce a pensar que afectaría negativamente el mantenimiento de las razas de maíz que para esta época se habían convertido en alimento básico para las culturas precolombinas.

Además, la evidencia arqueológica y estudios de linajes basados en la morfología, sugieren que algunas razas de maíz que existieron en épocas tempranas ya no existen en la actualidad (Benz, 1997). Este hecho podría esconder algunas relaciones importantes para la clarificación y postulación de progenies o linajes primitivos de maíz. Por otra parte, si la introducción del maíz a Sudamérica desde su lugar de origen, ocurrió hace bastante tiempo, las relaciones entre las poblaciones ancestrales de maíz y sus derivados podrían ser irreconocibles después de procesos independientes de selección y diversificación.

Muchas preguntas quedan abiertas y algunas están en la discusión y polémica. Similarmente, hubo preguntas polémicas sobre el origen unicéntrico o multicéntrico del maíz a partir del teocintle, y aunque ya está entendido, preguntas sobre la importancia que históricamente ha representado el flujo génico entre algunas subespecies de teocintle y maíz actualmente no han quedado resueltos. Las dataciones micro y macro botánicas no logran identificar un punto en común del origen geográfico del maíz primitivo. Los datos genéticos pueden apoyar la búsqueda de respuestas a las preguntas que aún quedan, pero



es necesario encontrar la mejor manera para integrar esta clase de datos con los datos botánicos, arqueológicos y registros históricos. Con todo lo anterior es innegable que el maíz ha sido un eje fundamental en las culturas del pasado y en las actuales, que ha conducido al desarrollo y mejoramiento continuo de innumerables técnicas para cultivarlo, y no hay que perder de vista el valor intrínseco que han representado de las razas de maíz a través de todo el continente Americano.

### **3.5 Conclusiones**

Dado que la caracterización tradicional basada en el fenotipo puede presentar problemas debido a la influencia de factores ambientales, la posibilidad de estudiar la diversidad genética entre poblaciones heterogéneas de maíz mediante el uso de marcadores moleculares (SSR) y el uso de muestras conjuntas de ADN se confirmó y en definitiva contribuye a conocer mejor la diversidad genética existente en las poblaciones nativas de maíz.

Se avaló el uso del conjunto de SSRs en el estudio de la diversidad genética a nivel Latinoamérica con la validación realizada en las poblaciones de maíces mexicanos mediante los análisis de agrupamiento, coordenadas principales y estructura poblacional, apoyando de esta manera caracterizaciones previamente realizadas y en algunos casos respaldando orígenes y parentescos.

A partir de las relaciones encontradas entre los maíces mexicanos, de Centroamérica, Caribe y Sudamérica se hace evidente la introducción del maíz desde Mesoamérica hacia el Caribe y Sudamérica mediante una dispersión progresiva, y de acuerdo también con la evidencia arqueológica, el maíz siguió diferentes rutas de difusión, atravesando tierra firme en una geografía diversa, siguiendo rutas marítimas y/o costeras vía Pacífico y Caribe, y alcanzando la cadena montañosa andina.

Los movimientos precolombinos y poscolombinos de las razas de maíz indudablemente han estado ligados a las migraciones humanas, destacando que en un estudio de la diversidad y dispersión del maíz se deben de considerar una gama de procesos sociales y económicos de gran complejidad de las culturas precolombinas y actuales, así como

eventos históricos que han sido determinantes para el establecimiento de la diversidad presente en las razas de maíz en Latinoamérica.

Una continuación lógica del presente estudio sería profundizar en la caracterización molecular de restos arqueológicos de maíz y su comparación con las poblaciones actuales podrían aportar luz sobre las cronologías de las introducciones e introgresiones de los diferentes tipos de maíz en épocas antiguas.

## **4. CARACTERIZACIÓN FENOTÍPICA Y GENÉTICA DE UN PANEL DE MAÍCES TROPICALES PRECOCES DE INTERÉS PARA ESTUDIOS DE ASOCIACIÓN DE TOLERANCIA A LA SEQUÍA.**

### **4.1 Introducción**

El aumento de la productividad del maíz, en los trópicos no son una elección sino una necesidad. Muchos países en los trópicos están sometidos a una fuerte presión demográfica y encaran una alta tasa de incremento de la población. Estos miles de millones de personas no solo deben ser alimentados a niveles de sustentación sino que deben urgentemente mejorar su nivel nutricional, especialmente los sectores de menos recursos donde el maíz es recurso básico. El potencial para incrementar la producción de maíz que satisfaga esas necesidades alimenticias es sin duda superior en el caso del maíz que en el de otros cereales. En los trópicos bajos de Latinoamérica el maíz aporta hasta un 40% de las proteínas y el 50% de las calorías en las dietas de algunos países centroamericanos, donde la desnutrición afecta severamente a casi la mitad de su población infantil menor de cinco años (FONTAGRO, 2008).

El maíz se siembra en más de 4 millones de hectáreas en Centroamérica, Colombia, Venezuela, Perú, Ecuador y Bolivia, con un rendimiento promedio de aproximadamente  $2,5 \text{ t ha}^{-1}$  (FONTAGRO, 2008). En la actualidad, la productividad de maíz podría estar negativamente afectada por los cambios ambientales asociados al cambio climático siendo el más evidente el aumento en las temperaturas a la que los cultivos de maíz serán expuestos (World Bank, 2007). La sensibilidad del rendimiento en el cultivo a los estreses abióticos como elevadas temperaturas (solas o asociadas al estrés hídrico o de nutrientes) afectarán de manera negativa los componentes agronómicos de este carácter.

Para el 2025 se estima que alrededor de 1,8 millones de personas sufrirán por escasez de agua y dos tercios de la humanidad se verán afectados por lo que la FAO ha definido como estrés hídrico (Ribaut et al., 2008). En agricultura, el término sequía, refiere a la condición en la cual la cantidad de agua disponible a través de la precipitación o irrigación es insuficiente para mantener las necesidades de transpiración del cultivo

(Tuberosa, 2011). Entre las diferentes definiciones de la resistencia a la sequía que se han formulado durante las pasadas décadas, cabe resaltar, la definición original propuesta por Levitt (1972) donde, se clasifican las estrategias para mitigar los efectos negativos del déficit de agua. La primera sería la “evasión o escape a la deshidratación”, definida como la capacidad de mantener la hidratación de la planta o celular bajo estrés, y la tolerancia a la deshidratación, esta última expresada como la capacidad de la planta de mantener su funcionalidad en un estado de deshidratación.

Características morfo-fisiológicas como, por ejemplo, la profundidad de las raíces, la floración temprana, ajuste osmótico, entre otros, estarían relacionadas con evitar la deshidratación y características como la re-movilización de los carbohidratos solubles del tallo y la acumulación de protectores moleculares que permitirían mantener al menos parcialmente la funcionalidad de la planta en estados severos de deshidratación se relacionarían con la tolerancia a la deshidratación (Tuberosa, 2011). La tolerancia a sequia también puede definirse como la capacidad de ser más productivo bajo condiciones de estrés, donde la expresión de la tolerancia depende en gran medida del estado de desarrollo de la planta en el cual el estrés es aplicado y de la intensidad y duración de éste (Ribaut et al., 2008).

Se ha demostrado que la sequía tiene un impacto devastador sobre la productividad de muchas especies de cultivo (Sawkins et al., 2006; Heisey y Morris, 2006). La mayoría del maíz tropical es sembrado bajo condiciones de secano, y el uso de riego para mitigar la escasez de agua es raramente una opción. La combinación de una creciente competencia por el agua, unida a la tendencia de mayores temperaturas globales, y probablemente más cambios extremos climáticos regionales y estacionales en algunas zonas, conllevan a que el desarrollo de variedades de maíz con una mayor resistencia a estrés hídrico y mayor eficiencia en el uso del agua se convierta en un objetivo de alta prioridad para los principales programas de mejoramiento de maíz y otras plantas, tanto en los sectores público y privado (Collins et al., 2008).

En un programa de mejoramiento, la evaluación de la resistencia a la sequía es uno de los problemas críticos que un mejorador debe enfrentar. La cuidadosa selección de los sitios representativos de las condiciones prevalentes en el área objetivo de evaluación, porque

la presencia e intensidad de la sequía en ambientes de secano es impredecible siendo la selección para tolerancia a sequía un experimento único e irreplicable (Ribaut et al., 2008). Sin embargo, en este sentido, la decisión de la fecha de siembra, el uso de una irrigación controlada, es decir, la regulación de la severidad del déficit de agua en estos ambientes proporcionan una alternativa efectiva para la evaluación de la tolerancia a sequía (Campos et al., 2004). De otro lado, la selección para rendimiento en ambientes bajo óptimas condiciones que permite seleccionar plantas con alto rendimiento que van a tener un mejor rendimiento bajo condiciones de sequía, generalmente es menos efectiva que la selección directa de individuos con mejor desempeño en ambientes bajo condiciones de estrés (Ribaut et al., 2008).

Una buena caracterización fenotípica es un punto crítico para cualquier clase de actividad experimental. Una colección de datos fenotípicos precisa se verá reflejada en una mayor heredabilidad del carácter evaluado (Tuberosa, 2011). La heredabilidad se refiere a la porción de la variabilidad fenotípica explicada por los efectos genéticos que pueden ser heredados a través de las generaciones propagadas sexualmente (Falconer, 1981); en otras palabras, mide la importancia relativa de la varianza genética como determinante de la varianza fenotípica. La heredabilidad de los caracteres varía de acuerdo con la composición genética de los materiales bajo evaluación, las condiciones en las cuales los materiales están siendo evaluados y la precisión y exactitud de los datos fenotípicos (Tuberosa, 2011). Un valor de heredabilidad bajo puede indicar varias cosas: que un gran número de genes regulan el carácter de interés, que la caracterización fenotípica no es precisa atribuible al error experimental, y/o a que una proporción significativa de la variación del carácter se debe al medio ambiente (Interacción Genotipo – Ambiente,  $G \times E$ ), (Flint-Garcia et al., 2005). La heredabilidad para rendimiento bajo condiciones de estrés, en gran medida depende de dos factores: (i) la existencia de genes para resistencia a sequía en la población que son efectivos en los ambientes de estrés en los que la selección se lleva a cabo (Blum et al., 2001) y (ii) el grado de control sobre la homogeneidad de las condiciones generales de estrés y por lo tanto la calidad de los datos fenotípicos (Blum, 2005). En condiciones de estrés severo, la heredabilidad de los caracteres que determinan el rendimiento de los cultivos bajo condiciones de sequía usualmente tienen valores bajos

entre 0,30–0,40, o un valor intermedio entre 0,40-0,70 si la  $G \times E$  no es muy grande o significativa (Tuberosa, 2011). Sin embargo, con una adecuada selección de material genético y la minimización de la varianza del error por buen manejo del experimento, la heredabilidad para rendimiento bajo condiciones de estrés puede ser alta y la selección eficaz (Blum, 2005). La determinación de la heredabilidad es importante porque es un buen indicador del potencial de la selección fenotípica y la capacidad de analizar apropiadamente las bases genéticas de los caracteres (Falconer, 1981; Melchinger et al., 2004).

#### **4.1.1 Mejora analítica y rendimiento de grano**

El rendimiento de grano puede ser modelado como el producto del agua transpirada, la eficiencia en el uso del agua y el índice de cosecha (Harvest Index, HI, por sus siglas en inglés), de esta manera, caracteres secundarios pueden ser clasificados en base a su contribución a uno o más de estos factores (Ribaut et al., 2008). El rendimiento de grano y su respuesta al estrés son rasgos complejos que implican una interacción genética a largo plazo, es decir, durante el ciclo de cultivo y obviamente en particular durante el periodo de estrés. La mejora convencional basada en la selección fenotípica por el rendimiento no es lo más apropiado para condiciones de sequía debido a las interacciones del genotipo por estación y genotipo por localidad, y la dificultad de obtener un estrés hídrico homogéneo en el campo incrementando la  $G \times E$  que trae como consecuencia una baja heredabilidad del rendimiento (Araus et al., 2002; Monneveux y Ribaut, 2006). Como alternativa y para mejorar la eficiencia de la selección, surge la denominada mejora analítica, que consiste en incorporar por encima de la selección para rendimiento el valor de caracteres secundarios en la selección (Latiffé et al., 2003). Estos caracteres secundarios deben cumplir algunos requisitos, como por ejemplo, estar correlacionados con el rendimiento de grano de manera significativa en condiciones de estrés, el carácter debería estar menos afectado por el ambiente que el rendimiento de grano (heredabilidad alta), debe ser posible medir el carácter de manera rápida, fiable y económica, debe ser posible también, evaluarlo en plantas individuales o en parcelas y debe existir suficiente variabilidad genética para el carácter a seleccionar y finalmente no

tener ningún impacto negativo sobre selección en condiciones normales (Bolaños et al., 1993; Edmeades et al., 1997; Araus et al., 2008).

De manera adicional, el desarrollo de técnicas moleculares que permiten la identificación de regiones genómicas asociadas con la expresión de caracteres cualitativos o cuantitativos asociados con rasgos fenotípicos, así como, los avances tecnológicos en las mediciones fisiológicas han abierto la posibilidad de nuevos enfoques en el uso de caracteres secundarios en programas de mejoramiento (Monneveux y Ribaut, 2006). En esta área, caracteres morfológicos (componentes del rendimiento, parámetros de floración, estructura del sistema radicular) y parámetros fisiológicos (como el papel del Acido Abscísico, ABA), ciclo celular, mantenimiento de los foto asimilados) asociados con tolerancia a sequía en maíz, así como su heredabilidad, han sido bien documentados (Sawkins et al., 2006). En esta especie de caracteres secundarios se incluyen fechas de floración masculina, femenina y la diferencia entre estas dos floraciones, el ASI (Anthesis Silking Interval, por sus siglas en inglés) que son caracteres fenológicos que afectan el índice de cosecha. Teniendo en cuenta lo anteriormente mencionado, la diferencia entre la floración femenina y masculina (ASI), es un excelente carácter secundario debido a que suele estar altamente y negativamente correlacionado con el rendimiento de grano y tiene una heredabilidad alta (Bolaños y Edmeades, 1996), su variabilidad genética es alta bajo sequía (Ribaut et al., 1996) y por lo tanto es ampliamente utilizado para la selección de maíz en sequía (Araus et al., 2010). Caracteres como senescencia, contenido de clorofila y enrollamiento de la hoja pueden también ser útiles como caracteres secundarios y son particularmente eficientes en una preliminar selección entre un gran muestra de germoplasma susceptible (Ribaut et al., 2008).

En el enfoque de mejora analítica, los caracteres valiosos o de interés específico pueden ser combinados en un Índice de Selección (IS). El IS es una traducción cuantitativa de un concepto de ideo tipo; para su definición son escogidos pesos para los diferentes caracteres basados en la varianza, heredabilidad, y la correlación genética con el rendimiento (Araus et al., 2011). El cálculo del IS para sequía se basa en los valores fenotípicos (P) de diferentes caracteres que deben ser estandarizados antes de calcular la

combinación lineal puesto que sus escalas de medición son diferentes, la estandarización usual es:

$$P_i = (x_{ij} - m_i) / s_i$$

Donde  $m_i$  y  $s_i$  son la media y la desviación estándar del carácter  $i$  en la población,  $x_{ij}$  es el valor del carácter  $i$  medido en el genotipo  $j$ .

Un Índice de Selección en su forma simple puede ser escrito como:

$$IS = b_1P_1 + b_2P_2 + \dots b_nP_n$$

Donde  $P_i$  es el valor fenotípico estandarizado observado del carácter  $i$  y  $b_i$  es el peso dado al carácter en el Índice de Selección (Bänziger et al., 2000).

Se han establecido valores típicos de pesos para algunos caracteres secundarios en maíz, +3, -2, -2, -2 y -1, para número de mazorcas por planta, ASI, senescencia, tamaño de la inflorescencia masculina y enrollamiento de hoja, respectivamente (Bänziger et al., 2000). Los Índices de Selección son continuamente redefinidos, no solo teniendo en cuenta el ambiente objetivo de selección sino también la incorporación de nuevos caracteres secundarios y las innovaciones tecnológicas para medirlos (Bänziger et al., 2000; Cerón-Rojas y Sahagún- Castellanos 2005; Araus et al., 2011).

#### **4.1.2 Tolerancia a la sequía y análisis de mapeo por asociación**

La tolerancia a sequía es uno de muchos caracteres complejos de importancia agronómica y evolucionaria en particular en países en desarrollo donde la disponibilidad de agua y/o a sistemas de riego es una limitante mayor. Como todos los caracteres cuantitativos su variación fenotípica es regulada por: i) QTLs (*quantitative trait loci*, por sus siglas en inglés), regiones sobre cromosomas cuya variación alélica está asociada con parte de la variación de un carácter poli génico; ii) interacción entre genes o QTLs (por ejemplo, epistasias); iii) el ambiente y iv) la interacción entre QTLs y el ambiente (Zhu et al., 2008). Entender las bases genéticas de la variación fenotípica tiene como objetivo fundamental conectar el genotipo con el fenotipo, esto significa identificar estas variantes



funcionales (*loci*, y los alelos de cada *locus*) ligadas a diferencias en un rasgo o carácter del fenotipo (Oraguzie y Wilcox, 2007).

Con avances recientes en la tecnología genómica y el interés por explotar la diversidad natural existente de los cultivos de importancia económica, como el maíz, y junto con el desarrollo de métodos robustos de análisis estadísticos, existe un gran interés en utilizar estudios de asociación genética para identificar genes responsables de la variación cuantitativa de variación fenotípica (Buckler y Thornsberry, 2002), y hasta hace poco se han identificado genes involucrados en la expresión de caracteres complejos como lo es la resistencia a la sequía (Setter et al., 2011). Sin embargo, la arquitectura genética de estos caracteres complejos es difícil de elucidar por el gran número de genes que lo pueden estar afectando, y los efectos que puedan estar causando cada uno de ellos. No obstante, hay un número importante de estudios de QTLs que se han publicado y se ha identificado un número limitado de genes que afectan a un carácter en cuestión (Salvi y Tuberosa, 2005).

Con la identificación de marcadores moleculares asociados a *loci* que codifican tanto para características cualitativas como para QTLs, se plantea la posibilidad de realizar selección asistida por marcadores, o comúnmente llamado mejoramiento molecular. Esta se basa en conjugar la variabilidad fenotípica y genotípica como fuente de información de la variabilidad existente y en la utilización de una variable genética como criterio selección, en este caso, marcadores moleculares asociados a la característica de interés (Ribaut y Hoisington, 1998; Dekkers y Hospital, 2002; Moose y Mumm, 2008). El mejoramiento molecular en términos generales es usado para describir varias estrategias modernas de mejoramiento que incluyen: (i) la Selección Asistida por Marcadores (MAS, por sus siglas en inglés), selección de alelos específicos para caracteres condicionados por pocos *loci*; ii) Retrocruzas Asistidas por Marcadores (MABC, por sus siglas en inglés), referente a la transferencia de un número limitado de *loci* de una base genética a otra, incluyendo transgenes; iii) Selección Recurrente Asistida por Marcadores (MARS, por sus siglas en inglés), se refiere a la identificación y selección de varias regiones genómicas involucradas en la expresión de caracteres complejos con el fin de desarrollar genotipos de comportamiento superior; y aunque se encuentra en un estado experimental

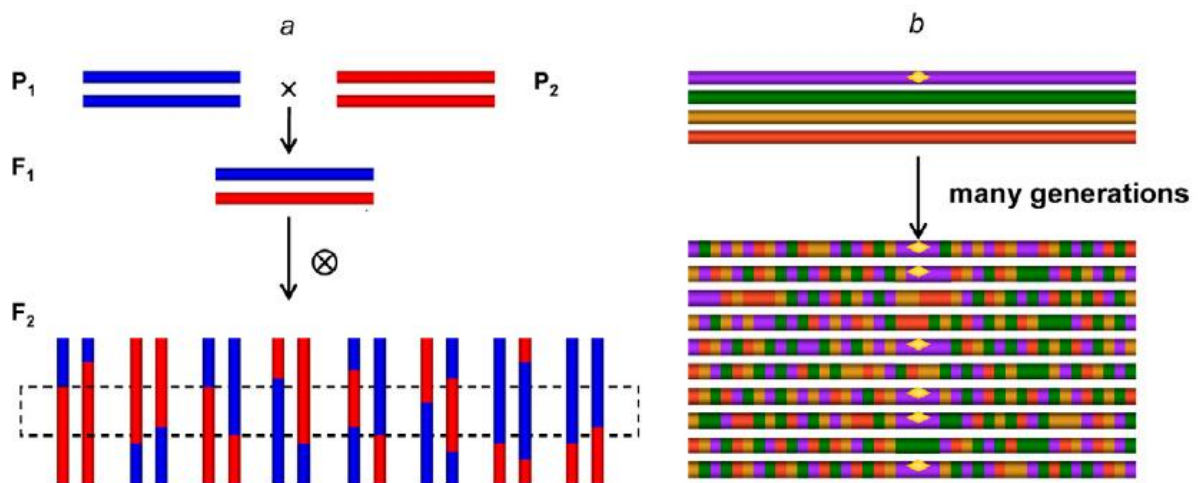
para plantas: iv) Selección Genómica (GWS, por sus siglas en inglés), es la selección basada en marcadores pero sin la identificación *a priori* de un conjunto de marcadores asociados con el carácter de interés (Ribaut et al., 2010, Delannay et al., 2012)

Sin embargo, la identificación de QTLs en un fondo genético específico (por ejemplo, población biparental) tiene un potencial predictivo muy limitado en otro fondo genético por la misma complejidad de parámetros poli génicos como la sequía. Por lo tanto la Selección Recurrente Asistida por Marcadores, MARS, (Eathington et al., 2005), o la re-estimación de los efectos de la variación alélica de los QTLs a través de los ciclos de selección en un programa de mejoramiento (Mapping as You Go, MAYG, por sus siglas en inglés, Podlich et al., 2004) son los métodos de selección molecular más eficientes para mejorar tolerancia a la sequía. Para entender las bases genéticas de parámetros complejos desde un punto de vista no reduccionista los análisis de ligamiento y mapeo por asociación representan una opción muy atractiva y han demostrado su potencia para discernir estos rasgos complejos (Figura 4.1). Las características generales de estos estudios involucran un panel o germoplasma en estudio que es caracterizado fenotípicamente y genotípicamente para detectar asociaciones estadísticas entre polimorfismos genéticos y la variación de caracteres heredables. Aunque los análisis de ligamiento y mapeo por asociación se basan en la co-herencia de polimorfismos funcionales y variantes de ADN, los dos métodos tienen sus diferencias (Zhu et al., 2008).

El ligamiento es la asociación física entre dos o más *loci*, lo que implica una baja recombinación y por lo tanto una mayor probabilidad de herencia conjunta, de otra manera, es la tendencia de los alelos de los *loci* que están cercanos entre sí a heredarse juntos como un bloque (haplotipo). Su análisis consiste en mapear los *loci* en individuos con parentesco y pedigrís de ascendencia conocida y encontrar un carácter que esté asociado al gen o grupo de genes. Este método es muy bueno para la evaluación de los efectos genéticos pero tiene la desventaja de baja resolución en el mapeo de la población (Zhu et al., 2008). De otro lado, el mapeo por asociación también es llamado “asociación genética” o “mapeo por desequilibrio de ligamiento”. Este enfoque utiliza los eventos de recombinación ancestrales en poblaciones naturales para hacer la asociación marcador-

fenotipo y aprovecha la diversidad genética natural; los métodos de asociación evalúan si determinados alelos en una población se encuentran en fenotipos específicos con mayor frecuencia de lo esperado (Flint-Garcia et al., 2005). Este enfoque tiene las ventajas de mayor resolución en el mapeo de la población y puede ir aumentando drásticamente, además de un número potencialmente elevado de alelos por *locus* puede ser estudiado al mismo tiempo (frente a tan sólo dos en el análisis de ligamiento), sin embargo este método no es lo mejor para identificar efectos genéticos en los *loci* de interés (Flint-Garcia et al., 2005; Zhu et al., 2008; McMullen et al. 2009).

**Figura 4.1** Comparación esquemática entre análisis de ligamiento (a) y análisis de mapeo por asociación (b); tomada de Zhu et al. (2008).



La actuales aplicaciones de los análisis de asociación incluyen: i) el mapeo por asociación de genes candidatos, que relaciona polimorfismos en los genes candidatos seleccionados que supuestamente han jugado un rol en el control de la variación fenotípica para caracteres específicos e involucra solo la secuenciación de los genes candidatos (Flint-Garcia et al., 2005; Zhu et al., 2008) y ii) la exploración del genoma que estudia la variación genética en el genoma para encontrar asociaciones para varios rasgos complejos (Rish y Merikangas, 1996), en este caso, se utilizan marcadores tipo SNP en todo el genoma a una densidad apropiada (Flint-Garcia et al., 2003).

Algunos de los estudios de asociación se han llevado a cabo en maíz, para tiempo de floración (Thornsberry et al., 2001), composición de grano (Wilson et al., 2004),

contenido de ácido oleico (Belo et al., 2008), y contenido de carotenoides (Harjes et al., 2008), enfermedades como el tizón norteño de la hoja (Poland et al., 2011), y hasta algunos genes de sequía (Setter et al., 2011), entre otros. Aunque estudios de asociación han permitido la identificación de varios *loci* para caracteres de interés agronómico, el impacto en programas de mejoramiento está todavía limitado y la validación de los alelos favorables identificados por asociación queda en una etapa crítica (Varsney et al., no publicado, 2012)

Para buscar las asociaciones marcador-carácter, dos metodologías han sido rutinariamente usadas, la primera está usualmente basada en la caracterización genotípica de una población completa segregante con marcadores que eventualmente cubren el genoma de la especie en estudio o solamente los genes que son de mayor interés, y una evaluación fenotípica a gran escala que exige ser detallada y precisa, y que muchas veces logísticamente es difícil para algunos caracteres ya que la adaptabilidad de los genotipos se confunden con su rendimiento (Xu et al., 2008). El segundo enfoque involucra una selección selectiva de individuos tomados de los extremos de la distribución fenotípica de una caracterización fenotípica detallada y precisa como en el primer caso (Lebowitz et al., 1987), para luego ser evaluados con marcadores moleculares, es decir, solo los individuos que representan los extremos altos y bajos de la distribución fenotípica son caracterizados genotípicamente (Xu et al., 2008). Esto significa una asociación marcador-carácter inferida por las diferencias en las frecuencias alélicas encontradas en los individuos extremos (Stuber et al., 1980).

La idea de tomar extremos de poblaciones en segregación para identificar alelos o genes de interés es conocida desde hace más de 20 años, se utilizó primero bajo el nombre de “*bulked segregant analysis*” y fue utilizada en la identificación de genes mayores para parámetros de calidad o genes involucrados en estreses bióticos y abióticos, por ejemplo en la identificación de marcadores ligados a la resistencia en cebada para *Rhynchosporium secalis* (Barua et al., 1993); detección y validación de QTLs, por ejemplo, la identificación de genes relacionados con la resistencia a la salinidad en tomate (Foolad y Jones, 1993). De manera más reciente autores han sugerido también el empleo de fenotipos contrastantes para identificar genes involucrados en caracteres más

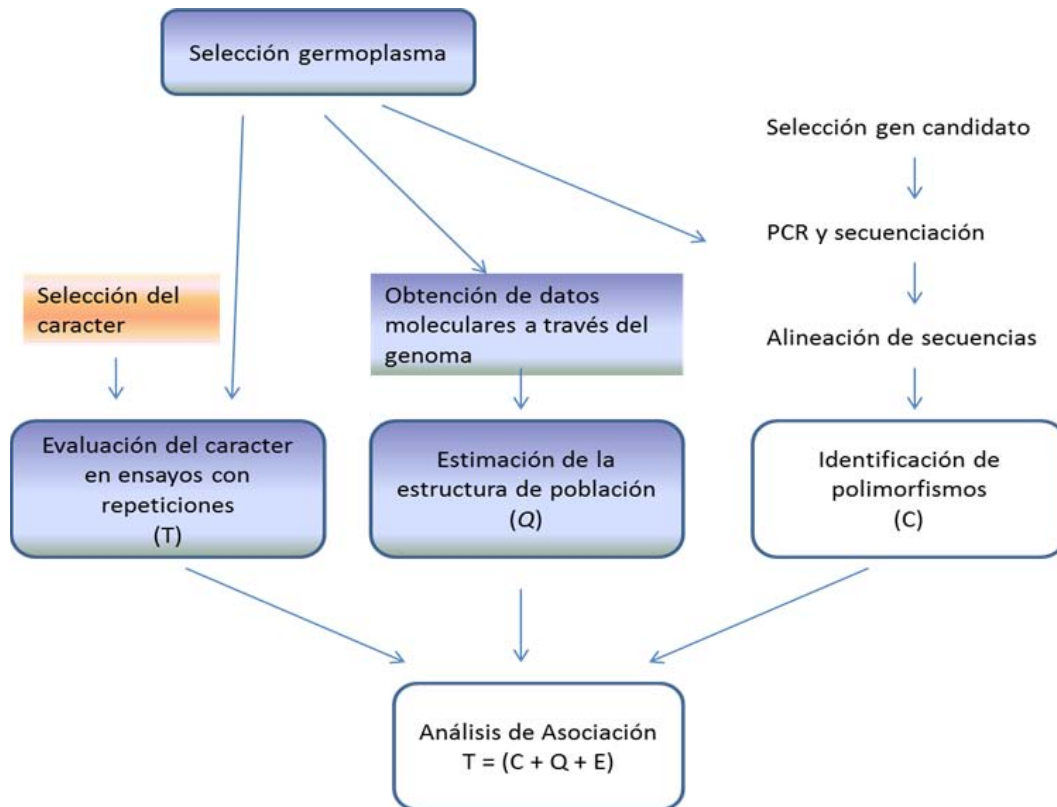
cuantitativos combinando métodos basados en asociación y Desequilibrio de Ligamiento (LD), sin embargo, aunque atractivo este método presenta algunas limitantes y todavía no está comprobado al 100% (Lu et al., 2011).

En cuanto a la optimización de la caracterización genotípica selectiva Gallais et al., (2007) ha indicado que para un tamaño de muestra dado de plantas caracterizadas fenotípicamente la proporción óptima de individuos para el estudio genotípico selectivo es alrededor del 30% para cada extremo de la distribución fenotípica. Esto podría indicar que las plantas escogidas en los extremos fenotípicos serían poseedoras de una acumulación de alelos favorables en múltiples *loci* con varios efectos aditivos (Xu et al., 2008). En nuestro caso, tomamos el principio de esta estrategia para explorar el descubrimiento de *loci* involucrados en el desempeño agronómico bajo condiciones de estrés hídrico y riego normales. Esta metodología puede ser potente cuando se aplica a accesiones del banco de germoplasma ya que cada accesión tiene su propia diversidad intrínseca y podrían encontrarse genes de adaptación a tolerancia a sequía y no solo de vigor. Entonces, tomando accesiones contrastantes desde un punto de vista fenotípico y combinándolo con datos genotípicos de diferentes regiones del genoma podría permitir la identificación de *loci* involucrados en la expresión de los parámetros de interés y cuáles de los alelos identificados en estos *loci* sean favorables o desfavorables desde un punto de vista agronómico.

#### **4.1.3 Consideraciones para un estudio de mapeo por asociación**

Cuando se va a iniciar un estudio de mapeo por asociación se deben considerar los aspectos genéticos de la especie, el germoplasma disponible, los recursos en caracterización genotípica o secuenciación y costos de la caracterización genotípica ya sea dirigida al mapeo para genes candidatos o a la exploración de todo el genoma (Figura 4.2). También se deben contemplar el diseño experimental para la caracterización fenotípica con réplicas a través de varios años y en múltiples localidades, todo dirigido a obtener una colección de datos de alta calidad lo que es esencial para esta clase de estudios, así como la identificación de polimorfismos candidatos juntamente con unas herramientas estadísticas potentes (Flint-Garcia et al., 2005; Yu et al., 2008).

**Figura 4.2** Esquema que ilustra los pasos involucrados en un estudio de mapeo por asociación, adaptada de Flint-Garcia et al. (2005). Pasos realizados en esta tesis doctoral son mostrados en azul, y discusión de especial interés en naranja.



La selección del germoplasma para un análisis de asociación es un punto crítico (Zhu et al., 2008), al tener un panel con pobre estructura poblacional y parentescos familiares evitaría las asociaciones falsas positivas que pueden ser generadas en el análisis de desequilibrio de ligamiento debido a la desigual distribución de alelos dentro de las poblaciones (Flint-Garcia et al., 2005). Por lo tanto, es importante determinar la subestructura genética si se tiene en un panel de asociación, porque se puede incluir en el modelo de análisis y se evitarían las asociaciones falsas positivas (Hansey et al., 2011), por esta razón es necesario la obtención de la matriz de ancestro  $Q$  (Flint-Garcia et al., 2005). Para el análisis de asociación se cuenta con una variedad de programas; TASSEL (Bradbury et al., 2007; Zhang et al., 2009) es el más comúnmente usado en plantas, y

frecuentemente es actualizado con los nuevos modelos que son desarrollados (Yu et al., 2008).

En cuanto al germoplasma a estudiar, o panel, generalmente las poblaciones de plantas pueden ser clasificadas para análisis de mapeo por asociación en uno de los siguientes 5 tipos de grupos (Yu y Buckler, 2006; Zhu et al., 2008): i) panel ideal, sin estructura poblacional y parentescos familiares, ii) panel con múltiples familias, iii) panel con estructura poblacional, iv) panel con estructura poblacional y parentescos familiares y v) panel con rigurosa estructura poblacional y relaciones familiares. Debido a las adaptaciones locales, selección, historia en mejoramiento, para muchas de las especies de plantas en el panel de asociación serán clasificadas de tipo cuatro. Aunque también se pueden clasificar de acuerdo a la fuente del material como colecciones provenientes de un banco de germoplasma, poblaciones sintéticas o germoplasma elite (Bressegello y Sorrells, 2006).

Un mapeo por asociación con germoplasma diverso puede identificar alelos “superiores” que todavía no han sido capturados en el desarrollo de un programa de mejoramiento y puede mantener la introgresión de estos alelos en las líneas mejoradas (Zhu et al., 2008), aunque demasiada diversidad podría generar un problema de adaptabilidad, por ejemplo, algunos materiales no florearían ni alcanzarían su madurez en el ambiente de evaluación (Hansey et al., 2011). En relación al tamaño de muestra, los estudios realizados en maíz son diversos, para tiempo de floración 375 líneas y 275 razas (Camus-Kulandaivelu et al., 2006); en composición de grano 102 líneas diversas (Wilson et al., 2004); contenido de ácido oleico, 553 líneas elite (Belo et al., 2008), y en contenido de carotenoides 282 líneas diversas (Harjes et al., 2008). La mayoría de los estudios anteriores en mapeo por asociación se han basado en el análisis de líneas endogámicas, es decir, líneas que tienen ascendencia común o pertenecen a la misma familia o linaje, siendo material homocigoto y de fácil caracterización genotípica. De manera general un número mínimo de 300 genotipos contrastantes parece ser más apropiado para estudios de asociación de parámetros cuantitativos cuando se trata de accesiones de Banco de Germoplasma o poblaciones la variabilidad genética se encuentra a dentro y a través de las accesiones.

Enfoques multidisciplinarios que vinculen diferentes esquemas de mejoramiento, fisiología y genética molecular pueden llevar a entender mejor sobre la respuesta del maíz al déficit de agua (Ribaut et al., 2004). Con una mejor comprensión de los mecanismos fisiológicos y la base genética de la respuesta del maíz a la sequía, debe ser cada vez más posible identificar, seleccionar y transferir genes y alelos clave para construir genotipos con la tolerancia a la sequía (Ribaut et al., 2008). De otro lado, estudiar las variaciones genéticas y el papel de éstas en los caracteres importantes presentes en la diversidad genética disponible de maíz como en los maíces tropicales, podrían ayudar a determinar el papel de genes en algunos rasgos agronómicos adicionales al rendimiento como resistencia a sequía, potencial del maíz para producir etanol, eficacia en uso de fertilizantes, entre otros. De otro lado, mejorar y ampliar la base genética del germoplasma base utilizada hasta ahora en mejoramiento podría impactar de manera significativa la producción de maíz en el futuro (Salwana y Pollack 2006).

La exploración de la diversidad genética existente en las poblaciones nativas de maíz no ha sido muy examinada en estudios de asociación. Sin embargo ahora se está despertando un interés generalizado en el estudio específico de caracteres de interés en los materiales nativos que en su conjunto son poseedores de adaptación a múltiples condiciones ambientales y agroecológicas. En este sentido, resultados obtenidos en la evaluación de las características agronómicas en más de 14.000 accesiones de maíces Latinoamericanos y de Estados Unidos en los proyectos LAMP y GEM (Latin American Maize Project y Gemplasm Enhancement of Maize, por sus siglas en inglés respectivamente) indicaron que alelos de fuentes de germoplasma exótico pueden ser combinados con genes de líneas elite de la franja maicera de los Estados Unidos para generar nuevos materiales útiles para programas de mejoramiento (Salwana y Pollack, 2006).

Con el propósito de capturar una substancial diversidad molecular y funcional presentes en los maíces tropicales, además de explorar la utilización de poblaciones heterogéneas de maíz para un estudio de mapeo por asociación para tolerancia a la sequía, en esta tesis, se ha estudiado la diversidad genética de un panel de germoplasma diverso (70 accesiones en total) y la variabilidad fenotípica bajo condiciones de estrés y regadío normal en una sola localidad como punto de referencia. Dada la evidencia y la



importancia de la utilización de caracteres secundarios como criterios de selección en material tolerante a la sequía, se calculó un Índice de Selección basado en componentes de rendimiento y caracteres fisiológicos con el fin de tener una distribución fenotípica que permitiera seleccionar las mejores y peores accesiones con respuesta adaptativa. Seguidamente también se hace una discusión amplia sobre el direccionamiento de estudios de asociación que podrían llevarse a cabo en el panel estudiado apoyada por los análisis de variabilidad fenotípica calculada debido a la estructura de población y heredabilidad.

## **4.2 Materiales y Métodos**

### **4.2.1 Germoplasma**

Para este estudio se seleccionaron 56 accesiones de poblaciones nativas de maíz tropical precoz, de la colección resguardada en el Banco de Germoplasma de Maíz del CIMMYT y colectadas en México, Centroamérica, Sudamérica y el Caribe (Tabla 4.1). Esta selección se basó en la colección nuclear de 120 poblaciones de maíces tropicales obtenida de los ensayos de evaluación y caracterización en las estaciones experimentales de Agua Fria-Puebla y Tlaltizapán-Morelos en México. Estas poblaciones fueron evaluadas en la estación seca y bajo dos regímenes de riego; tomándose en cuenta diversidad, caracteres agro-morfológicos como la floración, características de planta y mazorca, largo y ancho de grano, senescencia, resistencia a enfermedades y rendimiento (detalles de la clasificación en razas de éstas poblaciones nativas se encuentra más adelante en la Tabla 4.9). Estos ensayos previos fueron conducidos por el Banco de Germoplasma del CIMMYT.

Al conjunto inicial de accesiones, se adicionaron 10 accesiones pertenecientes a los complejos genéticos o *poles* creados por el Banco de Germoplasma en su programa de pre-mejoramiento y 4 accesiones desarrolladas por el grupo de fisiología del Programa Global de Maíz de CIMMYT hasta completar un total de 70 accesiones. En la Tabla 4.2 se encuentra la descripción detallada del germoplasma proveniente de los programas de pre-mejoramiento del Banco de germoplasma y del Programa de Fisiología del CIMMYT.

**Tabla 4.1** Fuentes de origen del panel de maíces tropicales precoces de interés para estudios de asociación de tolerancia a la sequía

<u>Fuentes del germoplasma</u>	<u>Número de accesiones</u>
Argentina	6
Antillas	2
Brasil	5
Cuba	1
Haití	1
México	28
Paraguay	4
Puerto Rico	4
República Dominicana	1
Salvador	1
Uruguay	1
Venezuela	2
CIMMYT	14
Total	70

#### **4.2.2 Condiciones experimentales**

Los experimentos de campo se llevaron cabo en la estación experimental de Tlaltizapán en el Estado de Morelos México (18°41'N, 99°07'W, 940 metros sobre el nivel del mar). Los datos climáticos fueron documentados en la estación meteorológica de Tlaltizapán para el periodo del experimento (Tabla 4.3). Los ensayos fueron sembrados el 29 del mes de Noviembre del 2007 en parcelas de una accesión por surco, de 5 m de longitud con una distancia de 0,25 m entre plantas y una separación entre surcos de 0,75 m, lo que representó una densidad de 6,67 plantas por m<sup>2</sup>. Las plantas se cosecharon el 8 del mes de Abril del 2008 a los 131 días después de la siembra. El diseño experimental fue un Lattice de bloques incompletos al azar con 3 repeticiones por accesión. Se sembraron dos semillas por golpe y se raleó una unos 14 días después de la germinación.

##### **4.2.2.1 Régimen de riego**

Se aplicó un primer riego a la siembra realizado por aspersión para los dos regímenes de riego contrastado evaluados y así asegurar la germinación de todo el ensayo.

**Tabla 4.2.** Descripción del germoplasma de los programas de pre-mejoramiento del banco de Germoplasma y del Programa de Fisiología del CIMMYT\*.

Tipo de material	Descripción general	Descripción del germoplasma de Origen	Comentarios
<b>Material programa de pre-mejoramiento Banco de Germoplasma</b>			
Pool 21	Tropical Intermedio Cristalino Amarillo	Basado en materiales de México, Cuba, República Dominicana, India, Colombia, Argentina, Costa Rica, Uruguay, San Vicente y Surinam.	El énfasis de la selección se hace para mejorar el precio de venta y la resistencia al barrenador de la caña de azúcar ( <i>Diatrea saccharalis</i> ).
Pool 22	Tropical Intermedio Dentado Amarillo	Basado en materiales de México, Cuba, República Dominicana, Antigua, Brasil, Perú, India, Colombia, Puerto Rico, Argentina y varios países centroamericanos. Argentina, Costa Rica, Uruguay, San Vicente y Surinam.	Tiene buen rendimiento potencial y tipo de planta, se le mejora para la resistencia a la pudrición del tallo.
<b>Pool 21 x Pool 22 (♂)</b>	Cruza	Pool 21 y Pool 22	Híbrido que sirve como referencia para evaluar y seleccionar las mejores líneas de los pools que lo forman
<b>Pool 21 x Pool 22 (Best testcrosses)</b>	Cruza	Pool 21 y Pool 22	Compuesto balanceado de las mejores cruza
<b>Pool 22 x Pool 21 (Best testcrosses)</b>	Cruza	Pool 21 y Pool 22	Compuesto balanceado de las mejores cruza
Pool 20	Tropical Intermedio Dentado Blanco	Basado en materiales principalmente de Filipinas, India, sureste de Asia y una pequeña fracción de México y varios países centroamericanos.	Se le mejora para incrementar la resistencia a la pudrición de la mazorca
CML 494	Línea autógama desarrollada en CIMMYT Tropical Intermedio Dentado Blanco	Basado en poblaciones de La Posta Sequía	
<b>Pool 20 x CML-494 (Best testcrosses)</b>	Compuesto balanceado de las mejores cruza	Pool 20 y CML 494	
Pool 19	Tropical Intermedio Cristalino Blanco	Cerca del 30% de este pool está formado por materiales resistentes al Mildiú filipino del maíz causado por <i>Peronosclerospora philippinensis</i> , también contiene materiales de Cuba, México, Argentina, Honduras, El Salvador, Colombia, Ecuador, India, Centro y Sureste de Estados Unidos.	El énfasis de la selección se hace para mejorar el precio de venta y la resistencia al barrenador de la caña de azúcar ( <i>Diatrea saccharalis</i> ).

## Continuación Tabla 4.2

Tipo de material	Descripción general	Descripción del germoplasma de Origen	Comentarios
<b>Material programa de pre-mejoramiento Banco de Germoplasma</b>			
CML 495	Línea autógama desarrollada en CIMMYT Tropical Tardío Cristalino Blanco		
<b>Pool 19 x CML-495 (Best testcrosses)</b>	Cruza	Pool 19 y CML 495	Compuesto balanceado de las mejores cruzas
Pool 26	Tropical Tardío Dentado Amarillo	Basado en materiales de México, Colombia, área del Caribe, y países centroamericanos, así como también de una fracción del germoplasma de la faja maicera de los Estados Unidos	Buen rendimiento potencial, se le mejora para obtener la resistencia al gusano cogollero ( <i>Spodoptera frugiperda</i> )
CML 451	Línea autógama desarrollada en CIMMYT Tropical Tardío Semicristalino Amarillo		
<b>Pool 26 x CML-451 (Best testcrosses)</b>	Cruza	Pool 26 y CML 451	Compuesto balanceado de las mejores cruzas
Pool 23	Tropical Tardío Cristalino Blanco	Basado en selecciones a partir de cruzas entre materiales de México, Colombia, área del Caribe, Guatemala, Panamá, y otros países centroamericanos, India, Tailandia y Filipinas	Excelente rendimiento y se le mejora para incrementar la resistencia a la pudrición del tallo
Pool 24	Tropical Tardío Dentado Blanco	Basado principalmente en materiales pertenecientes a la raza Tuxpeño de México y algunos materiales de países Centroamericanos	Relativamente alta resistencia a la pudrición del tallo y mazorca, excelente tipo de planta y rendimiento potencial, se le mejora específicamente para incrementar la resistencia al gusano cogollero ( <i>Spodoptera frugiperda</i> ).
<b>Pool 23 x Pool 24 (♂)</b>	Cruza	Pool 23 y Pool 24	Híbrido que sirve como referencia para evaluar y seleccionar las mejores líneas de los pools que lo forman
<b>Pool 24 x Pool 23 (Best testcrosses)</b>	Cruza	Pool 23 y Pool 24	Compuesto balanceado de las mejores cruzas

## Continuación Tabla 4.2

Tipo de material	Descripción general	Descripción del germoplasma de Origen	Comentarios
<b>Material programa de pre-mejoramiento Banco de Germoplasma</b>			
<b>Pool 23 x Pool 24 (Best testcrosses)</b>	Cruza	Pool 23 y Pool 24	Compuesto balanceado de las mejores cruzas
<b>Tuxpeño Crema 1</b>	Población Tropical Tardía Dentado Blanco	Compuesta principalmente por colecciones de la raza Tuxpeño de México,	Tolerante a la mayoría de las enfermedades foliares
<b>Material programa de Fisiología</b>			
<b>OPV-DTPYDR</b>	Variedad de Polinización Abierta, tolerante a sequía y altas temperaturas Tropical Intermedio Amarillo Dentado	DTP-1 y DTP-2. Poblaciones intermedias de color y textura de grano mixtos, estas a su vez derivadas de poblaciones seleccionadas por alta resistencia a diferentes estreses en general, susceptibles a algunas enfermedades foliares y resistentes al acame.	Están siendo mejoradas especialmente para tolerancia a sequía y altas temperaturas con buenos resultados
<b>OPV-DTPWDR</b>	Variedad de Polinización Abierta, tolerante a sequía y altas temperaturas Tropical Intermedio Blanco Dentado	Derivada de la Población 43- La Posta, población derivada de recombinación de líneas de Tuxpeño	Mejorada especialmente para tolerancia a sequía y altas temperaturas
<b>OPV-LPSDR</b>	Variedad de Polinización Abierta tolerante a sequía y altas temperaturas	Sintético combinado de OPV-DTPW y LPSDR para sequía	Está siendo mejorada especialmente para tolerancia a sequía y altas temperaturas con buenos resultados

El riego de control, fue regado 10 veces por gravedad en un promedio de 5 horas durante todo el ciclo de cultivo. El régimen con déficit de agua fue regado por goteo 6 veces durante todo el ciclo de cultivo, a saber: el primer riego se realizó en el día 22 después de la siembra (DDS<sub>22</sub>) por 3 horas; el segundo, 46 días después de la siembra por 5 horas (DDS<sub>46</sub>); el tercero, 64 días después de la siembra por solo una hora (DDS<sub>64</sub>). El cuarto riego fue realizado en 76 días después de germinación (DDS<sub>76</sub>), el quinto y sexto riego fueron realizados en los días por 6 y 3 horas respectivamente, en los días 88 y 100 días después de la siembra (DDS<sub>88</sub> y DDS<sub>100</sub>). Estas relaciones de riego corresponden aproximadamente a un total de 1500 mm y 325 mm para el riego de control y régimen con déficit de agua respectivamente. Durante todo el ciclo de cultivo no se recibió precipitación. Las medias de todas las accesiones para antesis fueron 92 y 88 días después de la germinación para condiciones de regadío normal y condiciones de estrés respectivamente.

**Tabla 4.3.** Registros climatológicos para el periodo comprendido entre Diciembre 2007 y Abril 2008 en la Estación experimental del CIMMYT de Tlaltizapán (Edo. de Morelos, México). Se presentan datos de temperatura máxima, mínima y media (T), humedad relativa máxima y mínima (HR), flujo de radiación máxima y mensual para cada mes (Rad), y promedios mensuales de precipitación (P).

Mes	P (mm)	T <sub>max</sub> (°C)	T <sub>min</sub> (°C)	T <sub>med</sub> (°C)	HR <sub>max</sub> (%)	HR <sub>min</sub> (%)	Rad <sub>max</sub> (KW m-2)	Rad <sub>total</sub> (MJ m-2)
Diciembre	0,0	32,6	8,7	20,8	96,2	14,2	0,9	15,9
Enero	0,0	33,3	6,0	20,3	92,0	12,1	1,1	17,8
Febrero	0,0	33,2	11,2	22,2	87,3	18,0	1,1	20,3
Marzo	0,0	37,1	6,3	21,7	79,5	5,9	1,4	22,6
Abril	0,0	37,7	13,5	25,6	72,8	8,7	1,0	23,6

#### 4.2.2.2 Caracteres agronómicos y floración

Las fechas de floración masculina y femenina fueron medidas como número de días transcurridos desde la siembra y el momento en que el 50% de las plantas en cada bloque liberaban polen o tenían los estigmas visibles respectivamente. El ASI, se calculó como la

diferencia entre la floración femenina y la floración masculina. Las plantas fueron cosechadas a mano en la maduración fisiológica y seguidamente se midió el rendimiento de grano (ajustado a 15% de humedad y un porcentaje de desgrane del 80%) y sus componentes agronómicos (mazorcas por planta, granos por mazorca, granos por hectárea, y el peso de 100 granos).

Se registraron las alturas de planta, alturas medias de las mazorcas medidas desde la base de la planta al nudo que porta la mazorca más alta; se contaron el número de plantas con acame de raíz y tallo, número de plantas y mazorcas cosechadas, además del número de mazorcas con pudrición. Los caracteres cualitativos se calificaron en una escala de 1-5; en aspecto de planta se califican características como altura de planta y mazorca, uniformidad de las plantas, daño de insectos y enfermedades, y acame, donde 1 es excelente y 5 deficiente. En aspecto de mazorca, calificación 1 indica que las mazorcas cosechadas están libres de enfermedades e insectos, con tamaño uniforme y buen llenado de grano; calificación 5 indica que las mazorcas son deficientes en los aspectos evaluados. En textura y dureza de grano, una calificación 1 indica mazorcas con granos duros y calificación 5 granos suaves. Pudrición de mazorca, calificación 1 indica que no hubo pudrición, mientras que una calificación de 5 indica severa pudrición, y de manera análoga se califica enrollamiento de hoja. De manera adicional se marcaron 5 plantas en cada surco con el fin de tener una caracterización fenotípica por individuo, es decir, por cada accesión se evaluaron de manera individual 15 plantas (5 plantas por repetición). Los parámetros estudiados en esas 15 plantas fueron: los mismos componentes agronómicos de rendimiento, clorofila medida con un aparato portátil, fechas de floración masculina y femenina, altura de planta y mazorca, y peso de tallos.

#### **4.2.2.3 Crecimiento y senescencia de plantas**

La senescencia durante el llenado de grano se evaluó midiendo el contenido de clorofila en la hoja por encima de la mazorca, en un muestra de 5 plantas por surco y repetición, usando un clorofilómetro portátil (SPAD-502, Minolta, Tokio, Japón) el cual mide las diferencias de la densidad óptica a dos longitudes de onda (650 y 940 nm). Se llevaron a cabo dos

mediciones, la primera, se realizó 98 días después de la siembra, (DDS<sub>98</sub>), y la segunda, en un intervalo de 20 días, en el día 118 después de la siembra (DDS<sub>118</sub>).

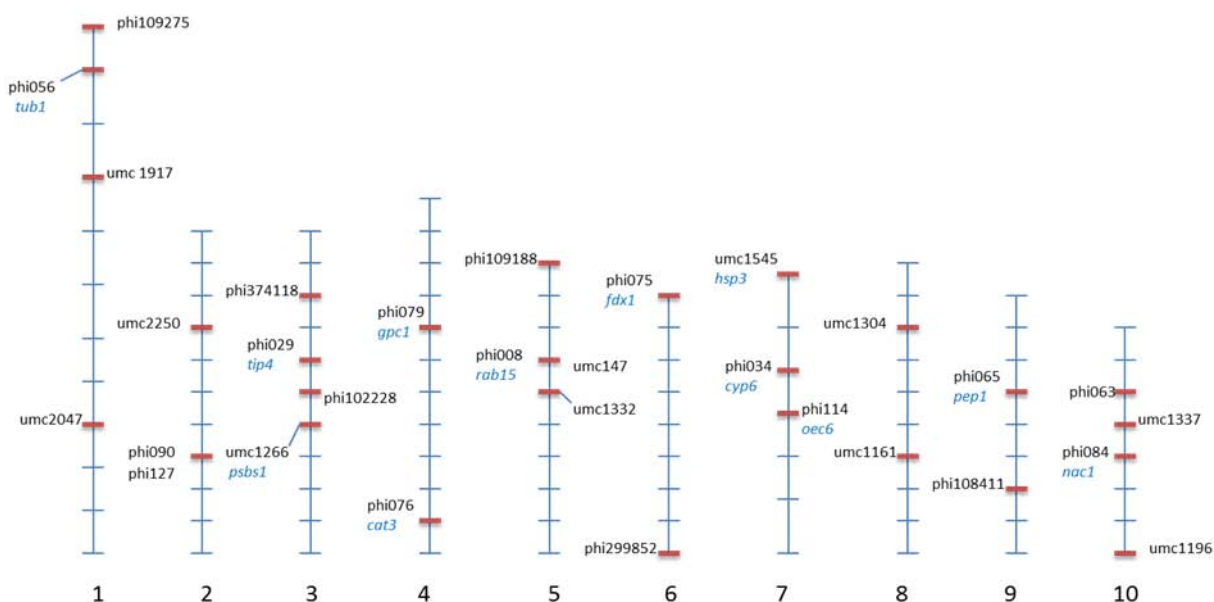
La senescencia también se evaluó visualmente durante el llenado de grano, en dos tomas, la primera, 102 días después de la siembra (DDS<sub>102</sub>) y la segunda 120 días después de la siembra (DDS<sub>120</sub>), en la base a las hojas aun verdes en la planta. Esta medición se realizó en una escala de 1-10, dividiendo por 10 el porcentaje de área foliar total estimada que ha muerto, es decir, 1 = 10% del área foliar muerta y 10= 100% área foliar muerta.

#### **4.2.3 Caracterización genotípica**

Para el análisis genético fueron seleccionados 30 microsatélites (Figura 4.3, Anexo 7.3). En este caso se realizaron las extracciones de ADN para las 70 accesiones con el mismo protocolo que para muestras conjuntas para su análisis de diversidad genética y estructura poblacional siguiendo el procedimiento descrito en detalle el capítulo 3, ([http://www.cimmyt.org/english/docs/manual/protocols/abc\\_amgl.pdf](http://www.cimmyt.org/english/docs/manual/protocols/abc_amgl.pdf), página 7, se pueden consultar detalles también en Anexo 7.2). De manera simultánea pero independiente se extrajo el ADN para cada una de las 5 plantas seleccionadas por surco (caracterizadas fenotípicamente) para su probable empleo posterior en el estudio de mapeo por asociación. Las reacciones de PCR se llevaron a cabo utilizando el mismo protocolo descrito en el capítulo anterior y descrito en: [http://www.cimmyt.org/english/docs/manual/protocols/abc\\_amgl.pdf](http://www.cimmyt.org/english/docs/manual/protocols/abc_amgl.pdf), (página 48, ver detalles también en Anexo 7.4). Se utilizó el secuenciador automático de capilares ABI 3100 (Applied Biosystems, Foster City, CA) para determinar el tamaño de los fragmentos en pares de bases, productos de la reacción de PCR. Los tamaños de los fragmentos fueron calculados de manera automática con el programa GeneScan® 3.1 (PerkinElmer/Applied Biosystems, Foster City, CA) y Genotyper® 2.1 (PerkinElmer/Applied Biosystems, Foster City, CA), de igual manera que el capítulo anterior fueron utilizados los programas Freqs-R (Franco et al., 2005) y FtoL-R (Franco et al., 2007) para obtener finalmente los individuos dialélicos simulados para cada uno de los 15 individuos que componen cada una de las muestras conjuntas de ADN (detalles en la sección 3.2.1).



**Figura 4.3** Esquema de los microsatélites utilizados a través de todos los cromosomas y se muestran los genes de los que se han derivado algunos de ellos. Información obtenida de: <http://www.maizegdb.org/>



#### 4.2.4 Análisis de datos genotípicos

Se calcularon para el panel inicial (70 accesiones), las mejores y peores 20 accesiones a través de las 70 a ccesiones (panel seleccionado), y para los diversos grupos que lo componen según clasificación de material (*pooles* de pre-mejoramiento del Banco de Germoplasma, material mejorado del Subprograma de Fisiología de Maíz, y las poblaciones nativas): índices de diversidad genética como heterocigocidad esperada, Índice de Fijación ( $F_{ST}$ ) por parejas (solo para grupos según clase de material), y número de alelos encontrados utilizando el programa PowerMarker (Liu y Muse, 2005). El programa estadístico Darwin 5.0 (Perrier y Jacquemoud-Collet, 2006) fue usado para obtener el Análisis de Coordenadas Principales (ACoP) basado en la matriz de distancias genéticas obtenida con PowerMarker.

El programa Structure 2.2 (Pritchard et al., 2000) fue utilizado para analizar los datos generados con los microsatélites y evaluar la estructura poblacional. El análisis de

estructura fue analizado para panel inicial, accesiones de poblaciones nativas en panel inicial, panel seleccionado (los extremos) y nativas en panel seleccionado. Se utilizó el valor de mínimo ancestro para pertenecer a un grupo del 51%. No se les asignó *a priori* la pertenencia a ninguno de los grupos, a cada población se le permitió variar ( $K=1$  hasta  $K=12$ ) y agrupar con los individuos que estuvieran más relacionados genéticamente basados en los datos de microsatélites. No fue posible determinar el valor de  $K$  directamente y se utilizó nuevamente el procedimiento reportado por Evanno et al. (2005).

#### 4.2.5 Análisis de datos fenotípicos

El cálculo del Índice de Selección (IS) se realizó tomando en cuenta la evaluación fenotípica del germoplasma en condiciones de estrés, con el propósito de tener una distribución fenotípica de las 70 accesiones iniciales (panel inicial), y seleccionar 40 accesiones de los extremos (panel seleccionado). Este panel seleccionado estaría representado por 20 accesiones de cada uno de los extremos de la distribución inicial catalogadas como las mejores y peores accesiones con respuesta adaptativa bajo régimen de estrés y que representara aproximadamente el 50% del panel original. El cálculo del IS, se basó en el método de valores y vectores característicos (ESIM, por sus siglas en inglés), que involucra el primer componente principal (CP) y el vector y la raíz características de la matriz de varianza-covarianza fenotípica de los caracteres de interés, donde los elementos del vector característico del primer CP determinan la proporción con la que los caracteres respectivos contribuyen al Índice de Selección (Cerón-Rojas y Sahagún-Castellanos 2005; Cerón-Rojas et al., 2006). De este modo:

$$(S - \lambda I) \beta = 0$$

Donde  $S$  es la matriz de varianza-covarianza fenotípica de los caracteres de interés,  $\lambda$  y  $\beta$  son respectivamente el vector característico y raíz característica de  $S$ , los elementos de los vectores característicos son los coeficientes de los valores fenotípicos. El IS se calculó con un programa (ESIMIndex) desarrollado en CIMMYT y escrito en código R (<http://cran.r-project.org>). En este respecto se determinó un ideal tipo con el objetivo de obtener una selección dirigida que reuniera características como incremento en rendimiento de grano, mazorcas por planta y contenido de clorofila y una disminución en ASI, enrollamiento de

hojas y senescencia, los pesos asignados fueron +5, +3, +1, -2, -1 y -2 respectivamente. Esta combinación de parámetros de rendimiento y parámetros secundarios críticos para la evaluación de sequía permite una evaluación integrada de la tolerancia a sequía.

Como un intento de la estimación de la tolerancia a la sequía del material para rendimiento de grano sin los efectos del rendimiento potencial, se calculó el Índice de Tolerancia a la Sequía, ITS, (DTI, por sus siglas en inglés) basado en el índice simple propuesto por Fischer y Maurer (1978) y está definido como la razón entre el rendimiento en condiciones de estrés (CE) y el rendimiento en condiciones normales (CN):

$$ITS = \frac{\text{Rendimiento (CE)}}{\text{Rendimiento (CN)}}$$

La heredabilidad en sentido amplio ( $H^2$ ) para una sola localidad se calculó utilizando el programa Fieldbook IMIS, que usa un código R (<http://cran.r-project.org>), propiedad de CIMMYT, Int, y disponible en: <http://dtma.cimmyt.org/index.php/information-tools/software-download>, utilizando la siguiente ecuación:

$$H^2 = \frac{\sigma_g^2}{\sigma_p^2} = \frac{\sigma_g^2}{\sigma_g^2 + \frac{\sigma_E^2}{nr}}$$

Donde:

$\sigma_g^2$ : corresponde al componente de la varianza genotípica

$\sigma_p^2$ : componente de la varianza fenotípica

$\sigma_E^2$ : componente de la varianza del error

$nr$ : número de repeticiones.

Los cálculos de heredabilidad para los diferentes caracteres, fueron conducidos clasificando las accesiones para condiciones de estrés de la siguiente forma: i) para el panel inicial (70 accesiones); ii) accesiones de poblaciones nativas presentes en panel inicial (56 accesiones) y iii) accesiones de material mejorado presentes en el panel inicial (14 accesiones). La

heredabilidad se calculó para rendimiento de grano, ASI, número de mazorcas por planta, contenido de clorofila (SPAD<sub>98DDS</sub> y SPAD<sub>118DDS</sub>), enrollamiento de hoja (EH<sub>104DDS</sub>), y senescencia (Sen<sub>102DDS</sub> y Sen<sub>120DDS</sub>). Para el ensayo de regadío normal se calculó la heredabilidad para rendimiento, ASI y número de mazorcas por planta; en este caso para el total de las accesiones y para las 56 accesiones de poblaciones nativas.

Los efectos de la variabilidad fenotípica debido a la estructura de población fueron analizados mediante un modelo estadístico de regresiones lineales múltiples utilizando el software de análisis GenStat (13 edición), bajo dos perspectivas. La primera de ellas, teniendo en cuenta solamente la información de número de grupo al que pertenecieron las accesiones, este número de grupo es dado por el análisis de estructura de población (ver detalles adelante) y utilizando este número como co-variable. La segunda perspectiva implicó el empleo de los componentes de la matriz de proporción ( $Q_{l-k}$ ) correspondiente al número de grupo definido en el punto anterior y para cada caso la matriz  $Q$  se extrajo de los análisis de estructura de población y fue utilizada también como co-variable. Estos análisis fueron estudiados para el panel inicial, accesiones nativas totales presentes en el panel inicial, panel seleccionado y nativas presentes en panel seleccionado. Los efectos de variabilidad fenotípica debida a la estructura de población ( $R^2$ ) fueron evaluados para rendimiento de grano, ASI, número de mazorcas por planta, contenido de clorofila (SPAD<sub>98DDS</sub> y SPAD<sub>118DDS</sub>), enrollamiento de hoja (EH<sub>104DDS</sub>), Senescencia (Sen<sub>102DDS</sub> y Sen<sub>120DDS</sub>). Teniendo presente que la matriz  $Q$  no es el único camino para evaluar la estructura de población, pues recientemente se ha propuesto como una rápida y eficiente alternativa usar el Análisis de Componentes/Coordenadas Principales como diagnóstico de la estructura de poblacional (Patterson et al., 2007), por este motivo, de manera adicional y sólo como comprobación se realizó un análisis de prueba para el panel inicial tomando en cuenta la matriz de datos del décimo componente principal, utilizados también como co-variables para el análisis de variabilidad fenotípica debido a la estructura poblacional para los mismos caracteres mencionados.

## **4.3 Resultados**

### **4.3.1 Rendimiento promedio de los diferentes tipos de accesiones**

En la Tabla 4.4 se exponen los rendimientos promedio y la desviación estándar, valor máximo y mínimo para las accesiones y poblaciones nativas obtenidos en los ensayos bajo estrés y condiciones de regadío normal.

**Tabla 4.4** Rendimientos promedio (Rend) en  $\text{tm ha}^{-1}$  y desviación estándar correspondiente, valores máximo y mínimo del rendimiento para cada material en condiciones de estrés y regadío normal tomando en cuenta todo el germoplasma (general), se presentan también los valores de rendimiento para condiciones de estrés pero únicamente tomando en cuenta las accesiones con antesis entre 85 y 95 días.

Material	Condiciones de estrés						Regadío normal		
	General			85 ≤ Antesis ≤ 95 días			General		
	Rend $\text{tmha}^{-1}$	Máx	Mín	Rend $\text{tmha}^{-1}$	Máx	Mín	Rend $\text{tmha}^{-1}$	Máx	Mín
Banco Germoplasma	0,34 ± 0,23	0,88	0,09	0,25 ± 0,16	0,36	0,15	6,74 ± 1,83	9,07	4,90
Subprograma Fisiología	0,68 ± 0,25	0,87	0,47	0,68 ± 0,25	0,87	0,47	6,77 ± 1,82	8,62	5,09
Materiales nativos	0,07 ± 0,07	0,30	0,00	0,05 ± 0,07	0,20	0,00	3,31 ± 1,80	6,79	0,85

Para tener una evaluación adecuada del rendimiento bajo sequía, eliminando el efecto del escape de sequía que sucede cuando las accesiones tienen una floración precoz y por lo tanto no están realmente expuestas al estrés durante la floración, y el efecto contrario se aplica a las accesiones tardías que no producen nada por un estrés demasiado fuerte, se eliminaron los extremos de la distribución para floración, es decir, con este ejercicio se generó una pérdida promedio del rendimiento para los *pooles* del Banco de Germoplasma de 26% y para las poblaciones nativas de 29%, pues se eliminaron 3 y 18 accesiones de los *pooles* del Banco y de las poblaciones nativas respectivamente. Lo mas importante es que se eliminaron las accesiones que presentaban un rendimiento alto debido a un escape al estrés como se demostró a través de la disminución de los extremos mas altos de los rendimientos, en particular para las accesiones del banco por las cuales se obtuvo una reducción de la máxima de 0.88 a 0.36  $\text{tm ha}^{-1}$ .

Comparando los resultados del rendimiento en condiciones de estrés y condiciones normales es evidente que la intensidad del estrés fue severa ya que en promedio el rendimiento en condición de sequía corresponde a un 10% del rendimiento en condiciones normales. Con la restricción en antesis, el máximo rendimiento en las poblaciones nativas basado es un poco a bajo del promedio de las poblaciones del Banco y de un tercio del promedio de los materiales mejorados de Fisiología. Con esta intensidad de estrés severa, pero con todavía una buena segregación del rendimiento bajo sequía, y más que todo de los parámetros secundarios, se puede asegurar que los mecanismos de tolerancia a sequía están activados y contribuyen al rendimiento en conjunto con los mecanismos que contribuyen al vigor de las plantas.

#### **4.3.2 Clasificación de las accesiones según Índice de Selección evaluado bajo condiciones de estrés**

La distribución fenotípica del rendimiento y desempeño agronómico para las 70 accesiones se obtuvo de la caracterización fenotípica a través de todas las accesiones y como resultado del cálculo del Índice de Selección bajo condiciones de estrés. En la columna IS-número (Tabla 4.5), las accesiones se presentan en un rango de 1 hasta 70, donde las accesiones con IS-Número 1, e IS-Número 70 representan la mejor y peor accesión con respuesta adaptativa según el Índice de Selección bajo condiciones de estrés. La columna IS+Antesis (Tabla 4.5), expresa la selección de 40 accesiones, en orden ascendente, de 1 hasta 20 están representadas las 20 mejores y en orden descendente de 20 hasta 1, las 20 peores (panel seleccionado). Este panel seleccionado se clasificó teniendo en cuenta el Índice de Selección (IS) y la precocidad de las accesiones; es decir, floración masculina en un rango entre 85 a 95 días, y de esta manera evitar la confusión del escape y la tolerancia a la sequía del material frente al estrés aplicado como se explicó en la sección anterior. La columna IS+Nativos fue expresada en orden ascendente solamente para las poblaciones nativas de 1 hasta 56 tomando como base la columna IS-número general.

La accesión OPV-DTPW/LPSDR (ID 25) del Subprograma de Fisiología fue la mejor en términos de respuesta adaptativa en IS-número e IS+Antesis. Cuando sólo se tiene en cuenta el IS-número, 15 de las 16 accesiones (exceptuando por Tuxpeño Crema 1, ID 47)

provenientes del Banco de Germoplasma y del Subprograma Fisiología, se encuentran entre las 20 mejores. Para las accesiones seleccionadas en IS+Antesis quedan fuera 3 *pooles* del Banco de germoplasma de las 20 mejores seleccionadas en IS-número por tener floración masculina menor a 85 días.

Mirando la clasificación IS+Antesis la mejor población nativa fue PUER 5 (ID 36) proveniente de Puerto Rico y clasificada como una población compuesta de una mezcla de Cristalinos Cubanos. Es importante enfatizar otra vez que la precocidad es realmente un factor muy importante en los ensayos de sequía y se debe tener en cuenta en el análisis. Por ejemplo, la población ARZM 18009 de la raza Cristalino Colorado proveniente de Argentina y GUAD 15 proveniente del archipiélago de Guadalupe y de la raza Cristalino Costeño Tropical están listadas como las accesiones que presentan la mejor respuesta adaptativa bajo condiciones de estrés ya que obtienen los ITS más altos de todas las accesiones nativas 0,28 y 0,21 respectivamente. Estos índices altos resultan de la combinación de un muy bajo rendimiento en condiciones normales ( $0,85$  y  $1,50 \text{ tm ha}^{-1}$ ) y un rendimiento alto en condiciones de estrés comparado con otras accesiones nativas ( $0,24$  y  $0,31 \text{ tm ha}^{-1}$ ). Sin embargo, estos resultados bajo condiciones de sequía reflejan una evasión a la sequía ya que estas accesiones cuentan con una floración masculina de 78 y 82 días respectivamente. De manera general la mayoría de las accesiones precoces tienen un ITS alto (por ejemplo: NVOL 9, con un ITS de 0,10 y una floración de 76 días) y por lo tanto tienen que ser eliminadas de la selección como se hizo.

#### **4.3.3 Caracterización y desempeño agronómico de las accesiones bajo estrés y condiciones normales**

Considerando el rendimiento en condiciones de estrés únicamente, las accesiones con mayor rendimiento pertenecen al grupo del programa de pre-mejoramiento del Banco y del Subprograma de Fisiología, Pool 21 x Pool 22 (Best testcrosses) y OPV-DTPW/LPSDR respectivamente. La distribución fenotípica señala una amplia variabilidad del rendimiento basados en la desviación estándar de este carácter que a su vez es una estimación indirecta de la variación dentro de cada accesión o población. En este sentido, el rango más amplio

es para las poblaciones nativas, seguida por el material del Banco de Germoplasma y con menor rango de distribución el material mejorado del Subprograma de Fisiología.

Comparando la variación estándar de los rendimientos dentro de cada accesión (Tabla 4.5) con la desviación estándar dentro del grupo al que pertenecen las accesiones (Tabla 4.4) bajo condiciones de estrés, en muchos casos y para los tres tipos de material aquí estudiados es más alta la variación dentro de una misma accesión que dentro su grupo correspondiente.

En condiciones de estrés, el material de Fisiología obtuvo el doble de rendimiento medio que los *pooles* del Banco de Germoplasma, y aproximadamente 9 veces el rendimiento medio de las poblaciones nativas. Sin embargo, se tiene que evaluar si esta diferencia es el resultado del mecanismo de adaptación a sequía o del vigor de las plantas *per se*, ya que se están comparando material mejorado con material nativo. Para contestar a esta pregunta se calculó el Índice de Tolerancia a Sequía (ITS) (Tabla 4.5).

Considerando los resultados del ITS, que no rebasa el valor de 0,13 (OPV-DTPYDR) para los genotipos floreado entre 85 y 95 días, se puede concluir que el mejor rendimiento de las accesiones de fisiología o los *pooles* del Banco de Germoplasma provienen más que todo de su vigor ya que también tienen un rendimiento mucho superior en condiciones normales. Sin duda la accesión que presenta un vigor muy alto es OPV-DTPW/LPSDR, ya que tiene casi el mayor rendimiento a través de los diferentes regímenes de irrigación y florea a los 88 días.

En cuanto a las accesiones que corresponden a las poblaciones nativas, y considerando el nivel de estrés severo, si se eliminan los extremos de la distribución para floración los ITS también están muy bajos, ya que los rendimientos en condiciones de sequía son mínimos. Sin embargo, el nivel de estrés fue apropiado considerando el propósito del experimento que era distinguir los extremos en términos de tolerancia a sequía.



**Tabla 4.5** Accesiones identificadas con número ID, rendimientos para condiciones de estrés (CE) y riego normal (CN) con su desviación estándar correspondiente, Índice de tolerancia a la sequía (ITS, definido como: CE/CN) normales). Distribución fenotípica de las 70 accesiones en el estudio bajo condiciones estrés basada en el Índice de Selección (IS) en orden ascendente (1-70) (IS-número), y la distribución fenotípica basada en IS y floración masculina (antesis) denotada como IS + ASI, en orden ascendente para las 20 mejores y en orden ascendente las 20 peores. Distribución fenotípica solo para las poblaciones nativas se denota como IS-Nativo, basado en los datos que se usaron para calcular el IS-número. ASI: Intervalo entre las floraciones masculina y femenina; MPP: Mazorcas por planta; EH: Enrollamiento de hojas (calificación 1 indica que no hay enrollamiento, 5 indica severo enrollamiento); SPAD<sub>DDS</sub>: Contenido de clorofila en las hojas medido con un equipo portátil, DDS indica días después de la siembra, Sene<sub>DDS</sub>: senescencia visual, 1 = 10% del área foliar muerta y 10= 100% área foliar muerta. \*\*OPVs del Subprograma de Fisiología, \**pooles* del Banco de Germoplasma y complejo Tuxpeño Crema, accesiones restantes representan las poblaciones nativas.

ID	Nombre de Accesoión	CE Rendimiento $\text{tm ha}^{-1}$	CN Rendimiento $\text{tm ha}^{-1}$	ITS	ASI (días)	Número MPP	SPAD198DDS	SPAD2118DDS	EH (1-5)	Sen1102DDS	Sen2120DDS	IS	Antesis	IS-Número	IS+ Antesis	IS Nativo
25**	OPV-DTPW/LPSDR	0,87 ± 0,29	8,62 ± 1,11	0,10	-1,00	0,53	38,70	22,13	3,00	2,67	3,67	0,83	88	1	1	
19**	OPV-DTPWDR	0,75 ± 0,19	6,61 ± 1,99	0,11	1,67	0,51	34,51	11,00	3,00	3,67	4,67	0,72	86	2	2	
8*	Pool 21 x Pool 22 (Best testcrosses)	0,90 ± 0,77	4,90 ± 0,63	0,18	1,33	0,48	36,05	9,87	2,33	3,67	7,67	0,67	80	3		
32**	OPV-LPSDR	0,47 ± 0,21	6,75 ± 1,63	0,07	4,67	0,42	35,67	16,57	2,67	2,67	3,00	0,57	92	4	3	
16**	OPV-DTPYDR	0,64 ± 0,21	5,09 ± 0,98	0,13	1,33	0,39	31,57	10,03	4,00	3,00	5,00	0,54	85	5	4	
21*	Pool 20 x CML-495 (Best testcrosses)	0,36 ± 0,16	9,07 ± 0,72	0,04	7,33	0,28	39,36	15,30	1,67	3,33	5,33	0,34	86	6	5	
6*	Pool 21 x Pool 22 (♂)	0,46 ± 0,58	5,90 ± 1,15	0,08	5,50	0,28	24,60	5,60	3,33	4,33	6,67	0,30	80	7		
13*	Pool 22 x Pool 21 (Best testcrosses)	0,38 ± 0,22	6,12 ± 0,78	0,06	6,00	0,30	26,87	4,83	3,33	4,33	6,67	0,29	81	8		
58*	Pool 23 x Pool 24 (Best testcrosses)	0,26 ± 0,13	5,76 ± 0,93	0,05	2,00	0,16	34,15	12,97	3,00	2,33	4,67	0,22	90	9	6	
38*	Pool 26 x CML-451 (Best testcrosses)	0,25 ± 0,23	8,10 ± 0,71	0,03	6,00	0,15	33,98	14,60	4,67	2,67	3,67	0,22	90	10	7	
12	GUAD 15	0,31 ± 0,22	1,50 ± 1,41	0,21	-0,33	0,29	22,93	6,90	2,67	5,33	8,67	0,22	82	11		1
44*	Pool 24 x Pool 23 (Best testcrosses)	0,22 ± 0,11	8,98 ± 1,06	0,02	2,67	0,15	36,28	12,73	3,67	2,67	4,33	0,20	90	12	8	
36	PUER 5	0,13 ± 0,04	2,57 ± 0,79	0,05	1,00	0,14	27,77	9,13	2,67	2,67	4,67	0,14	89	13	9	2

**Continuación Tabla 4.5**

ID	Nombre de Acésión	CE Rendimiento tm ha <sup>-1</sup>	CN Rendimiento tm ha <sup>-1</sup>	ITS	ASI (días)	Número MPP	SPAD198DDS	SPAD2118DDS	EH (1-5)	Sen1102DDS	Sen2120DDS	IS	Antesis	IS- Número	IS+ Antesis	IS Nativo
23*	Pool 19 x CML-494 (Best testcrosses)	0,14 ± 0,17	6,93 ± 1,25	0,02	3,67	0,09	39,37	18,27	2,33	2,00	4,33	0,13	88	14	10	
66	PAZM 6053	0,17 ± 0,15	5,03 ± 1,60	0,03	4,00	0,14	37,15	11,97	4,33	4,00	6,00	0,13	90	15	11	3
39	PUER GP1	0,20 ± 0,13	4,77 ± 2,25	0,04	3,67	0,17	32,75	7,13	2,33	3,67	7,00	0,12	88	16	12	4
40*	Pool 23 x Pool 24 (♂)	0,07 ± 0,02	6,64 ± 2,78	0,01	-2,00	0,09	38,13	13,27	4,67	3,33	4,67	0,11	92	17	13	
9	NAYA 171	0,22 ± 0,14	2,19 ± 1,60	0,10	4,50	0,20	20,67	4,43	4,00	5,33	9,00	0,11	82	18		5
15	TAMA 12	0,18 ± 0,19	4,29 ± 0,60	0,04	2,50	0,19	28,62	9,47	3,67	4,00	8,00	0,10	84	19		6
50	GUAD GP2	0,08 ± 0,08	6,50 ± 1,56	0,01	5,50	0,13	32,78	9,03	3,67	3,00	5,00	0,10	91	20	14	7
64	GUAD 9	0,01 ± 0,17	3,90 ± 1,26	0,00	5,33	0,11	30,40	10,67	3,33	2,67	4,67	0,09	92	21	15	8
55	HAIT GP3	0,09 ± 0,09	6,20 ± 0,95	0,01	3,50	0,10	33,03	12,17	2,00	3,33	5,67	0,07	98	22		9
5	NVOL 9	0,20 ± 0,13	2,03 ± 0,52	0,10	3,00	0,27	10,82	4,30	4,33	6,67	9,67	0,07	76	23		10
1	ARZM 18 009	0,24 ± 0,27	0,85 ± 0,40	0,28	6,33	0,21	20,08	5,00	5,00	6,00	9,33	0,07	78	24		11
47*	TUXPEÑO CREMA 1	0,04 ± 0,06	4,92 ± 1,35	0,01	4,00	0,11	31,03	6,53	2,67	3,00	5,00	0,06	92	25	16	
57	VENE 783	0,00 ± 0,01	2,54 ± 0,97	0,00	0,67	0,09	31,42	12,03	3,00	4,33	4,67	0,06	97	26		12
60	JALI 285	0,00 ± 0,01	2,39 ± 0,69	0,00	4,00	0,04	34,50	4,80	3,00	3,67	5,67	0,05	94	27	17	13
30	JALI 281	0,06 ± 0,06	3,24 ± 0,22	0,02	5,00	0,09	34,56	9,37	3,00	3,00	5,67	0,04	87	28	18	14
70	JALI 280	0,11 ± 0,10	3,10 ± 1,01	0,04	4,33	0,08	35,10	14,90	2,67	3,33	6,33	0,04	90	29	19	15
20	NVOL 22	0,12 ± 0,04	2,91 ± 1,18	0,04	4,67	0,11	28,77	7,57	3,67	4,33	6,33	0,04	86	30	20	16
28	TAMA 36	0,18 ± 0,11	3,36 ± 1,95	0,05	4,50	0,13	27,42	8,13	3,67	5,67	8,00	0,04	88	31		17
26	PAZM 14107	0,14 ± 0,02	4,49 ± 3,45	0,03	5,00	0,06	33,98	20,13	3,67	3,67	4,33	0,03	97	32		18
14	ARZM 06 072	0,11 ± 0,10	1,50 ± 1,40	0,07	3,50	0,15	29,67	5,03	3,67	5,33	8,67	0,03	82	33		19
52	PAZM 10043	0,04 ± 0,05	4,42 ± 0,58	0,01	3,00	0,05	34,17	14,13	2,67	3,67	5,00	0,02	93	34		20
11	SONO 34	0,14 ± 0,13	1,20 ± 1,07	0,12	-0,67	0,19	18,72	5,20	4,33	6,00	9,33	0,02	72	35		21
22	TAMA 37	0,11 ± 0,12	2,77 ± 0,61	0,04	7,00	0,15	25,40	5,77	3,33	5,67	7,33	0,02	86	36		22
51	PAZM 2019	0,01 ± 0,02	3,01 ± 1,66	0,00	0,00	0,01	37,13	13,20	4,00	3,00	4,33	0,01	93	37		23
27	ANTI GP2	0,09 ± 0,10	4,26 ± 0,40	0,02	4,00	0,10	30,48	13,93	3,67	4,67	7,33	0,01	90	38		24
33	BRAZ 1327	0,07 ± 0,09	2,36 ± 1,30	0,03	6,00	0,05	34,58	11,77	4,00	3,00	5,67	0,00	90	39		25
17	JALI 397	0,14 ± 0,03	3,85 ± 0,59	0,04	7,67	0,11	26,35	4,43	3,00	5,00	7,33	0,00	81	40		26
65	NAYA 157	0,01 ± 0,02	2,26 ± 0,61	0,00	0,00	0,03	32,70	10,67	4,33	3,33	5,00	-0,01	90	41		27

**Continuación Tabla 4.5**

ID	Nombre de Acesión	CE Rendimiento tm ha <sup>-1</sup>	CN Rendimiento tm ha <sup>-1</sup>	ITS	ASI (días)	Número MPP	SPAD198DDS	SPAD2118DDS	EH (1-5)	Sen1102DDS	Sen2120DDS	IS	Antesis	IS- Número	IS+ Antesis	IS Nativo
48	VENE 484	0,00 ± 0,00	2,75 ± 0,35	0,00	3,50	0,05	34,60	7,37	4,67	4,33	5,67	-0,01	92	42		28
7	SALV 66	0,12 ± 0,10	1,28 ± 0,88	0,09	3,33	0,17	21,40	5,17	3,33	6,33	9,67	-0,01	76	43		29
18	ARZM 06 056	0,10 ± 0,10	1,42 ± 0,35	0,07	2,50	0,14	20,13	3,77	4,00	5,67	8,00	-0,01	83	44		30
31	GUAD 6	0,04 ± 0,02	4,00 ± 0,90	0,01	5,33	0,08	27,89	7,70	2,33	3,33	6,33	-0,02	88	45		31
45	BRAZ 2314	0,06 ± 0,05	6,30 ± 1,29	0,01	7,00	0,04	35,83	9,53	2,00	4,00	6,00	-0,02	91	46	20	32
59	JALI 286	0,05 ± 0,10	4,12 ± 0,37	0,01	15,00	0,08	30,02	12,80	2,67	3,33	5,33	-0,02	91	47	19	33
49	CUBA 82	0,00 ± 0,00	4,29 ± 1,76	0,00	4,50	0,03	34,35	12,13	3,67	4,00	4,67	-0,02	95	48	18	34
62	GUAD 3	0,02 ± 0,03	3,90 ± 0,97	0,01	8,00	0,08	25,45	10,43	2,67	4,00	5,33	-0,02	89	49	17	35
61	PUER 19	0,01 ± 0,01	4,17 ± 1,62	0,00	8,67	0,04	29,85	12,27	3,00	2,67	4,33	-0,02	93	50	16	36
54	ARZM 06 079	0,02 ± 0,03	3,88 ± 0,62	0,01	0,00	0,07	27,73	4,43	3,00	4,00	6,67	-0,03	90	51	15	37
68	BRAZ 1192	0,01 ± 0,00	2,09 ± 0,86	0,00	-3,00	0,04	32,57	9,77	3,00	3,67	7,00	-0,03	91	52	14	38
56	NAYA 174	0,03 ± 0,05	2,65 ± 0,95	0,01	-3,00	0,03	28,21	7,77	2,67	4,33	6,33	-0,03	92	53	13	39
67	TAMA 125	0,01 ± 0,01	2,71 ± 1,92	0,00	1,50	0,05	29,63	7,07	4,33	4,67	6,67	-0,04	94	54	12	40
46	SINA 30	0,02 ± 0,02	3,96 ± 0,35	0,01	4,00	0,03	28,31	3,73	4,67	4,33	5,33	-0,04	93	55	11	41
2	ARZM 18 004	0,15 ± 0,07	0,89 ± 0,36	0,17	4,00	0,16	14,87	4,70	4,67	6,33	9,67	-0,04	77	56		42
43	PUER GP6	0,03 ± 0,03	0,39 ± 0,36	0,08	6,50	0,05	31,35	5,00	3,00	3,33	6,33	-0,05	91	57	10	43
69	CRIC 81	0,02 ± 0,03	2,78 ± 1,89	0,01	5,50	0,07	30,12	4,70	3,67	4,00	7,00	-0,05	90	58	9	44
34	SINA 74	0,03 ± 0,06	3,55 ± 1,41	0,01	2,50	0,01	34,72	11,27	4,67	4,67	6,33	-0,05	92	59	8	45
63	SONO 72	0,06 ± 0,06	2,69 ± 1,29	0,02	6,00	0,03	30,78	8,83	3,67	4,00	6,67	-0,06	92	60	7	46
53	BRAZ 2394	0,01 ± 0,01	4,68 ± 1,66	0,00	10,00	0,07	35,75	7,67	4,00	4,33	7,33	-0,06	93	61	6	47
3	ARZM 1124	0,05 ± 0,03	1,59 ± 0,62	0,03	2,67	0,14	21,85	5,23	3,67	5,00	10,00	-0,06	78	62		48
24	GUAD GP1	0,02 ± 0,02	5,33 ± 0,24	0,00	8,00	0,04	35,18	7,00	3,67	3,67	6,67	-0,06	87	63	5	49
29	TAMA GP4	0,01 ± 0,01	1,48 ± 0,20	0,01	-3,00	0,04	26,82	3,90	4,33	5,00	8,67	-0,09	88	66	3	52
42	NAYA 178	0,00 ± 0,01	2,77 ± 1,27	0,00	9,00	0,03	26,88	10,57	3,33	4,00	6,33	-0,10	87	67	2	53
10	URUG 630	0,04 ± 0,05	0,86 ± 1,02	0,05	5,00	0,12	18,40	4,00	5,00	6,00	9,67	-0,10	80	68		54
35	BRAZ 2306	0,00 ± 0,00	6,79 ± 0,67	0,00	4,00	0,00	24,36	5,93	2,33	3,33	6,33	-0,10	93	69	1	55
4	SINA 29	0,04 ± 0,03	1,51 ± 0,35	0,03	6,67	0,08	16,45	3,47	4,67	7,33	9,33	-0,16	78	70		56

De hecho, hay que recordar que este germoplasma no está mejorado y que cuenta con más mecanismos de sobrevivencia como producir al mínimo algunas semillas que de tolerancia según el concepto de mejoramiento, a saber: seguir produciendo la cantidad máxima de semilla aunque la planta detecta condiciones desfavorables que pueden comprometer su descendencia.

Por esta misma razón quizá no es tan importante fijarse en el rendimiento *per se* pero si en el Índice de Selección y los valores de los parámetros secundarios para sequía, por esto fueron seleccionadas las 20 mejores y peores en función del Índice de Selección que incluye los parámetros de sequía como el ASI, el contenido de clorofila o el número de mazorcas por planta. En este sentido las accesiones BRAZ 1192, NAYA 174, PAZM 2019 y VENE 783 se consideran interesantes ya que tienen un ASI de -3, -3, 0 y 0,67 días respectivamente, y con una floración algo tardía de 91, 92, 93 y 97 días. Aunque estas accesiones casi no tuvieron rendimiento en condiciones de estrés, deben tener alelos favorables para la sincronía entre floración femenina y masculina bajo sequía y por lo tanto estos alelos que pueden ser identificados por asociación y podrían tener un impacto significativo sobre rendimiento una vez introducidos en fondo genético elite. Este mismo hecho aplica para senescencia, por ejemplo la accesión PUER 19 puede ser relevante, y también tiene una floración tardía de 93 días. También accesiones interesantes que podrían presentar alelos favorables para senescencia y que están entre las “mejores” nativas son: PAZM 6053 Y PUER GP1 con floraciones de 88 y 90 días respectivamente.

Sin embargo, hay también varias accesiones nativas interesantes desde un punto de vista genético para el rendimiento. Accesiones como TAMA36 (88 días de floración) que presenta un rendimiento de 0,18 y 3,36  $\text{tm ha}^{-1}$  para estrés y condiciones normales respectivamente, con un ITS de 0,05, parece una accesión muy interesante para buscar genes de tolerancia a sequía en material nativo. Las accesiones PUER GP1 y TAMA 12 son dos accesiones de poblaciones nativas “atractivas”, ya que producen 0,2 y 0,18  $\text{tm ha}^{-1}$  respectivamente en condiciones de estrés, 4,8 y 4,3  $\text{tm ha}^{-1}$  respectivamente en condiciones normales, PUER GP1 tiene una floración de 88 días, y TAMA 12 es un poco precoz con 84 días, sin embargo estas dos accesiones son excepcionales en condición de sequía ya que la siguiente accesión en rendimiento produce 0,12  $\text{tm ha}^{-1}$ . Mirando sus rendimientos en los

dos ambientes, estrés y condiciones normales, sin ninguna duda estas dos accesiones tienen alelos favorables para sequía y también alelos de vigor. Bajo condiciones de sequía hay varias accesiones provenientes de Tamaulipas – México con perfiles interesantes (TAM A12, TAMA 36 y TAMA 37) que “coincidentalmente” son de la raza Tuxpeño, raza de gran importancia desde el punto de vista de su influencia en otras razas de maíz agrícola productivas. Los rendimientos bajo riego normal para poblaciones nativas como por ejemplo las accesiones BRAZ 2306, GUAD GP2, BRAZ 2314 y HAIT GP3 están por encima de las 6 t m ha<sup>-1</sup> (Tabla 4.5) pero casi no producen bajo condiciones de sequía, este hecho, probablemente es el resultado de una floración tardía (93 días para BRAZ 2306 y hasta 98 para HAIT GP3). Sin embargo estas accesiones tienen un potencial alto de identificar nuevos alelos para vigor.

#### **4.3.4 Heredabilidad**

La heredabilidad para el rendimiento en todos los casos (Tabla 4.6), fue menor en condiciones de estrés (en un rango de 0,58 - 0,79) que en riego normal (0,82 - 0,91). Para el carácter de número de mazorcas por planta en condiciones de estrés como para riego normal la heredabilidad tuvo valores en un rango intermedio - alto (0,63 - 0,86).

Todos los valores de heredabilidad para ASI fueron menores que para rendimiento, mazorcas por planta y senescencia. Las mayores diferencias en heredabilidad, bajo los dos regímenes evaluados, fueron para los contenidos de clorofila (SPAD<sub>98DDS</sub> y SPAD<sub>118DD</sub>) que se registraron para las accesiones nativas en condiciones de riego normal. En el carácter enrollamiento de hoja evaluado para condiciones de estrés, se registró el doble del valor de heredabilidad para el material mejorado que para el panel inicial y las poblaciones nativas. El caso contrario es para las medidas de senescencia (Sen<sub>102DDS</sub> y Sen<sub>120DDS</sub>) cuyos valores son menores si se compara el material mejorado comparado con panel inicial y accesiones nativas. Estos valores altos de heredabilidad reflejan una alta calidad en la caracterización fenotípica, un componente esencial para cualquier estudio genético. De manera general, una heredabilidad un poco más alta en condición de sequía en comparación de la heredabilidad en un riego normal es lo esperado, para parámetros secundarios de sequía ya que tienen una mejor segregación y varianza genética en condiciones de estrés.

Sin embargo, es lo contrario para el rendimiento, con una heredabilidad mayor en condición de riego normal y menor en condiciones de estrés, este hecho es debido al impacto del ambiente y más importante todavía en condiciones de estrés. Por esto, es primordial considerar parámetros secundarios en la selección (alta correlación con el rendimiento, fácil de medir y con heredabilidad alta aunque en condiciones de estrés).

**Tabla 4.6** Heredabilidad ( $H^2$ ) calculada para rendimiento de grano, mazorcas por planta, ASI, contenido de clorofila (SPAD<sub>98DDS</sub> y SPAD<sub>118DDS</sub>), enrollamiento de hoja\* (EH<sub>104DDS</sub>) y senescencia\* (Sen<sub>102DDS</sub> y Sen<sub>120DDS</sub>) en condiciones de estrés y riego normal para el panel inicial, las accesiones nativas del panel inicial y material mejorado incluido en el panel inicial.

Carácter	Heredabilidad ( $H^2$ )					
	Condiciones de estrés			Regadío normal		
	panel inicial	nativas	material mejorado	panel inicial	nativas	material mejorado
Rendimiento	0,79	0,58	0,60	0,91	0,84	0,82
Mazorcas por planta	0,86	0,68	0,80	0,74	0,63	0,80
ASI	0,42	0,40	0,52	0,52	0,26	0,60
SPAD <sub>98DDS</sub>	0,72	0,69	0,52	0,61	0,37	0,59
SPAD <sub>118DDS</sub>	0,67	0,58	0,63	0,75	0,46	0,56
EH <sub>104DDS</sub>	0,37	0,32	0,64	*	*	*
Sen <sub>102DDS</sub>	0,82	0,83	0,33	*	*	*
Sen <sub>120DDS</sub>	0,86	0,84	0,71	*	*	*

#### 4.3.5 Diversidad genética

Los 30 SSRs usados para medir la diversidad genética fueron polimórficos a través de las 70 poblaciones y un total de 216 alelos fueron detectados con una media de 7,2 alelos por locus (Tabla 4.7), donde el 56% son alelos raros (frecuencias menores del 5%). Dentro de estos alelos raros se identificaron 59 alelos únicos, es decir, identificados en una sola clase de germoplasma: *poles* del Banco de Germoplasma, o material mejorado del Subprograma de Fisiología, o poblaciones nativas; de este total, el 95% de alelos únicos se encuentran en

las poblaciones nativas. De la misma manera, si se compara el panel inicial con el panel seleccionado, se detectaron 13 alelos raros en el panel principal que no están presentes en el panel seleccionado (Tabla 4.7). En cuestión de diversidad genética, la frecuencia de heterocigotos, y tomando en cuenta la frecuencia alélica y número de alelos, hay una tendencia que indica que la diversidad es menor para las accesiones de fisiología en comparación de los *pooles* del Banco y las accesiones nativas. Esta reducción de diversidad es el producto de la selección, pero el hecho que la diferencia es apenas significativa parece indicar que la selección afecta la frecuencia alélica pero que no eliminó muchos alelos desfavorables para el desempeño agronómico.

**Tabla 4.7** Alelos encontrados y diversidad genética calculados como el promedio de la diversidad para cada grupo de accesiones o poblaciones de acuerdo a la clase de material, así como también para el panel seleccionado. \*\* Número de alelos únicos comparados entre grupos de accesiones, \*Número de alelos únicos comparados entre el panel seleccionado con el total; na, no aplica.

Material	Número de accesiones o poblaciones (individuos)	Alelos encontrados	Número de alelos únicos	Media número de alelos por locus	Diversidad genética $\pm$ DS
Banco de Germoplasma	10 (150)	155	3**	5,17	0,51 $\pm$ 0,03
Programa Fisiología	4 (60)	113	1**	3,77	0,45 $\pm$ 0,03
Poblaciones nativas	56- (840)	212	55**	7,07	0,52 $\pm$ 0,04
Panel seleccionado	40 (600)	203	na	6,77	0,51 $\pm$ 0,04
<b>Panel Inicial - Total</b>	70 (1050)	216	13*	7,20	0,52 $\pm$ 0,04

Se estimaron los índices de fijación  $F_{ST}$  por parejas para los grupos de accesiones clasificadas de acuerdo a la clase de material como indicador de diferenciación genética (Tabla 4.8), donde  $F_{ST}$  será 0 cuando las accesiones/poblaciones son idénticas en las frecuencias alélicas, y 1 cuando las frecuencias alélicas de las accesiones/poblaciones están

fijas para diferentes alelos (Mohammadi y Prasanna, 2003). De acuerdo con los intervalos para la interpretación de los valores de  $F_{ST}$  sugeridos por Wright (1965) se encontraron tres niveles de diferenciación genética entre los grupos de germoplasma.

Entre los *pools* provenientes del Banco de Germoplasma y las poblaciones nativas se encontró muy poca diferenciación genética (0,03) en un rango de 0,0 – 0,05, entre los *pools* del Banco de Germoplasma y el Subprograma de Fisiología se encontró moderada diferenciación (0,09) entre un rango de 0,05 – 0,15, y entre el Subprograma de Fisiología y las poblaciones nativas se encontró un valor de 0,32 ( $> 0,25$ ), lo que representa una diferenciación genética importante entre estos dos grupos, probablemente como resultado de una presión de selección alta para mejorar la accesiones de fisiología. Una presión de selección fuerte cambia las frecuencias alélicas en los *loci* involucrados en el desempeño agronómico de las plantas, y aunque no afecta de manera drástica la diversidad genética intrínseca de cada tipo de material por las razones mencionadas arriba, la reducción de *loci* heterocigotos en material mejorado como resultado del incremento de frecuencia alélica de alelos favorables, impacta mucho el índice de Fijación entre tipo de material.

**Tabla 4.8** Valores de Índices de Fijación  $F_{ST}$  calculados por parejas entre los grupos de poblaciones clasificados de acuerdo a clase de material.

	Banco de Germoplasma	Subprograma de Fisiología
Subprograma de Fisiología	0,09	
Poblaciones Nativas	0,03	0,32

#### 4.3.6 Cambios de frecuencia alélica

Se calcularon los valores de los cambios de frecuencia alélica entre los diferentes materiales (Anexo 7.9), *pools* del Banco vs accesiones nativas, fisiología vs poblaciones nativas y *pools* contra fisiología, tomando el valor de 10% como valor de significancia, basado en el valor medio del intervalo de moderada diferenciación genética sugerida por Wright (1965). Tomando en cuenta los tres tipos de comparaciones, de los 216 alelos totales, el 71% (154) no representaron cambios de frecuencia mayores al 10%. Sin embargo, se encontraron 101 cambios de frecuencia mayores al 10% con 27 de 30



marcadores, y si se toma en cuenta que 12 de los 30 SSRs utilizados en este estudio se derivan de genes previamente identificados, se encontraron 42 cambios de frecuencia alélica mayores al 10% en estos *loci* de los 101 cambios totales. A su vez, de estos 42 cambios de frecuencia, 15 de éstos están relacionados con genes de sequía o altas temperaturas (Tabla 4.9, Anexo 7.3).

**Tabla 4.9** Cambios de frecuencia alélica entre las diferentes clases de material tomando en cuenta todos los cambios detectados a través de los 30 marcadores y los cambios encontrados a través de los marcadores derivados de genes previamente identificados, se expone también el porcentaje máximo de cambio de frecuencia a través de todos los cambios detectados.

	<b><i>Pooles del Banco-Poblaciones nativas</i></b>	<b><i>Fisiología – Pooles del Banco</i></b>	<b><i>Fisiología-Poblaciones nativas</i></b>
Número de marcadores con los que se identificaron cambios de frecuencia alélica	12	25	18
Número total de cambios en frecuencias alélicas	16 (25,29%)	48 (46,67%)	37 (40,75%)
Número de cambios en frecuencias alélicas en <i>loci</i> relacionados con gen y máximo % de cambio en frecuencia alélica	8 (25,29%)	19 (34,33%)	15 (34,01%)
Número de cambios en frecuencias alélicas en <i>loci</i> relacionados con genes de sequía y altas temperaturas, máximo % de cambio en frecuencia alélica	2 (15,94%)	8 (34,33%)	5 (34,01%)

De los cambios totales (101), se registraron 16 cambios de frecuencia entre los *pooles* del Banco y las poblaciones nativas en 12 de 30 marcadores, siendo el valor más alto de 25,29%; 48 cambios de frecuencia detectados con 25 marcadores para *pooles* del Banco-Fisiología, encontrándose el valor más alto de cambio de frecuencia de 46,67%, esto implica que la selección para mejoramiento en condiciones normales (*pooles*) y sequía (fisiología) se hizo a través de diferentes alelos. Finalmente, se detectaron con 18 de 30 marcadores, 37 cambios de frecuencia alélica para Fisiología-poblaciones nativas con un valor máximo de 40,75% (Tabla 4.9).

Para cambios de frecuencia alélica a través de los marcadores derivados de genes previamente identificados, el número mayor de cambios fue entre materiales de Fisiología y los *pooles* del Banco de Germoplasma (19), con un valor máximo de 34,33%; y el menor número de cambios (8) entre *pooles* de Banco y poblaciones nativas, con un valor máximo de 25,29%. Para genes de tolerancia al a sequía y/o altas temperaturas el mayor número de cambios fue también entre materiales de Fisiología y los *pooles* del Banco de Germoplasma con un máximo de 34,33%, seguido por 5 cambios entre Fisiología y poblaciones nativas con valor máximo de 34,01%, y finalmente 2 cambios para *pooles* del Banco y poblaciones nativas y con un valor máximo de 15,94%.

Los cambios de frecuencia alélica a través de las tres clases de material y de los 30 marcadores dan información muy valiosa relacionada con alelos favorables o desfavorables cuyas frecuencias alélicas se han ido incrementando o disminuyendo a través de los ciclos de selección. La tabla que se presenta a continuación (Tabla 4.10), representa una selección del Anexo 7.9, para mostrar los diferentes casos y posibles interpretaciones de los cambios de frecuencias alélicas a través del germoplasma estudiado y el conjunto de SSRs utilizados en este trabajo de Tesis.

Esta sinopsis presenta y explica con 4 marcadores, alelos encontrados, las frecuencias alélicas, y cambios de frecuencia (expresada como porcentaje) entre los tipos de material. Para los dos primeros marcadores, umc1266 y umc1332, no se encontraron cambios importantes (>10%) en las frecuencias alélicas, sin embargo, representan dos situaciones diferentes. La composición alélica del marcador umc 1266 no es común en material tan diverso ya que uno de los alelos, el 130, se encuentra casi fijo a través de las tres diferentes clases de germoplasma con una frecuencia alélica arriba de 94%. Este alelo tiene que ser extremadamente importante ya que su frecuencia no fue afectada por la evolución o la selección a través del tiempo, quizás podría ser un alelo importante para la germinación o el desarrollo de las plantas. En el segundo caso, el marcador umc1332, las frecuencias alélicas de los diferentes alelos detectados se encuentran constantes a través de las tres clases de germoplasma (ver en particular alelos 117, 139, y 142 y frecuencias alélicas marcadas con rojo).

**Tabla 4.10** Frecuencias alélicas encontradas para *pooles* del Banco de Germoplasma, OPVs del Subprograma de Fisiología y poblaciones nativas, cambios de frecuencia alélica expresada como porcentaje entre estas clases de germoplasma.

SSR	Alelo	Frecuencias alélicas			Cambios de frecuencia (%)		
		<i>pooles</i> del Banco de germoplasma	OPVs de Fisiología	Poblaciones nativas	<i>pooles</i> - nativas	OPVs - <i>pooles</i>	OPVs - nativas
umc1266	119	0,06	0,00	0,01	5,51	-6,30	-0,78
umc1266	<b>130</b>	<b>0,94</b>	<b>1,00</b>	<b>0,98</b>	-4,79	6,30	1,50
umc1266	143	0,00	0,00	0,00	-0,33	0,00	-0,33
umc1266	147	0,00	0,00	0,00	-0,39	0,00	-0,39
umc1332	107	0,03	0,01	0,02	0,85	-1,83	-0,98
umc1332	114	0,00	0,00	0,01	-1,21	0,00	-1,21
umc1332	<b>117</b>	<b>0,35</b>	<b>0,37</b>	<b>0,37</b>	-1,79	1,67	-0,12
umc1332	121	0,02	0,00	0,00	1,97	-2,33	-0,36
umc1332	130	0,02	0,06	0,01	0,73	3,83	4,56
umc1332	132	0,01	0,00	0,01	0,06	-1,33	-1,27
umc1332	135	0,01	0,00	0,02	-1,64	-0,67	-2,30
umc1332	<b>139</b>	<b>0,22</b>	<b>0,22</b>	<b>0,20</b>	2,58	-0,67	1,91
umc1332	<b>142</b>	<b>0,34</b>	<b>0,35</b>	<b>0,35</b>	-1,55	1,33	-0,21
phi063	<b>154</b>	0,04	0,18	0,04	0,91	13,33	14,24
phi063	156	0,00	0,00	0,00	-0,14	0,00	-0,14
phi063	159	0,00	0,00	0,02	-1,56	0,00	-1,56
phi063	<b>161</b>	0,13	0,60	0,19	-5,92	46,67	40,75
phi063	166	0,00	0,00	0,01	-0,88	0,00	-0,88
phi063	<b>170</b>	<b>0,56</b>	<b>0,18</b>	<b>0,49</b>	6,24	-37,78	-31,54
phi063	<b>174</b>	<b>0,13</b>	<b>0,04</b>	<b>0,16</b>	-2,93	-8,89	-11,81
phi063	178	0,04	0,00	0,04	0,57	-4,44	-3,88
phi063	182	0,02	0,00	0,01	0,83	-1,85	-1,02
phi063	186	0,00	0,00	0,02	-2,11	0,00	-2,11
phi063	190	0,00	0,00	0,00	-0,27	0,00	-0,27
phi063	214	0,02	0,00	0,01	0,36	-1,85	-1,50
phi063	218	0,04	0,00	0,00	3,94	-4,07	-0,14
phi063	222	0,00	0,00	0,00	-0,14	0,00	-0,14
phi063	226	0,01	0,00	0,00	1,11	-1,11	0,00
phi076	157	0,00	0,00	0,01	-0,72	-0,33	-1,05
phi076	<b>159</b>	0,12	0,00	0,15	-2,98	-12,33	-15,31
phi076	163	0,01	0,00	0,00	0,67	-0,67	0,00
phi076	165	0,02	0,00	0,14	-12,41	-1,67	-14,07
phi076	168	0,00	0,02	0,00	-0,19	1,67	1,48
phi076	170	0,84	0,63	0,68	15,94	-21,00	-5,06
phi076	<b>176</b>	<b>0,01</b>	<b>0,35</b>	<b>0,01</b>	-0,32	34,33	34,01

Los resultados parecen indicar que la presión de selección no generó ningún cambio de frecuencia, sugiriendo que este *locus* no es importante para rendimiento o desempeño agronómico en condición de riego normal y de sequía. Para los siguientes marcadores, phi063 y phi076 se encontraron cambios frecuencia alélica mayores al 10% (marcados con amarillo), para alelos favorables (incremento de la frecuencia) y desfavorables (disminución de la frecuencia) tomando como referencia la composición alélica en el material del Subprograma de Fisiología. Con el marcador phi063, los alelos 154 y 161 incrementan frecuencias alélicas alrededor del 13% y 44% respectivamente debido a la presión de selección que se ha ejercido sobre este material comparado con los *pooles* del Banco y las poblaciones nativas (marcados con azul). Análogamente, la disminución de frecuencias alélicas marcadas con verde se pueden observar para los alelos 170 y 174.

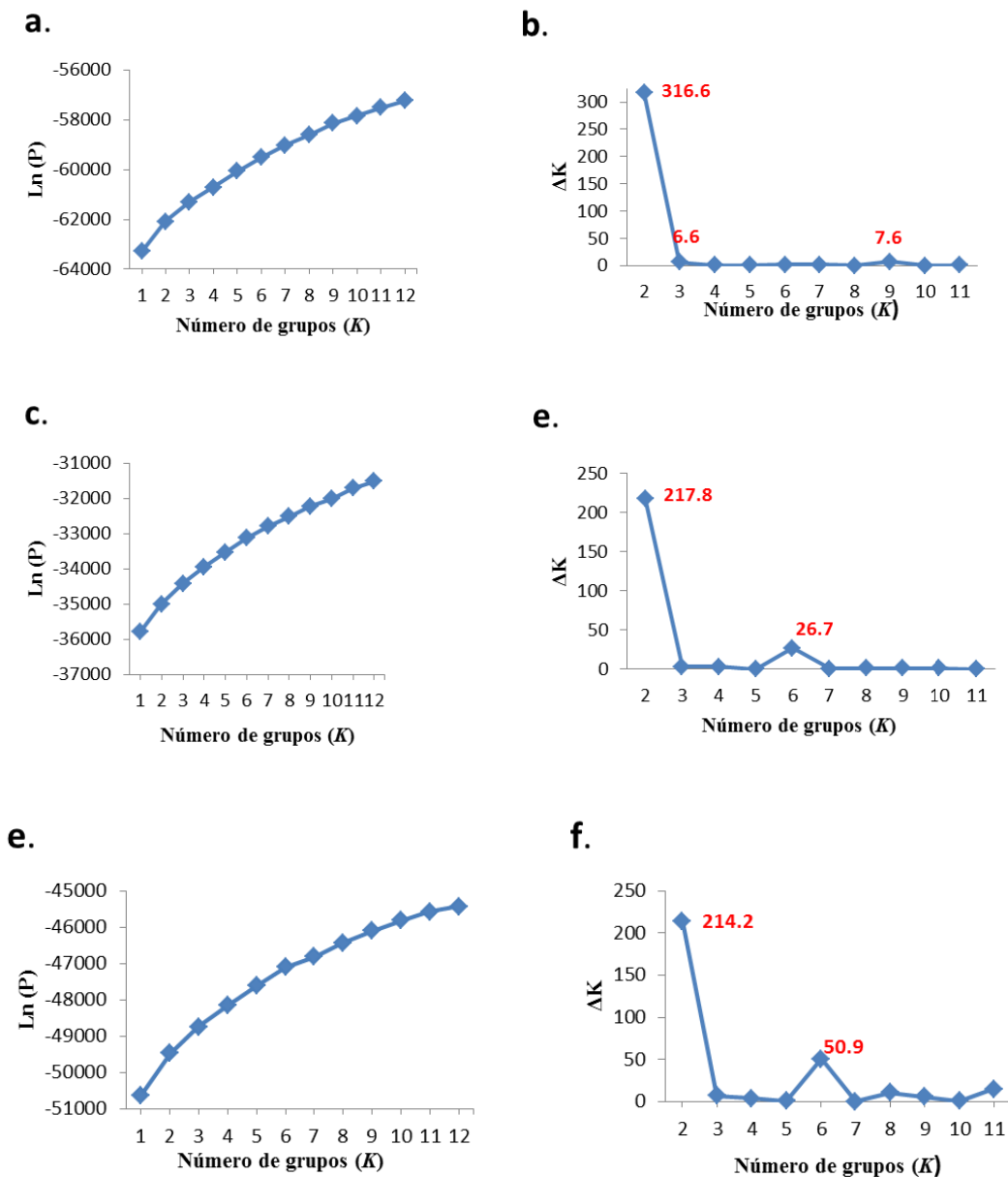
El marcador phi076 identifica un *locus* de particular interés para mejoramiento en sequía/altas temperaturas, pues además de que este *locus* está ligado a un gen (*catalase3*) involucrado en la regulación de la catalasa, una enzima importante para el mantenimiento de procesos celulares normales bajo condiciones de estrés oxidativo debido a altas temperaturas (Anexo 7.3), se puede observar un incremento drástico de la frecuencia de un alelo favorable (176) en el material de Fisiología. Este alelo esta presente en solamente 1 de 100 plantas en los *pooles* del Banco y poblaciones nativas mientras que en las OPVs mejoradas para sequía 35 de 100 tienen este alelo, demostrando su gran valor ligado a tolerancia a sequía.

#### **4.3.7 Análisis de estructura poblacional**

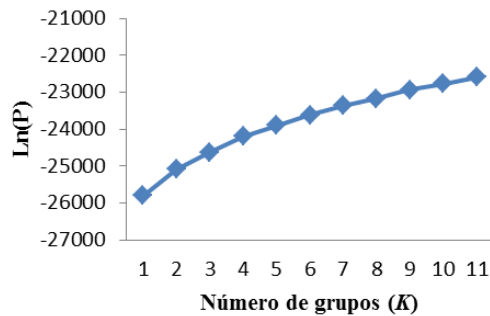
En la evaluación de la estructura poblacional no fue posible determinar directamente el valor de  $K$  en ninguno de los casos (Figuras 4.4a, 4.4c, 4.4e y 4.4g), pero utilizando la metodología reportada por Evanno et al. (2005), descrita en detalle en el capítulo anterior, para el panel inicial de 70 poblaciones los valores más altos de  $\Delta K$ , aunque con poca diferencia, fueron encontrados en  $K=3$  y  $K=9$  (Figura 4.4b); para el panel seleccionado basado en SI+antesis (Figura 4.4d); así como, para las poblaciones nativas el valor más alto para  $\Delta K$  fue para un valor de  $K=6$  (Figura 4.1f) y para accesiones de poblaciones nativas en

panel seleccionado fue de un valor de  $K=4$  (Figura 4.4h). La estructura más marcada fue encontrada en el panel de selección y las poblaciones nativas.

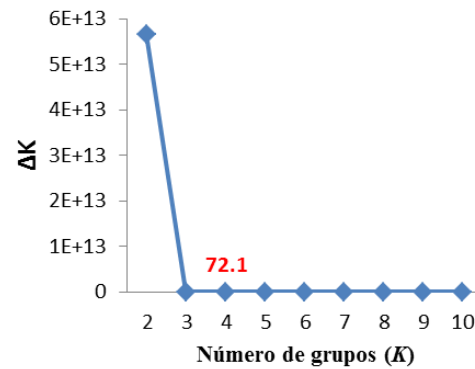
**Figura 4.4** Gráficas del Logaritmo de la probabilidad y  $\Delta K$  para el análisis de estructura poblacional, (a) y (b) para panel inicial de 70 accesiones; (c) y (d) para panel seleccionado de 40 accesiones, (e) y (f) para solo poblaciones nativas, y (g) y (h) para nativas en el panel seleccionado. Se graficaron los valores promedios de  $\ln[\text{prob}(\text{dato})]$  de 3 corridas independientes para cada valor de  $K$ , de la misma manera estos valores fueron utilizados para el cálculo de los valores  $\Delta K$ . Con color rojo están marcados los valores más altos encontrados para  $\Delta K$ .



**g.**



**h.**



En el panel inicial para los valores de  $K$  el 87,14% y 78,57% de las accesiones pertenecieron a uno de los 3 y 9 grupos, respectivamente, en cada análisis independiente de estructura, con promedios altos de ancestro de 78,56% y 79,78%, demostrando la robustez del agrupamiento. En el panel inicial, cuando se analizan tres grupos principales, se identifican 27 accesiones mixtas para el grupo 1, siendo la gran mayoría accesiones mexicanas, una caribeña del Archipiélago de Guadalupe, dos poblaciones de maíces dentados de Brasil y una población venezolana de la raza Puya. En el grupo 2 está formado por 19 accesiones, el común son las poblaciones caribeñas, pero también se encuentran poblaciones de Paraguay, Brasil, Costa Rica y México. El grupo 3 está formado por 13 accesiones, la generalidad en este caso está dada por materiales Sudamericanos de Argentina y Uruguay con una accesión mexicana y otra caribeña. El material de mejoramiento del Banco de germoplasma se encuentra presente en los tres grupos (Tabla 4.11), mientras que las 4 accesiones del Subprograma del Fisiología se encuentran en un solo grupo (3). Esta distribución de los *pooles* del Banco explica por qué la estructura del panel inicial no es tan marcada en comparación a los otros tipos de material.

Tomando un valor de  $K=9$ , en el panel inicial (Tabla 4.11), el primer grupo está conformado por 5 accesiones de materiales pre-mejorados o mejorados de Fisiología; el grupo 2 comprende 6 accesiones de diversos tipos de material, a saber: nativos caribeños o de la parte continental o accesiones con algún ciclo de mejora. Los maíces Cristalinos Colorados argentinos y Cateto Sulino de Uruguay están conformando el grupo 3 con 5 accesiones. Cuatro accesiones de maíces cristalinos caribeños de Puerto Rico y Antigua

conforman el grupo 4. El grupo 5 contiene 8 accesiones de poblaciones nativas procedentes de El Salvador, México, una población nativa de Brasil y accesiones del Banco de Germoplasma. El sexto grupo comprende 5 poblaciones mexicanas de razas diversas como Blando de Sonora, Tuxpeño, Vandéño, y *pooles* del Banco y material de Fisiología.

El grupo 7 posee 8 accesiones de materiales nativos de las Islas Vírgenes, República Dominicana, Archipiélago de Guadalupe, Brasil, Costa Rica y una accesión de Tuxpeño mexicano. El grupo 8 está conformado por 4 accesiones nativas de de la raza Caribeño Precoz del Archipiélago de Guadalupe, un Cristalino Amarillo Naranja de Argentina y un Tuxpeño mexicano. Finalmente el grupo 9 comprende 8 accesiones, en su mayoría nativas mexicanas de las Tabloncillo, Tuxpeño, Vandéño y una población de la raza Puya de Venezuela. Estos análisis de estructura poblacional, admitiendo 3 y 9 grupos, serán tomados en cuenta simultáneamente para estudiar los efectos en la variabilidad fenotípica debido a la estructura de población.

En la Tabla 4.11 también se incluyen las 20 mejores y peores accesiones, para el panel seleccionado y la pertenencia a cada uno de los 6 grupos encontrados en el análisis de estructura ( $K=6$ ). En este panel de germoplasma los *pooles* del Banco de Germoplasma se encuentran distribuidos en 4 de 6 grupos y uno agrupado; así mismo las accesiones de la raza Tuxpeño de México se encuentran distribuidas en 5 de los 6 grupos definidos. El grupo 2 es el único que exclusivamente contiene de las mejores accesiones con respuesta adaptativa; en este caso 3 son de las 4 variedades de polinización abierta del Subprograma de Fisiología mientras que los grupos restantes agrupan indistintamente accesiones catalogadas como “buenas” y “malas” en la distribución fenotípica. En este último caso, la descripción y definición de las accesiones en los grupos formados no se manifiesta de una manera clara teniendo en cuenta lo arriba mencionado, dado que en el caso de los *pooles* del Banco de Germoplasma su base genética es tan amplia y en algunos casos parte de esta base tiene materiales con orígenes comunes al resto de las accesiones aquí estudiadas, es el caso de materiales nativos de México, Centroamérica y área del Caribe, por lo que de alguna manera el análisis de estructura poblacional no es capaz de dilucidar de manera eficiente este aspecto (detalles en Tabla 4.2).

**Tabla 4.11** Análisis de estructura poblacional para el panel inicial, el panel seleccionado, poblaciones nativas y nativas en panel seleccionado. Número de grupo correspondiente a cada accesión. Mínimo porcentaje de pertenencia un grupo 51%. Las razas para poblaciones nativas están especificadas.

ID	País	Nombre accesión	Raza	Panel inicial (70)		Panel seleccionado (40)		Nativas (56)	Nativas en panel seleccionado (29)
				K=3	K=9	K=6	K=6	K=4	
1	ARG	ARZM 18 009	Cristalino Colorado	3	3		2		
2	ARG	ARZM 18 004	Cristalino Colorado	3	3		2		
3	ARG	ARZM 1124	Cristalino Colorado	3	3		2		
4	MEX	SINA 29	Tabloncillo	1	9		4		
5	MEX	NVOL 9	Tuxpeño	1	9		4 <sup>§</sup>		
6	MEX	<b>Pool 21 x Pool 22 (♂)</b>		3	1				
7	SLV	SALV 66	Salvadoreño	1	5		1		
8	MEX	<b>Pool 21 x Pool 22 (Best testcrosses)</b>		3	1				
9	MEX	NAYA 171	Tabloncillo	1	5		1 <sup>§</sup>		
10	URY	URUG 630	Cateto Sulino	3	3		2		
11	MEX	SONO 34	Blando de Sonora	3	6		1		
12	GLP	GUAD 15	Cristalino Costeño Tropical	sg	sg		sg <sup>§</sup>		
13	MEX	<b>Pool 22 x Pool 21 (Best testcrosses)</b>		1	5				
14	ARG	ARZM 06 072	Cristalino Amarillo Naranja	3	8		3		
15	MEX	TAMA 12	Tuxpeño	1	sg		sg <sup>§</sup>		
16		<b>OPV-DTPYDR</b>		3*	1*	2*			
17	MEX	JALI 397	Tuxpeño	1	5		1		
18	ARG	ARZM 06 056	Cristalino Amarillo Naranja	3	3		2		
19		<b>OPV-DTPWDR</b>		3*	6*	4*			
20	MEX	NVOL 22	Tuxpeño	1*	9*	3*	4	4	
21	MEX	<b>Pool 20 x CML-494 (Best testcrosses)</b>		1*	5*	4*			
22	MEX	TAMA 37	Tuxpeño	1	6		6		
23	MEX	<b>Pool 19 x CML-495 (Best testcrosses)</b>		1*	6*	3*			
24	GLP	GUAD GP1	Caribeño Precoz	3**	8**	6**	3	2	
25		<b>OPV-DTPW/LPSDR</b>		3*	1*	2*			
26	PRY	PAZM 14107	Tupipy	sg	sg		4		
27	ATG	ANTI GP2	Cristalino Costeño Tropical	2	4		5		
28	MEX	TAMA 36	Tuxpeño	1	sg		sg		
29	MEX	TAMA GP4	Tuxpeño	1**	9**	3**	4	4	
30	MEX	JALI 281	Tuxpeño	2*	sg	1*	sg	3	
31	GLP	GUAD 6	Caribeño Precoz	1	8		3		
32		<b>OPV-LPSDR</b>		3*	1*	2*			
33	BRA	BRAZ 1327	Dentado Rugoso	sg	2		3		
34	MEX	SINA 74	Tuxpeño	1**	5**	4**	1	4	
35	BRA	BRAZ 2306	Dentado Paulista	1*	sg**	sg**	sg	sg	



### Continuación Tabla 4.11

ID	País	Nombre accesión	Raza	Panel inicial (70)		Panel seleccionado (40)	Nativas (56)	Nativas en panel seleccionado (29)
				K=3	K=9	K=6	K=6	K=4
36	PRI	PUER 5	Mezcla (Cristalino Cubano)	2*	4*	1*	5 <sup>§</sup>	1
37	VIR	SCRO GP1	St Croix	2**	7**	5**	6	3
38	MEX	<b>Pool 26 x CML-451 (Best testcrosses)</b>		1*	2*	3*		
39	PRI	PUER GP1	Cristalino Cubano	2*	4*	1*	5 <sup>§</sup>	1
40	MEX	<b>Pool 23 x Pool 24 (♂)</b>		2*	2*	3*		
41	DOM	RDOM GP12	Chandelle	2	7		6	
42	MEX	NAYA 178	Tuxpeño	sg**	sg**	6**	1	2
43	PRI	PUER GP6	Tusón	2**	7**	1**	6	3
44	MEX	<b>Pool 24 x Pool 23 (Best testcrosses)</b>		2*	2*	5*		
45	BRA	BRAZ 2314	Dentado Paulista	1**	5**	5**	1	4
46	MEX	SINA 30	Tuxpeño	1**	sg**	6**	1	2
47	MEX	<b>TUXPEÑO CREMA 1</b>		1*	5*	4*		
48	VEN	VENE 484	Puya	2	2		6	
49	CUB	CUBA 82	Tusón	2**	sg**	1**	6	3
50	GLP	GUAD GP2	Caribeño Precoz	sg*	8*	6*	3 <sup>§</sup>	2
51	PRY	PAZM 2019	Tupipy	sg	sg		sg	
52	PRY	PAZM 10043	Tupipy	sg	sg		4	
53	BRA	BRAZ 2394	Dentado Paulista	2**	7**	5**	6	3
54	ARG	ARZM 06 079	Cristalino Amarillo Naranja	sg**	sg**	sg**	sg	4
55	HTI	HAIT GP3	Haitiano Amarillo	2	7		6 <sup>§</sup>	
56	MEX	NAYA 174	Tuxpeño	1**	9**	6**	sg	2**
57	VEN	VENE 783	Puya	1	9		sg	
58	MEX	<b>Pool 23 x Pool 24 (Best testcrosses)</b>		2*	sg*	sg*		
59	MEX	JALI 286	Tuxpeño	1**	8**	6**	3	2
60	MEX	JALI 285	Vandeano	sg*	6	sg*	1	2
61	PRI	PUER 19	Cristalino Costeño Tropical	2**	4**	1**	5	1
62	GLP	GUAD 3	Cristalino Costeño Tropical	2**	7**	5**	6	3
63	MEX	SONO 72	Tuxpeño	1**	sg**	sg**	1	2
64	GLP	GUAD 9	Caribeño Precoz	1*	8*	6*	3 <sup>§</sup>	2
65	MEX	NAYA 157	Vandeano	1	9		4	
66	PRY	PAZM 6053	Tupipy	2*	sg*	3*	sg <sup>§</sup>	4
67	MEX	TAMA 125	Tuxpeño	1**	9**	3**	4	4
68	BRA	BRAZ 1192	Dentado Rugoso	2**	2**	3**	4	4
69	CRI	CRIC 81	Salvadorenño	2**	7**	5**	6	3
70	MEX	JALI 280	Tuxpeño	2*	7*	5*	6	3

\*Accesiones que se encuentran entre las 20 mejores en IS+Antesis definido anteriormente. \*\* Accesiones que se encuentran entre las 20 peores en IS+Antesis definido anteriormente. § Poblaciones nativas que se encuentran entre las 10 mejores en IS-Nativo definido anteriormente. sg Accesiones que no pertenecieron a ningún grupo para cada análisis independiente de estructura.

El promedio de ancestro para las poblaciones nativas fue de 82,14% para el 80,85% de las poblaciones que pertenecieron a uno de los seis grupos. La Tabla 4.11 también expone las 10 mejores poblaciones basadas en IS-Nativo definido anteriormente y el grupo al que pertenecieron según análisis de estructura poblacional. En este caso, las 10 mejores poblaciones se encontraron distribuidas en 5 grupos de 6, y una de ellas en poblaciones no agrupadas.

Tanto la clasificación de las poblaciones nativas como para el panel seleccionado no es del todo clara, sin embargo, se pueden interpretar algunos grupos con características generales por su origen geográfico, a saber: se puede definir el grupo 1 como un compendio de 5 Tuxpeños y 3 razas de México (Tabloncillo, Vandeño y Blando de Sonora), con la raza Salvadoreño del Salvador y Dentado Paulista de Brasil (Tabla 4.11). El grupo 2 es una conjunción de 4 poblaciones de maíces del sureste de Sudamérica representado por Cristalinos de Argentina además de la población de la raza Cateto Sulino de Uruguay. El grupo 3 reúne poblaciones mixtas, 4 poblaciones de la raza Caribeño Precoz de Guadalupe con una población de Tuxpeño de México, Dentado Rugoso de Brasil y Cristalino Amarillo Naranja de Argentina. El grupo 4 contiene 6 poblaciones mexicanas de las razas Tuxpeño, Vandeño y Tabloncillo con una población de Brasil y dos accesiones de la raza Tupipy de Paraguay. Los grupos 3 y 4 podrían representar las relaciones entre los maíces de las tierras bajas de México con los maíces caribeños y del sureste de Sudamérica encontradas también el capítulo anterior. Las 4 poblaciones que conforman el grupo 5 son todas caribeñas de maíces Cristalinos, y el grupo 6 es un grupo mixto de maíces caribeños de Islas Vírgenes, República Dominicana, Puerto Rico, Cuba, Haití, Guadalupe con poblaciones de Venezuela, Brasil, Costa Rica y México.

Es interesante apuntar que la accesiones que no agrupan con otras accesiones tampoco lo hacen de manera general a través los análisis de estructura analizados, demostrando que su particularidad y aislamiento no es tanto un problema de agrupamiento *per se*, realmente tiene que ver más con una genética distinta a las otras accesiones. Por lo tanto, estas accesiones representan un material de interés específico con el potencial de contener alelos raros y específicos. Por ejemplo PAZM 2019 no agrupa en ninguno de los análisis de estructura realizados. Las accesiones BRAZ 2306 y ARZM 06 079 se encuentran en los 5

tipos de análisis y agrupan únicamente un vez y se quedan aisladas 4 veces. Para darle peso a esta hipótesis reportamos que PAZM 2019 y ARZM 06 079 son 2 accesiones con un ASI de 0 bajo condiciones de sequía (Tabla 4.5), y como se indicó pueden ser fuentes muy interesantes de alelos para este carácter secundario importante para tolerancia a sequía, y BRAZ 2306 es la accesión con el mejor rendimiento en condiciones normales  $6,79 \text{ tm ha}^{-1}$  y por lo tanto debe tener alelos de vigor muy relevantes para mejoramiento.

#### **4.3.8 Análisis descriptivo del rendimiento basado en análisis de estructura poblacional**

La variabilidad fenotípica en condiciones de estrés como se mencionó anteriormente, señala un amplio rango en la variación del rendimiento calculado dentro de las poblaciones para las tres repeticiones y en los grupos formados por las poblaciones según clase de material. De igual manera esta variación está presente cuando se analiza el rendimiento en el panel inicial (Admitiendo  $K=3$  y  $K=9$ ), panel seleccionado y poblaciones nativas según los grupos formados por el análisis de estructura para cada caso y para condiciones de regadío normal (Tabla 4.12).

Bajo condiciones de estrés, las poblaciones nativas de los grupos 2 y 5 formados en su mayoría por maíces cristalinos mostraron los valores más altos de rendimiento y con una variabilidad fenotípica más amplia de este carácter mientras que, los grupos 4 y 6 presentaron los más bajos rendimientos (Tabla 4.12).

El grupo 6 de las nativas presenta el rendimiento promedio más bajo en condiciones de estrés ( $0,04 \text{ tm ha}^{-1}$ ), pero el más alto en condiciones de riego normal ( $4,18 \text{ tm ha}^{-1}$ ) indicando presencia de alelos interesantes para desempeño agronómico y adaptaciones a condiciones de riego normal que no tiene mucho que ver con vigor. El grupo 2, tiene uno de los rendimientos promedio más alto en condiciones de estrés ( $0,11 \text{ tm ha}^{-1}$ ) y el más bajo en condiciones de riego normal ( $1,12 \text{ tm ha}^{-1}$ ) estos valores de rendimiento dan un ITS promedio de 0,1, un valor alto en este estudio (Tabla 4.5) e indica la presencia de alelos favorables a sequía en este grupo.

**Tabla 4.12** Rendimientos promedio ( $\text{tm ha}^{-1}$ ) y desviación estándar correspondiente para el panel inicial ( $K=3$  y  $K=9$ ), panel seleccionado ( $K=6$ ) y poblaciones nativas ( $K=6$ ) bajo condiciones de estrés y riego normal basado en los grupos definidos por el análisis de estructura poblacional.

Grupos	Condiciones de estrés	Regadío normal
	Rendimiento $\text{tm ha}^{-1}$	Rendimiento $\text{tm ha}^{-1}$
<b>Panel inicial</b>		
1	$0,11 \pm 0,10$	$3,94 \pm 2,21$
2	$0,08 \pm 0,09$	$4,54 \pm 1,97$
3	$0,34 \pm 0,32$	$3,68 \pm 2,80$
Mixtos no agrupados	$0,07 \pm 0,09$	$3,48 \pm 1,99$
1	$0,67 \pm 0,21$	$6,25 \pm 1,51$
2	$0,10 \pm 0,11$	$5,16 \pm 3,12$
3	$1,12 \pm 0,08$	$1,12 \pm 0,35$
4	$0,11 \pm 0,08$	$3,94 \pm 0,96$
5	$0,17 \pm 0,14$	$4,65 \pm 2,50$
6	$0,23 \pm 0,30$	$3,97 \pm 2,61$
7	$0,04 \pm 0,04$	$4,53 \pm 1,22$
8	$0,06 \pm 0,03$	$4,23 \pm 1,67$
9	$0,06 \pm 0,07$	$2,26 \pm 0,54$
Mixtos no agrupados	$0,09 \pm 0,10$	$3,96 \pm 0,54$
<b>Panel seleccionado</b>		
1	$0,05 \pm 0,08$	$4,07 \pm 1,51$
2	$0,65 \pm 0,20$	$6,81 \pm 1,89$
3	$0,10 \pm 0,09$	$4,49 \pm 2,73$
4	$0,27 \pm 0,30$	$6,03 \pm 2,48$
5	$0,07 \pm 0,08$	$4,89 \pm 2,31$
6	$0,05 \pm 0,04$	$4,18 \pm 1,53$
Mixtos no agrupados	$0,08 \pm 0,11$	$4,49 \pm 1,92$
<b>Poblaciones Nativas</b>		
1	$0,09 \pm 0,06$	$3,02 \pm 1,72$
2	$0,10 \pm 0,08$	$1,12 \pm 0,61$
3	$0,08 \pm 0,02$	$3,96 \pm 1,83$
4	$0,05 \pm 0,06$	$2,66 \pm 1,62$
5	$0,11 \pm 0,08$	$3,94 \pm 1,51$
6	$0,04 \pm 0,04$	$4,18 \pm 1,57$
Mixtos no agrupados	$0,10 \pm 0,010$	$3,63 \pm 1,74$

En condiciones de estrés en el panel inicial el grupo 3 representado especialmente por las accesiones sudamericanas y el material del Subprograma de Fisiología presenta el mayor rendimiento y además una amplia variación dentro de su grupo, mientras que en condiciones normales el grupo de accesiones mixtas obtuvo mayor rendimiento. Con un rendimiento más alto en condiciones de sequía ( $0,34 \text{ tm ha}^{-1}$ ), tres veces más alto que el promedio de los otros grupos y el rendimiento un poco inferior a los otros dos grupos ( $0,36 \text{ tm ha}^{-1}$ ) este grupo contiene sin dudas alelos favorables para tolerancia a sequía. En el panel seleccionado, el grupo 2 presenta el mayor rendimiento para condiciones de estrés y regadío normal, y se encuentran en este grupo, de manera análoga al panel inicial la mayoría de las poblaciones del Subprograma de Fisiología. También el grupo 4 en panel seleccionado presenta el siguiente rendimiento más alto bajo estrés y condiciones normales, este grupo está formado por una accesión de fisiología, un *pool* del Banco, una accesión de Tuxpeño y la población Tuxpeño Crema, evidenciando especialmente el vigor que presenta este grupo y la reivindicación de la importancia de la raza Tuxpeño para los programas de mejoramiento.

El valor promedio del rendimiento en las 10 poblaciones mixtas que no pertenecieron a ninguno de los 6 grupos definidos también se encuentra entre los mayores promedios de rendimiento bajo estrés, cabe anotar que en estas poblaciones no agrupadas se encuentra la población GUAD 15 de la raza Cristalino Costeño Tropical (ID 12) que presentó el más alto rendimiento de todas las poblaciones nativas ( $0,30 \pm 0,22$ , Tabla 4.5). Contrariamente a condiciones de estrés, en condiciones normales, los grupos 3, 5 y 6 obtuvieron los rendimientos más altos. Especialmente, el grupo 2 presenta el menor rendimiento bajo condiciones normales, esto implicaría que estos maíces cristalinos del sur de Sudamérica poseen una acumulación de alelos desfavorables para vigor.

Este análisis descriptivo del rendimiento basado en análisis de estructura poblacional, además de identificar accesiones con alelos específicos interesantes, las accesiones que no agrupan bien, permiten también identificar grupos de accesiones con alelos de interés y enfocar la búsqueda de estos alelos en un subconjunto de las accesiones totales.

#### **4.3.8 Efectos en la variabilidad fenotípica debida a la estructura de población ( $R^2$ )**

Como se ilustra en la Tabla 4.13, para todos los caracteres utilizados en el cálculo del Índice de Selección se calcularon los efectos de variación fenotípica debido a la estructura poblacional, esto debido a la presencia de grupos dentro del panel inicial y del panel seleccionado discutido anteriormente. Las diferencias en los porcentajes de variación fenotípica en los caracteres medidos y teniendo en cuenta solo el número de grupo al que pertenecieron las accesiones en el estudio de estructura, variaron entre un 6% para número de mazorcas por planta y enrollamiento de hoja y un 24% para senescencia visual y siempre mayor cuando se incrementa el número de grupos presentes en el panel inicial de  $K=3$  a  $K=9$  (Tabla 4.13).

**Tabla 4.13** Estadística descriptiva y variación fenotípica en condiciones de estrés explicada por estructura de población ( $R^2$ ), expresada como porcentaje para 8 caracteres medidos en una localidad.

Carácter	Panel inicial (70)					Panel seleccionado (40)							
	Min	Max	Promedio ± DS	$R^2$ ACoP	$R^2$ número de grupo	$R^2$ matrix	$Q$	Min	Max	Promedio ± DS	$R^2$ número de grupo	$R^2$ $Q$ matrix	
					PC=10	K=3	K=9	K=3	K=9			K=6	K=6
Rendimiento (ton ha <sup>-1</sup> )	0	1,77	0,14 ± 0,20	43,0	36,0	57,0	17,0	46,0	0	1,14	0,14 ± 0,22	57,0	56,0
No, Mazorcas por planta	0	2,62	0,14 ± 0,20	49,0	26,0	32,0	24,0	43,0	0	0,69	0,12 ± 0,14	58,0	58,0
ASI (días)	-3	15	3,73 ± 4,70	20,0	15,0	22,0	9,0	16,0	-14	15	3,96 ± 4,89	20,0	*
SPAD1	6,02	44,1	29,86 ± 7,97	32,0	27,0	38,0	15,0	49,0	17,56	44,1	32,32 ± 5,85	14,0	*
SPAD2	1,36	29,06	9,03 ± 5,48	31,0	18,0	37,0	4,0	37,0	2,76	29,06	9,94 ± 5,37	15,0	14,0
Enrollamiento de hoja	1	23	3,52 ± 1,74	34,0	21,0	27,0	12,0	33,0	1	5	3,27 ± 1,07	20,0	0,4
SEN1	1	8	4,08 ± 13,8	50,0	23,0	47,0	17,0	52,0	1	6	3,53 ± 1,01	18,0	0,6
SEN2	2	10	6,42 ± 1,97	49,0	19,0	30,0	11,0	49,0	2	9	5,65 ± 1,52	21,0	0,6

\*No se encontró variabilidad fenotípica debido a la estructura poblacional

Similarmente, para el estudio conducido teniendo en cuenta la matriz de proporción, los porcentajes de variación fenotípica debido a la estructura del panel para los caracteres medidos fueron mayores para  $K=9$  que para  $K=3$ , y las diferencias entre las variaciones

fenotípicas tomando la matriz de proporción estuvieron en un rango mayor, entre un 7% para ASI y un 38% para la segunda medida de la senescencia visual ( $Sen_{120DDS}$ ). Como se había mencionado, se evaluó también a manera de prueba, la variación fenotípica debido a la estructura utilizando el décimo componente del ACoP, en este caso se observó una diferencia entre las variaciones fenotípicas a través de todos los caracteres cuando se utilizó la matriz de proporción de ancestro ( $Q$ ) cuando se es más estricto en la estructura ( $K = 9$ ) de solo un 5% dando de esta manera solidez a los resultados.

Para los dos enfoques de análisis de variación fenotípica en el panel inicial y para 9 grupos asumidos ( $K=9$ ), el porcentaje de variación fenotípica debido a la estructura fue superior al 10% para todos los caracteres, encontrándose una gran porcentaje de variación fenotípica para la gran mayoría de los componentes de rendimiento y caracteres secundarios, con el menor valor encontrado para ASI (22% y 16%) y el mayor valor para rendimiento (57% y 46%). Mirando los resultados obtenidos para el panel seleccionado ( $K=6$ ) y considerando la variación fenotípica expresada por la estructura del panel obtenida con la matriz de proporción, los porcentajes de variación fenotípica disminuyeron en más del 20% para los caracteres secundarios involucrados en el IS, pero se mantuvieron altos, alrededor del 50%, para los componentes del rendimiento, esto podría significar que la eliminación de las accesiones por fuera del rango de floración masculina ( $85 \leq \text{Antesis} \leq 95$  días) afecta la estructura poblacional inicial y por consiguiente la variabilidad debido a la estructura de los valores de los parámetros secundarios para las accesiones que se mantienen en el panel seleccionado.

De otro lado, y considerando los altos valores de variación fenotípica debido a la estructura del panel inicial y seleccionado para rendimiento de grano, se procedió a realizar un análisis similar, pero solamente para las accesiones de poblaciones nativas presentes en el panel seleccionado (29 de 40). En este caso, los porcentajes de variación fenotípica explicada por la estructura poblacional ( $K=4$ ) disminuyeron significativamente para los componentes del rendimiento (15%) (Tabla 4.14). Sin embargo, este nuevo panel de razas nativas seleccionadas explicaron más de la varianza fenotípica para algunos de los caracteres secundarios como senescencia (>25%). Observando estos interesantes resultados, se realizó el mismo análisis para las accesiones de poblaciones nativas totales, 56 del panel

inicial ( $K=6$ ). Los resultados indican que la estructura del panel no expresó variabilidad fenotípica fiable medir para rendimiento de grano, y baja variabilidad para mazorcas por planta (4%) y ASI (13%), mientras que se registraron altos valores para los otros caracteres secundarios. Estos resultados corroboran, a pesar del número de marcadores empleados en este estudio de exploración del uso de materiales nativos heterogéneos para asociación, que hay una alta probabilidad de encontrar alelos favorables para ASI y número de mazorcas por planta, como se discutió anteriormente analizando germoplasma interesante basado en la distribución fenotípica bajo condiciones de estrés.

**Tabla 4.14** Estadística descriptiva y variación fenotípica explicada por estructura de población ( $R^2$ ), expresada como porcentaje para 8 caracteres medidos en una localidad.

Carácter	Nativas seleccionadas (29)					Nativas totales (56)				
	Min	Max	Promedio $\pm$ DS	$R^2$ número de grupo $K=4$	$R^2$ $Q$ matriz $K=4$	Min	Max	Promedio $\pm$ DS	$R^2$ número de grupo $K=6$	$R^2$ $Q$ matriz $K=6$
Rendimiento (tm $ha^{-1}$ )	0	0,33	0,05 $\pm$ 0,08	*	15,0	0	0,56	0,07 $\pm$ 0,10	5,0	*
No, Mazorcas por planta	0	0,32	0,07 $\pm$ 0,07	6,3	8,0	0	2,62	0,11 $\pm$ 0,21	1,0	4,0
ASI (días)	-3	15	4,43 $\pm$ 5,42	*	3,0	-14	15	3,88 $\pm$ 4,98	2,0	13,0
SPAD1	17,6	41,2	31,01 $\pm$ 5,76	*	*	6,02	42,54	28,70 $\pm$ 7,97	9,0	33,0
SPAD2	2,76	22,1	8,42 $\pm$ 4,23	*	1,0	1,36	27,48	8,18 $\pm$ 4,89	4,0	8,0
Enrollamiento de hoja	2	5	3,28 $\pm$ 1,02	6,0	9,0	2	5	3,48 $\pm$ 1,08	6,0	25,0
SEN1	2	6	3,79 $\pm$ 0,94	1,0	38,0	1	8	4,31 $\pm$ 1,35	8,0	36,0
SEN2	3	9	6,12 $\pm$ 1,40	*	24,0	3	10	6,79 $\pm$ 1,89	8,0	31,0

\*No se encontró variabilidad fenotípica debido a la estructura poblacional

#### 4.4 Discusión

La evaluación fenotípica de accesiones de material nativo no es fácil dada a la alta variabilidad genética entre genotipos de una misma accesión y una adaptabilidad reducida debido a un fondo genético no mejorado, estos dos factores generan a una variación fenotípica muy alta dentro de, y a través de las accesiones (Warburton et al., 2008). Explicar esta variación fenotípica natural en términos de cambios en la secuencia del ADN representa un desafío principalmente por la gran interacción genotipo - ambiente (G x E)



generado por el problema de adaptabilidad de las accesiones, y una G x E que se incrementa cuando el material es evaluado bajo condiciones de sequía (Cooper et al., 2006). Sin embargo, el estudio y la exploración de la variabilidad genética de estos materiales para el descubrimiento de alelos favorables para el mejoramiento de plantas han despertado un gran interés desde las últimas décadas (Xiao et al., 1996; Tanksley and Couch 1997; Salwana y Pollack, 2006), y sigue siendo un objetivo para numerosos grupos de investigación (Glaszmann et al., 2010; Varshney et al., 2010; Xu et al., 2012). Considerando la diversidad de las accesiones nativas, el mapeo por asociación parece ser el método apropiado para la identificación de estos nuevos alelos. Esta estrategia es todavía más relevante para tolerancia a estreses abióticos si se toma en cuenta que los alelos que confieren tolerancia a estos estreses abióticos, como la sequía, se encuentran en bajas frecuencias en el germoplasma elite seleccionado para condiciones normales (Sawkins et al., 2006). Además, y de manera general, el material nativo se desarrolla en condiciones de campo precarias y tiene que haber desarrollado mecanismos de tolerancia a estreses abióticos para poder sobrevivir (material exótico y silvestre) y producir granos (poblaciones nativas cultivadas) en las condiciones del medio ambiente donde se encuentran (Dwivedi et al., 2008). Tomando en cuenta todos estos elementos la caracterización fenotípica y genotípica de materiales nativos y/o heterogéneos, en algunos casos exóticos, tiene el potencial de acceder a nueva diversidad genética presente en accesiones nativas a través del descubrimiento de variación alélica funcional asociada a caracteres agronómicos de interés (Couch et al 1997; Dwivedi et al., 2008)

#### **4.4.1 Método de selección del material: el poder del Índice de Selección**

La respuesta de la planta bajo condiciones de estrés es altamente compleja y como se reportó en el párrafo anterior la evaluación de materiales heterogéneos aporta un grado mayor de dificultad en el discernimiento de caracteres complejos como la sequía. En este estudio, a pesar de las grandes diferencias en fenología (fecha de floración), el régimen predicho para el establecimiento de la relación de riegos por déficit de agua fue satisfactorio; es decir, se estableció un estrés severo que fue confirmado dadas las diferencias en rendimiento, que representaron una disminución del más del 70% bajo condiciones de estrés. Estas condiciones de estrés severas son necesarias para activar los

mecanismos de tolerancia a la sequía, ya que hasta con una reducción del 50% en rendimiento son los mecanismos de vigor que siguen predominantes (Blum, 1988; Blum, 2006). Las diferencias en el rendimiento de grano y en los demás caracteres evaluados bajo condiciones de estrés permitieron obtener una distribución fenotípica basada en el cálculo de un Índice de Selección (IS) y la selección de las mejores y peores accesiones con respuesta adaptativa.

La selección de genotipos para ser utilizados como progenitores en programas de mejoramiento, o mejor dicho pre-mejoramiento, depende mucho de los caracteres de enfoque así como de los ambientes donde se va a cultivar el material mejorado al final del día (Xu, 2003). Teniendo en mente el objetivo de tener un panel de genotipos adecuados para llevar a cabo un estudio de mapeo por asociación en materiales heterogéneos, y tomando los costos que representaría (por número de muestras a evaluar), la decisión de utilizar un Índice de Selección basado en los datos fenotípicos para reducir el número de accesiones y que permitiera una selección selectiva (fenotipos contrastantes) apareció como una opción muy relevante. Con 40 accesiones seleccionadas y considerando la variación dentro de cada una de las accesiones, considerando además, que cada accesión esta representada por 15 individuos, no hay duda que la variación genética es apropiada para el análisis de asociación (Flint et al., 2005). Además, el panel seleccionado cubre la mayoría de la diversidad alélica del panel original como se demostró en el análisis de alelos raros. Se generó la idea de utilizar el IS como criterio de selección y clasificación considerando que el objetivo principal de un estudio de asociación con accesiones de un banco de germoplasma es identificar alelos favorables para caracteres adaptativos de interés que influyen sobre el rendimiento pero sin mucha esperanza de identificar alelos que tengan un impacto directo sobre el rendimiento como tal, que podría ser el objetivo de un estudio de asociación con material mejorado (Hansey et al., 2011), ya que el fondo genético de las accesiones nativas no es tan bueno para rendimiento.

Los índices de selección permiten separar genotipos con base en la evaluación simultánea de varios caracteres de interés. Cada método para calcularlos tiene una eficiencia diferente y el que proporcione la ganancia genética máxima y por unidad de tiempo y esfuerzo es el mejor (Cerón-Rojas y Sahagún-Castellanos 2005). En este respecto, el Índice de Selección

desarrollado por Smith (1936) requiere estimaciones de las varianzas y covarianzas de los valores genotípicos y la asignación, frecuentemente subjetiva, de pesos económicos de los caracteres involucrados en la selección. La elección de utilizar el método de Eigen análisis (ESIM) para construir Índices de Selección, con base en el primer componente principal asociado solamente a la matriz de covarianzas fenotípicas, se basó en que este método ha proporcionado mayor respuesta a la selección y no requiere ponderaciones económicas ni estimaciones de varianzas y covarianzas genotípicas (Cerón-Rojas y Sahagún-Castellanos 2005) por lo que está siendo ampliamente usado en programas de mejoramiento del CIMMYT.

#### **4.4.2 Diversidad genética en el panel inicial, panel seleccionado y poblaciones nativas**

El conjunto de las poblaciones nativas reúne la totalidad de la diversidad genética (0,52) encontrada en todo el conjunto de germoplasma estudiado (Tabla 4.7), pues el origen de estas poblaciones abarcan una gran diversidad geográfica, desde México hasta Paraguay y Argentina pasando por el Caribe, por lo que seguramente cuenta con diferentes respuestas al estrés y adaptaciones a las condiciones locales. Por estas características el análisis molecular que se presenta en este estudio es el primero que abarca material latinoamericano tan heterogéneo y diverso.

A través de los tres tipos de accesiones, material mejorado del Banco (*pooles*), de Fisiología y accesiones nativas se reportó gran proporción de alelos raros, bajas frecuencias para la mayoría de los marcadores considerados en este estudio. En otros estudios publicados también se han encontrado muchos alelos raros en accesiones del Banco de Germoplasma, donde el mayor número de alelos únicos fue para poblaciones tropicales entre materiales subtropicales de maduración precoz e intermedia y poblaciones de maíces de clima templado (Reif et al., 2006); en Warburton et al. (2008) se reportaron gran cantidad de alelos únicos, esta vez descritos como alelos únicos por *locus*, siendo para este caso el mayor número de alelos raros para poblaciones de mejoramiento (OPVs) incluyendo trópico, subtropical y tierras altas, seguido por algunas razas de maíz mexicano y el menor número de alelos únicos por *locus* fue para CMLs. En nuestro caso, la magnitud de los alelos únicos encontrados en las poblaciones nativas indican que hay probablemente una

diversidad genética amplia en las variedades locales que no está presente en el material mejorado el día de hoy (Liu et al., 2003; Dwivedi et al 2008). El hecho que la mayoría de esta diversidad se encuentra representada por alelos raros hace difícil su rastreo por métodos de mejoramiento convencional basados solo en la selección fenotípica. Por lo tanto, es importante determinar el valor genético de estos alelos a través de estudios moleculares como mapeo de asociación para poder identificar los que presentan un potencial agronómico y después introducirlos a través de MAS (Ribaut y Hoisington 1998) en material elite.

Un panel apropiado para mapeo por asociación debe abarcar tanta diversidad fenotípica y molecular muy amplia, considerando sin embargo material con una fenología comparable (precocidad, altura de planta, etc.) para reducir efectos no deseados como escape a la sequía o problemas de adaptabilidad a través de diferentes ambientes (Flint-Garcia et al., 2005). En este estudio, se observó un promedio de 7,2 alelos por locus en el panel inicial y 6,8 en el panel seleccionado a través de los 30 marcadores moleculares representando para el panel seleccionado el 94% de los alelos totales, y una diversidad genética de 0,52 y 0,51 para el panel inicial y seleccionado respectivamente (Tabla 4.7). Tanto el número de alelos por *locus* e índices de diversidad entre el panel inicial y el panel seleccionado indican que no se penaliza la diversidad genética con la reducción de 70 a 40 accesiones. La diversidad genética aquí encontrada (0,52) también fue menor a la diversidad encontrada en el capítulo anterior para la colección de 194 poblaciones de maíces latinoamericanos (0,62), evaluados con un número similar de microsatélites. Esto es lógico, ya que se trata de un subconjunto de accesiones con fenología un poco mas restringida ya que se consideró únicamente material tropical.

El número de alelos por *locus* en el panel inicial (7,20) y seleccionado (6,77) fueron intermedios, y los valores de diversidad genética (0,52 en el panel inicial y diferente grupos de accesiones) un poco menores si se comparan con algunos índices de diversidad genética en paneles para estudios de asociación. Por ejemplo se identificaron 6,5 alelos por *locus* y una diversidad genética de 6,1 sobre 55 SSRs en 375 líneas provenientes de Europa y América en un estudio de tiempo de floración en maíz (Camus-Kulandaivelu, et al., 2006); y 8,23 alelos por *locus* y diversidad genética de 0,65 sobre 82 SSRs en un nuevo panel de

asociación para diferentes caracteres de la mazorca, planta y grano, contenido de carotenoides, entre otros, formado por 155 líneas de maíz (Yang et al., 2010). Considerando la diversidad muy amplia del panel de accesiones nativas, y además que en este estudio se habría podido esperar una diversidad mayor en comparación de otros estudios publicados, es importante recordar la metodología muy rigurosa adoptada en nuestro estudio que reduce a casi cero la posibilidad de identificar alelos que no son reales. Estos alelos ficticios, también llamados falsos positivos, pueden resultar de una lectura inapropiada y/o laxista de los resultados de secuenciación, y/o el uso de marcadores que no pasaron por un proceso de validación rigurosa. El riesgo de identificar falsos alelos, e incrementar de manera artificial la diversidad esperada, incrementa con el uso de SSRs con una secuencia repetitiva corta (dinucleótidos). Entonces, y como se discutió en el capítulo anterior la diversidad encontrada en este estudio en comparación de otros estudios publicados puede ser menor considerando el número de microsatélites y a la naturaleza de éstos (di, tri, tetra nucleótida). Además, el programa aquí utilizado para el cálculo de la frecuencias alélicas (Freqs-R) (Franco et al., 2005) elimina artefactos como bandas *stutter*, mencionadas previamente, alelos falsos, basado en los datos experimentales de la optimización de los marcadores aquí utilizados; además, esta optimización y remoción de bandas fantasma no ha sido reportada en los otros estudios. Sin ninguna duda la probabilidad de encontrar falsos positivos con nuestra metodología es muy reducida.

#### **4.4.3 QTLs para sequía y cambios de frecuencia alélica e identificación de *loci* con un potencial agronómico**

Para identificar regiones en el genoma del maíz que afectan la morfología, metabolismo, y caracteres enzimáticos relacionados con la respuesta a la sequía, estudios de QTLs se han llevado a cabo dentro de poblaciones construidas de cruzas entre progenitores representando dos o más líneas contrastantes (Tuberosa et al. 2005; Ribaut et al. 2008; Setter et al., 2011). Ribaut et al. (2004, 2008) identificaron varios QTLs para rendimiento, componentes del rendimiento, ASI y caracteres asociados a tolerancia en sequía en varias poblaciones biparentales de líneas recombinantes (F2:3 y RIL). Con una de estas poblaciones de mapeo (RILs), Welcker et al. (2007) encontró que algunos de esos QTLs de sequía co-localizaron con aquellos caracteres de crecimiento de la hoja en déficit de agua, y

Messmer et al. (2011) en otra de las poblaciones encontró también una buena co-localización de QTL para senescencia y “Stay Green”. Capelle et al. (2010) identificaron QTLs para humedad del grano, contenido de ABA en embriones y endospermo en un etapa tardía de la embriogénesis y en la pérdida de humedad del grano; también se han identificado QTLs para la acumulación de ABA en las hojas que se relacionan con la morfología de la raíz en respuesta a la sequía (Tuberosa y Salvi, 2007). Estudios con poblaciones biparentales basados en Desequilibrio de Ligamento (LD) son muy apropiados para identificar *loci* involucrados en la expresión de los parámetros de interés tomando en cuenta la diversidad alélica de las dos líneas parentales (Bernado et al. 2008). Sin embargo no es un método apropiado para explorar nuevos alelos en panel de diversidad amplia, para este tipo de estudio el mapeo de asociación es el método de predilección (Rafalski et al., 2010).

El mapeo por ligamiento para detectar QTLs generalmente se focaliza en la construcción y análisis de familias que provienen de cruzas entre dos líneas puras endogámicas. Aunque el poder estadístico de detección de estos QTLs es alto, la resolución genética es pobre por el número limitado de eventos de recombinación que ocurren durante el desarrollo de la población de mapeo. Contrariamente, el mapeo por asociación (AM) y el análisis de asociación del genoma (Genome-Wide Association Analysis, GWAS) aprovechan los eventos históricos de recombinación que revelan las áreas de desequilibrio de ligamiento (LD) donde los marcadores moleculares permanecen asociados a caracteres de interés durante muchas generaciones. Sin embargo, el poder de detección del AM y del GWAS es mas bajo que el análisis de ligamiento (Buckler y Thornsberry, 2002; McMullen et al., 2009). Sin embargo, los dos métodos son muy complementarios ya que después de la identificación de alelos de interés a través de un mapeo por asociación es relevante confirmar el valor de estos alelos en poblaciones bi o multiparentales. Esta etapa de validación es crítica ya que el efecto genético de un alelo detectado por asociación puede cambiar de manera significativa en diferentes fondos genéticos. Hay que reconocer también que la etapa de la caracterización fenotípica en un panel de asociación sigue siendo un reto por la diversidad en la fenología del material y el efecto contundente de la adaptabilidad con el desempeño agronómico (Myles et al., 2009).

Comparando la composición alélica entre los *pooles* del Banco de Germoplasma y Fisiología con las poblaciones nativas se identificaron cambios en la frecuencia alélica entre los *pooles* del Banco de Germoplasma y Fisiología con las poblaciones nativas con la mayoría de los marcadores utilizados. Los cambios absolutos de frecuencia alélica encontrados entre materiales sugieren un aumento en la frecuencia de alelos favorables, o la disminución de la frecuencia de alelos no deseables, causados por una selección dirigida ya sea para un buen desempeño agronómico general (*pooles* del Banco contra accesiones nativas), o para caracteres específicos como resistencia a sequía en el caso del material del Subprograma de Fisiología contra accesiones nativas y *pooles* del Banco.

La magnitud de estos cambios no era dependiente de la naturaleza de los SSRs, ya que no hay diferencia significativa entre los marcadores neutrales y los marcados ligados a genes conocidos. Sin embargo, se encontraron algunos cambios de frecuencia alélica interesantes detectados entre los *pooles* del Banco - Fisiología, Banco - nativas, y Fisiología – nativas. Específicamente se citan algunos casos para genes relevantes para sequía y altas temperaturas: i) con el marcador phi008 se detectó el incremento significativo de un alelo favorable entre los *pooles* del Banco y Fisiología, este marcador se derivó del gen *rab15* (*responsive to abscisic acid15*) (<http://www.maizegdb.org/>), importante para la regulación del Acido Abscísico (ABA), una hormona asociada con la tolerancia a la sequía (Prioul et al., 1999; Tuberosa et al., 2002, Sharp, 2006; Setter et al. 2010); ii) el marcador phi076 también es derivado de un gen, *cat3* (*catalase3*) (<http://www.maizegdb.org/>), este gen está ligado a la regulación de la catalasa, una enzima importante para el mantenimiento de procesos celulares normales bajo condiciones de estrés oxidativo debido a altas temperaturas (Guan y Scandalio, 1995; Revilla et al., 2009), con este marcador se encontraron cambios en las frecuencias alélicas en 5 alelos de 7 identificados y para las tres clases de germoplasma; iii) el marcador *umc1545* ha sido derivado del gen *hsp3* (*heat shock protein3*) (<http://www.maizegdb.org/>) (Xiao et al., 2005), este gen está relacionado con la síntesis de un grupo de proteínas cuya expresión se incrementa cuando las células son expuestas a temperaturas elevadas, y relacionado con la expresión del establecimiento de grano bajo condiciones de estrés (Li et al., 2003); con este marcador se encontraron 3 cambios de frecuencia comparando datos a través de las accesiones del Banco – Fisiología

y Fisiología – nativas; iv) también, el marcador phi299852 ha sido ligado con la co-localización del gen candidato hps101 (*heat shock protein 101*) (Nieto et al., 2002; Hao et al., 2010), una proteína de gran importancia para mantener la integridad de las membranas bajo altas temperaturas, y con este marcador se encontraron dos cambios de frecuencia comparando para Banco-Fisiología y dos más para Fisiología-nativas.

Vale la pena comentar mas en detalle los resultados encontrados con el marcador phi076, un marcador que como se mencionó anteriormente es derivado del gen cat3 (*catalase3*). Mirando la composición alélica en este *locus* a través de los diferentes tipos de material se puede concluir que este marcador identifica un alelo raro extremadamente importante para tolerancia a sequía ya que su frecuencia alélica incrementó 35 veces en el material de Fisiología en comparación de las accesiones nativas y los *pooles* del Banco. Este resultado, que re-enfatiza el potencial de los alelos únicos de impactar en el desempeño agronómico de las plantas, aunque este caso es probablemente poco común ya que la probabilidad de incrementar de manera tan espectacular un alelo con una frecuencia original igual o inferior a 1% por solo selección fenotípica es muy baja. Considerando que más de la mitad de los alelos identificados en el material nativo son raros (56%) es muy probable que una buena parte de los alelos con potencial agronómico presente en las poblaciones nativas no hayan sido incorporados todavía en material elite. Este resultado destaca la necesidad de identificar estos alelos a través de estudios genéticos como el mapeo por asociación.

La idea de aprovechar cambios de frecuencia alélica en poblaciones a través de ciclos de selección para identificar alelos favorables y desfavorables para los caracteres seleccionados no es completamente nueva. Un estudio para cuantificar las frecuencias alélicas a través de 8 ciclos de selección recurrente de una familia de hermanos para tolerancia a la sequía fue conducido en CIMMYT y fue reportado en la primera reunión *InterDrought* (Ribaut et al., 1996). Los objetivos específicos de este estudio fueron: i) determinar la magnitud de los cambios de frecuencia alélica y la probable posición en el genoma de los alelos correspondientes; ii) establecer si estas regiones genómicas correspondían a QTLs previamente mapeados para ASI en una población F2 derivada principalmente de germoplasma Tuxpeño; y iii) investigar las implicaciones del análisis concerniente al mejoramiento de poblaciones bajo sequía usando marcadores moleculares.



Usando un conjunto de 20 RFLPs se identificaron cambios en diferentes *loci* que co-localizaron con QTLs relacionados con sequía en otros estudios, y concluyeron que indudablemente la identificación de alelos favorables, los que aumentaron su frecuencia alélica, como una respuesta a la presión de selección fue posible. Otro grupo exploró esta metodología (Labate et al., 1999) para examinar los efectos del mejoramiento en los cambios de frecuencia alélica con 82 RFLPs en dos poblaciones de maíz que se sometieron a selección recurrente recíproca, los resultados mostraron que un 30% de los alelos se extinguieron y un 10% estuvieron cerca de la fijación (frecuencia alélica aproximada de 1) después de 12 ciclos de selección en cada población. Los autores demuestran que estos cambios no pueden ser el resultado de la deriva genética y pero si resultan de la presión de selección aplicada a cada población. También en este estudio se enfatiza que el tamaño efectivo de la población es un parámetro clave para discernir los cambios de frecuencia alélicas debido a la deriva genética de aquellos que resultan de la selección.

Estos estudios, conducidos dentro de poblaciones seleccionadas a través de diferentes ciclos apoyan la metodología considerada en este trabajo de Tesis. Lo más sobresaliente del conocimiento expuesto aquí es que nunca se habían examinado cambios de frecuencia alélica entre material nativo y poblaciones mejoradas, y menos todavía a través de un panel tan diverso. Aunque en los estudios mencionados se reconoce la limitante de la baja densidad de datos genotípicos que limita mucho el potencial de la metodología ya que en estas poblaciones de polinización abierta el desequilibrio de ligamiento es muy pequeño, este limitante ya no será mas un problema con el desarrollo de nueva metodología de caracterización genotípica como se discute mas adelante.

La arquitectura genética de caracteres cuantitativos complejos es generalmente estudiada como el objetivo final en el mejoramiento del desempeño agronómico en los cultivos (Yang et al., 2010); marcadores funcionales son desarrollados y aplicados en programas de mejoramiento molecular después de que los alelos favorables son identificados por análisis de ligamiento o mapeo por asociación (Andersen y Lübberstedt 1993), en este sentido, los resultados aquí generados con SSRs son positivos y alentadores pues se infieren alelos favorables para algunos caracteres de interés. Sin embargo, y aunque algunos de los alelos identificados pueden ser utilizados directamente en programa de mejoramiento (alelo 176

para phi076), este estudio tiene que considerarse como un “proof of concept” y sin ninguna duda se requiere un estudio de mapeo por asociación como etapa siguiente para aprovechar y maximizar el potencial de la estrategia de identificación e introgresión de alelos favorables de material nativo a germoplasma elite. Además, un análisis de mapeo por asociación en estos materiales heterogéneos maximiza el uso de la diversidad genética presente en el maíz al poder utilizar fuera de los materiales elites frecuentemente utilizados en análisis de ligamiento, material diverso con o sin estructura, relaciones conocidas y/o desconocidas (Flint-Garcia et al., 2005; Hansey et al., 2011) como el germoplasma aquí estudiado. Sin embargo, los resultados de este estudio demuestran sin ninguna duda que la metodología funciona y que utilizando marcadores moleculares se puede identificar en accesiones nativas alelos favorables para caracteres de interés y finalmente para fines agronómicos. Estos alelos se pueden identificar mirando cambios de frecuencia alélica a través de material sometido a diferentes tipos de selección o conduciendo mapeo por asociación. Además del descubrimiento de *loci* de interés, los cambios de frecuencia alélica entre las diferentes clases de germoplasma pueden llevar al descubrimiento de genes candidatos involucrados en la expresión de caracteres de interés (Rafalski 2010), y la caracterización genética de los diferentes alelos contenida en cada accesión podría ser muy valiosa para programas de mejoramiento avanzados (Lu et al., 2010).

El valor de los estudios de asociación para identificar genes y alelos relacionados con la variación genética cuantitativa crece a medida que se desarrolla metodología para incrementar la densidad de los datos genotípicos (Hansey et al., 2011). Este incremento explotó en la mayoría de los cultivos con el descubrimiento de los marcadores SNPs, y la densidad de marcadores paso en algunos años de centenares a miles (Yan et al., 2011) y hasta medio millón en arroz (McCouch et al., 2010). Estas nuevas tecnologías hacen de los estudios de asociación del genoma completo (GWAS) una metodología muy atractiva y eficiente (Rafalski et al., 2010). Los bajos costos en secuenciación han motivado a los investigadores a desarrollar un método robusto y rentable de múltiple secuenciación llamado ‘Genotyping-By-Sequencing’ (GBS), y es posible que esta metodología reemplace la caracterización genotípica con marcadores en alta densidad para estudios de diversidad en los años que vienen (Elshire et al. 2011). Sin embargo, el análisis de datos con esta

metodología requiere de la disponibilidad de herramientas de bioinformáticas para la evaluación de la calidad / filtrado, agrupamiento y el montaje, alineación y anotación de secuencias de referencia los cuales pueden crear algún cuello de botella para la aplicación de la tecnología a gran escala.

Aunque la metodología está demostrada y los marcadores de alta densidad para genotipos están disponibles para maíz, el número de poblaciones de mapeo (paneles de asociación) que han sido publicadas sigue limitado considerando la importancia económica de este cultivo. Sin embargo, no hay duda que esta metodología es ampliamente aplicada en el sector privado y los resultados se mantienen en la confidencialidad en las grandes compañías semilleras (Yu y Buckler, 2006).

La primera población de mapeo por asociación pública consistió en 102 líneas diversas de maíz (Remington et al., 2001), nuevas versiones han sido caracterizadas y publicadas más recientemente; Thornsberry et al. (2001) fue el primer estudio en asociar los polimorfismos en la secuencia de ADN dentro del *locus D8* con el tiempo de floración. Estudios posteriores con la misma población de mapeo asociaron el gen candidato *su1* con el sabor dulce (Whitt et al., 2002), *bt2*, *sh1*, y *sh2* con la composición del grano y otros genes candidatos involucrados con la biosíntesis del almidón (Wilson et al., 2004). El escaneado completo del genoma del maíz fue satisfactorio en la identificación de genes clave en la síntesis de aceites (Belo et al., 2008), mientras que estudios de genes candidatos han permitido la disección de las vías bioquímicas que controlan la biosíntesis de carotenoides (Yan et al., 2010) y aceite (Zheng et al., 2008). Las mejores características de los enfoques de ligamiento y de asociación se pueden combinar a través del mapeo de asociación anidado (Nested Association Mapping, NAM), este análisis proviene del mapeo de QTLs a gran escala de una población de 5.000 líneas recombinantes de maíz (RILs) derivadas de cruzar cada una de 25 líneas diversas y endogámicas de maíz con B73 (Yu y Buckler, 2006), Las RILs son "anidadas" en el sentido de que todas ellas comparten un antepasado común, B73, pero cada población tiene un padre o la madre diferente. La estrategia NAM consta de una caracterización genotípica robusta de marcadores o secuenciación de los padres, los SNPs específicos de los padres se utilizan para clasificar cada segmento cromosómico que definen cada RIL en función de si se deriva de B73 o del padre o la

madre correspondiente. De esta manera, los datos de alta resolución genotípica pueden ser proyectadas desde los padres a la descendencia RIL, sin la necesidad de secuenciar toda la descendencia, se caracteriza fenotípicamente la descendencia para los caracteres complejos de interés y finalmente se realiza el análisis de asociación del genoma completo relacionando los caracteres fenotípicos con la caracterización genotípica proyectada en la progenie (Yu et al., 2008; McMullan et al. 2009). Esta metodología ha sido llevada satisfactoriamente identificando genes clave para floración (Buckler et al., 2009), arquitectura de la hoja (Tian et al., 2011), y resistencia a enfermedades (Poland et al., 2011)

Más recientemente y por primera vez para sequía Setter et al. (2011) tomaron 350 genotipos de materiales tropicales y subtropicales existentes en los programas de mejoramiento del CIMMYT para crear un panel con gran diversidad genética y utilizaron mapeo por asociación y genes candidatos para identificar *loci* involucrados con los niveles y acumulación de carbohidratos y ABA en la mazorca y en la inflorescencia masculina o panoja durante el estrés hídrico. Estos avances podrían significar no solo la explotación de la diversidad genética interpoblacional sino el acercamiento al aprovechamiento real de la diversidad genética intraespecífica presente en materiales heterogéneos de maíz como los materiales nativos

#### **4.4.3 Heredabilidad**

La probabilidad de identificar regiones cromosómicas relevantes y una estimación precisa de sus efectos depende de una buena caracterización fenotípica de los materiales genéticos utilizados originalmente para establecer las asociaciones genotipo-fenotipo (Tuberosa et al., 2011). En otras palabras, la efectividad de estos estudios, depende íntimamente de qué tan bien y con qué precisión ha sido evaluado fenotípicamente el carácter de interés en el panel en estudio. De hecho, una baja heredabilidad afecta la probabilidad de detectar la presencia de QTLs (Flint-Garcia et al., 2005; Tuberosa et al., 2011), porque de una manera práctica la heredabilidad representa el límite superior de la proporción de la variación fenotípica del conjunto de datos puede ser explicado en los estudios genéticos (Flint-Garcia et al., 2005). En este sentido, en el panel inicial aquí estudiado la heredabilidad calculada para rendimiento y mazorcas por planta fueron las dos altas ( $> 0,70$ ), en condiciones de estrés

0,91 y 0,74 para regadío, si se tiene en cuenta valores anteriormente reportados entre 0,30 – 0,70 (Tuberosa et al., 2011). Ribaut et al. (2008) también reportaron niveles de heredabilidad, sin embargo con valores un poco más altos (0,4-0,88) para componentes del rendimiento y caracteres secundarios. Los valores más altos se obtienen para caracteres que se miden de manera cuantitativa (peso de granos, ASI, número de mazorcas) y las más bajas para caracteres estimados (por ejemplo: senescencia, enrollamiento de las hojas). Se observa también que aunque la heredabilidad del rendimiento se reduce en general bajo condiciones de sequía, para caracteres secundarios puede incrementar en condición de sequía ya que tiene una mejor segregación en poblaciones biparentales. En este estudio es importante subrayar que la segregación de las accesiones fue muy diversa bajo ambos tipos de riego y que también el número de plantas medidas por cada accesión es mas alto comparando (surcos de 5 metros) con los estudios de Ribaut (surco de 2.5 metros), dato que puede impactar de manera favorable las heredabilidades.

Los valores altos reportados aquí, resultan de la combinación de una metodología muy robusta desarrollada a través de 20 años en CIMMYT, un número de plantas por accesión alto, y de la diversidad muy amplia del material (accesiones de Banco, fisiología y nativas) que genera una buena segregación fenotípica para los diferentes caracteres considerado en el estudio.

Para entender mejor la varianza genética y el impacto del G x E para los diferentes tipos de accesiones, la heredabilidad fue calculada para los diferentes tipos de material. El resultado de esta “disección” de las heredabilidades calculadas en el panel inicial en heredabilidades para material nativo y material mejorado para condiciones de estrés y regadío indican que el ensayo fue bien conducido y controlado a pesar de tener materiales nativos con alta variabilidad genética, que, en general son difíciles de ensayar. En esta disección de la heredabilidad, las accesiones nativas conservaron valores entre rangos intermedio- alto para los componentes de rendimiento ( $>0,40$ ), y para casi la mayoría de los caracteres secundarios. En algunos casos, para caracteres como enrollamiento de hoja y senescencia visual se obtuvieron valores menores a 0,40. Estos valores bajos podrían ser atribuibles a la escala subjetiva usada para la evaluación de estos caracteres durante la toma de datos y estos valores corresponden a lo que se encontró en la literatura.

#### 4.4.4 Efectos de la estructura de población

El efecto de la estructura de población en los caracteres de interés determinado por el porcentaje de variación del carácter explicada por la estructura es una indicación del poder de detección de alelos significativos en el panel que va a ser utilizado. El conocimiento en detalle de las relaciones genéticas entre individuos/poblaciones en un panel de asociación es un factor clave para evitar falsas asociaciones. La estructura de población (Matriz  $Q$ , o matriz de proporción de ancestro) estimada usando STRUCTURE (Pritchard et al., 2000) es una vía para corregir falsas asociaciones debido a las relaciones genéticas del material en estudio. Esto en algunos casos es difícil por la estimación del número de grupos ( $K$ ), como se ha visto en el desarrollo del capítulo anterior y en el presente. El valor más alto de  $\ln[\text{prob}(\text{dato})]$  entregado por STRUCTURE es el que generalmente se toma en la estimación de  $K$ , pero en situaciones reales pocos conjuntos de datos parecen confirmar el modelo (Yang et al., 2010). Por este motivo se siguió también en este capítulo la metodología reportada por Evanno et al. (2005) y explicada en el capítulo anterior (Anexo 7.6). Sin embargo, y a pesar de la dificultad para definir grupos estructurales presentes en el panel de estudio, el uso del ACoP para evaluar la variabilidad fenotípica debido a la estructura poblacional aportó solidez a los resultados cuando se tomó como co-variable la matriz  $Q$ .

Los resultados obtenidos muestran que el incremento del número de grupos asumidos en el análisis de estructura poblacional en un panel estudiado es determinante para el estudio de la variabilidad fenotípica debido a la estructura. Es decir, cuando se incrementa el número de grupos asumidos, se incrementa a su vez el porcentaje de variación debido a la estructura para el carácter bajo estudio, que podría derivar en un sobre ajuste del modelo de análisis. De otro lado, el uso de la matriz  $Q$  como un conjunto de co-variables para calcular los efectos de la estructura de población en la variabilidad fenotípica, más que el uso del número de grupo como co-variable, podría ser más relevante cuando la variabilidad para un carácter dado depende altamente del valor de grupos asumidos dados por el análisis de estructura ( $K$ ). Análogamente, si se usa ACoP, la variabilidad para un carácter estaría dada por el conjunto de datos generados para cada individuo/accesión donde cada componente principal describiría la pertenencia de una individuo/accesión a un grupo o su ascendencia

(Zhu et al., 2008). Utilizando los componentes de la matriz  $Q$  de proporción de ancestro como co-variable, los promedios de la variación fenotípica a través de todos los caracteres medidos donde se encontró variabilidad fenotípica están grandemente afectados, 37% y 21% en panel inicial (para  $K=9$ ) y panel seleccionado respectivamente indicando que la estructura de población es un factor importante que contribuye en la variación fenotípica del panel inicial y seleccionado en este estudio, sin embargo los efectos variaron dependiendo del carácter como se ha expuesto en otros estudios (Yang et al., 2010). Estos valores altos de variación fenotípica debida a la estructura de población implican la importancia de la matriz de proporción  $Q$  extraída del programa STRUCTURE y que debe ser incluida posteriormente en el modelo de análisis de asociación para evitar falsas asociaciones.

Existen grandes diferencias en la variabilidad fenotípica afectada por la estructura de población en caracteres como la fecha de floración. Esto es debido al germoplasma escogido para el panel de asociación. Por ejemplo, está pobremente afectado en el panel de asociación en Yang et al. (2010) ( $R^2 < 2\%$ ), este panel principalmente contiene líneas de clima templado provenientes de China y Estados Unidos; grandemente afectado en el panel formado por 302 líneas de maíz que representan parte de la diversidad genética presente en programas de mejoramiento del sector publico alrededor del mundo en Flint-Gracia et al. (2005) ( $R^2 > 32\%$ ). Moderadamente afectado debido a la estructura ( $R^2 < 22\%$ ) en Hansey et al. (2011) en un panel formado por 1411 líneas de fenología restringida. En el presente estudio, en el que se restringió la fenología en base a la fecha de floración, los porcentajes de variación fenotípica para ASI en panel inicial (16%) y para el panel seleccionado (donde no se encontró variación fenotípica debido a la estructura), apoyan los análisis aquí realizados cuando se restringió la fenología con el tiempo de floración masculina.

Las diferencias en rendimiento de las tres diferentes clases de germoplasma, nativas, *pooles* genéticos desarrollados por el Banco de Germoplasma y, especialmente, las variedades de polinización abierta (OPVs) tolerantes a la sequía del subprograma de Fisiología, que rindieron 10 y 2 veces más que las accesiones nativas y los *pooles* respectivamente en condiciones de sequía, podrían explicar las altas respuestas de variación fenotípica para los caracteres medidos causados por la estructura de población en panel inicial. Estas

respuestas de variación fenotípica podrían ser explicadas debido a los efectos de la intensa selección humana fenotípica que han sido objeto los *pooles* del Banco de Germoplasma y Fisiología. La comparación entre los valores de los porcentajes de variación fenotípica debido a la estructura del panel inicial con las nativas en el panel seleccionado y las nativas en el panel inicial sugieren que la presión de selección fue más alta durante el desarrollo de las OPVs y de los *pooles* del Banco de Germoplasma para rendimiento y número de mazorcas por planta pero fue reducida para los caracteres secundarios aquí estudiados.

Paralelamente, y relacionado con las diferencias entre las tres clases de germoplasma mencionadas en párrafo anterior, las significantes respuestas de variación fenotípica para los caracteres medidos causados por la estructura de población, podrían ser explicadas también por los grupos formados en el análisis de estructura del panel inicial y panel seleccionado. Se encontró, que en la mayoría de los casos se encuentran en un mismo grupo material mejorado con poblaciones nativas debido a orígenes comunes y/o razas presentes en las bases genéticas de estos materiales, y que además se estudian de manera independiente como poblaciones nativas; el caso más notorio es de la raza Tuxpeño de México que forma parte de la base genética de diferentes *pooles* genéticos o materiales mejorados como: Pool 23 x Pool 24 (♂), Pool 24 x Pool 23 (Best testcrosses), Pool 23 x Pool 24 (Best testcrosses), Tuxpeño Crema, OPV-LPSDR y OPV-DTPW/LPSDR (detalles en Tabla 4.2). También, relaciones poblacionales entre materiales de México-Centroamérica-Caribe-costa este de Sudamérica ya se han discutido en el capítulo anterior, sobre maíces Latinoamericanos, y se pone de nuevo en evidencia aquí, en la evaluación de materiales tropicales.

Con todas las relaciones interpoblacionales encontradas, las altas variaciones fenotípicas debido a la estructura para muchos de los caracteres medidos en panel inicial y seleccionado, las diferencias entre las tres clases de germoplasma y la identificación de poblaciones nativas como fuente de alelos para ASI y vigor sugeridos por la distribución fenotípica se evaluó la hipótesis de que la presencia de material mejorado incrementa la variación fenotípica explicada por la estructura poblacional.



Para probar esta hipótesis se hizo también un discernimiento de la variación fenotípica explicada por la estructura poblacional tomando solo las poblaciones nativas presentes en el panel inicial y seleccionado. En este caso, el porcentaje de variación fenotípica debido a la estructura disminuye más de un 80% para rendimiento y número de mazorcas por planta, ASI está pobremente afectado en nativas seleccionadas (3%) y moderadamente afectado en nativas del panel inicial (13%). En general, lo deseado para un análisis de asociación es una pobre variación fenotípica del carácter en cuestión debida a la estructura de población, porque con una alta variación debido a la estructura alelos funcionales cuya distribución coincide con la estructura no serán detectados cuando el modelo incluye las estimaciones de la estructura de población (Flint et al., 2005). Las variaciones fenotípicas de los 8 caracteres aquí estudiados a través de los diferentes enfoques de germoplasma sugieren un gran potencial de estudio para componentes del rendimiento y caracteres secundarios, la clave en el uso del germoplasma está en la flexibilidad que expresa el panel inicial, porque puede ser analizado en diferentes direcciones, es decir, enfatizando en los caracteres donde la variación fenotípica de un carácter específico debido al estructura es pobre o moderadamente afectada, porque mientras que en un sub panel  $i$  el carácter  $j$  podría exhibir una alta variación fenotípica debida a la estructura y en otro sub panel  $k$  la situación podría ser contraria y favorable para un estudio de asociación como se detallará en la sección siguiente.

#### **4.4.5 Observaciones carácter-panel de asociación**

Aunque los resultados presentados son muy alentadores, se puede inferir que análisis de asociación que capturen nuevos alelos para mejoramiento bajo sequía de accesiones del Banco de Germoplasma no es tampoco una tarea fácil. Aunque estos recursos genéticos se han caracterizado en “paneles de diversidad” que capturan una gran parte de la variación genética presente en estos materiales heterogéneos (Warburton et al., 2002; Reif et al., 2004; Dubriel et al., 2006; Warburton et al., 2008) su adaptación y fenología no ha sido tomada en cuenta (Hansey et al., 2011). Precisamente esta falta de información relacionada con los materiales heterogéneos representa uno de los puntos importantes de este trabajo de Tesis a nivel institucional y de investigación aplicada. Puntualizando, los *pooles* de Banco y OPVs del Subprograma de Fisiología del CIMMYT han sido evaluados varias veces en

campo o usados como controles en los ensayos de los respectivos programas pero no habían sido caracterizados genotípicamente; de manera contraria, algunas accesiones que corresponden a materiales nativos han sido caracterizadas genotípicamente por programas nacionales o en el CIMMYT pero con poca evaluación fenotípica que facilite la comparación de su desempeño agronómico y su potencial como fuente de germoplasma para caracteres de interés.

En miras de realizar un análisis de asociación, agrupar diferentes clases de germoplasma podría no ser óptimo, especialmente si algunas accesiones han experimentado presión de selección y otros no, y que generan una fuerte estructura de población que no es recomendada para un estudio de asociación. Sin embargo, este trabajo de Tesis resalta el hecho de que un panel germoplasma heterogéneo puede ser idóneo para llevar a cabo un estudio de asociación para algunos y no necesariamente todos los caracteres que se podrían estar tomando inicialmente en consideración. Respecto a los materiales nativos se demostró que son poseedores de variabilidad alélica importante para caracteres secundarios puesto de manifiesto aquí en la distribución fenotípica, la descripción del rendimiento bajo condiciones de estrés y condiciones normales, y en la variabilidad fenotípica debido a la estructura poblacional.

Un panel compuesto solo por las 56 accesiones de poblaciones nativas presentes en el panel inicial de 70 accesiones de maíz tropical, parece ser apropiado para identificar nuevos alelos a través de estudios de asociación relacionados con componentes del rendimiento bajo sequía, porque para rendimiento no se encontró variabilidad fenotípica debida a la estructura, para mazorcas por planta está pobremente afectado ( $R^2= 4\%$ ); y para sincronía de floración moderadamente afectado ( $R^2= 13\%$ ). Para análisis de asociación de caracteres secundarios como senescencia, enrollamiento de hoja, contenido de clorofila, el panel seleccionado compuesto por 40 accesiones parece ser más adecuado porque las variaciones fenotípicas explicadas por la estructura poblacional de estos caracteres están pobremente afectadas. El panel compuesto por las 29 accesiones de poblaciones nativas en el panel seleccionado, aparece también como un interesante panel de análisis de asociación para casi todos los caracteres estudiados (aparte de la senescencia) pues la variabilidad fenotípica

debido a la estructura estuvo por debajo del 15%. Sin embargo, en este panel se requieren más análisis para entender si el número limitado de accesiones podría sesgar el análisis de asociación.

#### **4.5 Conclusiones**

El material aquí estudiado compuesto por complejos genéticos o *pooles* desarrollados por el programa de pre-mejoramiento del Banco de Germoplasma del CIMMYT, material mejorado para resistencia a sequía del Subprograma de Fisiología del CIMMYT y poblaciones nativas mostró un amplio rango de variación fenotípica para rendimiento y los caracteres secundarios medidos bajo condiciones de estrés hídrico severo y riego normal. Los valores altos de heredabilidad para los caracteres evaluados demostraron la calidad de los datos fenotípicos.

Utilizando un Índice de Selección y considerando el tiempo de floración masculina del panel inicial se seleccionó un subconjunto de 40 accesiones, es decir, 20 accesiones tomadas de cada uno de los extremos de la distribución fenotípica donde 12 de las mejores accesiones provienen del material mejorado, y este panel seleccionado representa el 94% de la diversidad genotípica del panel inicial.

Como lo demuestran los resultados del Índice de Tolerancia a Sequía, el desempeño agronómico del material mejorado bajo condiciones de estrés hídrico refleja más la presencia de alelos favorables para mecanismos de vigor que mecanismos de tolerancia a sequía.

Considerando el desempeño agronómico de las accesiones del Banco para caracteres secundarios importantes para la tolerancia a la sequía, ya que no tiene mucho sentido mirar el rendimiento *per se* de este material no mejorado, se identificaron accesiones nativas con una sincronía floral muy buena y un contenido de clorofila alto bajo condiciones de sequía. Estos resultados muy importantes demuestran el potencial de estas accesiones como fuente de alelos favorables para la expresión de estos caracteres que contribuyen a impactar el rendimiento bajo estrés hídrico. Sin ninguna duda la introgresión de estos alelos de manera

integrada puede mejorar el desempeño agronómico de material elite cultivado hoy día en Latinoamérica.

La gran proporción de alelos raros identificados en este estudio, y el hecho de que para los *loci* que presentan un número alto de alelos, solo 2-3 de ellos representaban la mayoría de la diversidad en términos de frecuencia alélica, sugiere que estas accesiones nativas también fueron sometidas a selección natural o humana (campesinos) a través del tiempo. El concepto de la gran diversidad del maíz parece ser un concepto un poco relativo dependiendo de si se evalúa la diversidad por número de alelos *per se* en cada *locus* o si se consideran las frecuencias alélicas en un material dado.

Este estudio genético indica que una fuente de diversidad importante para mejoramiento se encuentra probablemente en los alelos raros y/o únicos de las accesiones nativas ya que la selección no los eliminó como lo demuestra el número de alelos por *locus* comparable a través de los diferentes tipos de material. Este dato demuestra también la limitación de la selección fenotípica para incrementar las frecuencias alélicas de alelos únicos a través de ciclos de selección; phi076 es una excepción en este contexto.

Considerando la magnitud de los alelos raros y/o únicos encontrados en las accesiones nativas, la dificultad de identificarlos y rastrearlos a través de una selección fenotípica, especialmente si están en repulsión con otros alelos favorables, y el impacto fuerte que podrían tener algunos de ellos sobre el desempeño agronómico de las plantas, se demostró con los resultados observados para el marcador phi076.

No hay ninguna duda que el uso de marcadores moleculares puede impactar de manera significativa el acceso a los alelos favorables que se encuentran en accesiones nativas permitiendo: i) identificar el potencial de estos alelos raros y/o únicos a través de estudios de asociación o cambios de frecuencia a través de los ciclos de selección y ii) rastrear estos alelos favorables en material elite (BC y MAS).

Los análisis de estructura poblacional del germoplasma, combinados con los datos fenotípicos permitieron identificar grupos accesiones con alelos específicos interesantes para vigor, sequía, escape de la sequía, y caracteres secundarios. Basados en estos

resultados, el panel inicial aquí estudiado podría utilizarse de una forma flexible, es decir, utilizando “sub paneles de asociación” que podrían ser usados para diferentes caracteres de interés como componentes del rendimiento y caracteres secundarios de manera independiente.

Este análisis identificó también accesiones con una fenología muy interesante pero que no agruparon con los diferente tipos de material considerado, incluyendo las accesiones mejoradas, indicando una diversidad genética con potencial agronómico que probablemente no fue explotada a través del mejoramiento ya que una proporción alta de los alelos de estas accesiones son únicos o muy raros a través de las otras accesiones consideradas en este estudio.

## 5. PERSPECTIVAS

Durante el desarrollo de este trabajo de Tesis fue evidente la falta de información relacionada con las características particulares de las razas de maíz de una forma compilada, práctica y accesible. Existen razas de maíz muy especiales, por ejemplo, la raza Chococeño del Pacífico Colombiano que germina en condiciones de inundación, condición limitante para muchas variedades mejoradas, la raza Lluteño del norte de Chile que presenta alta tolerancia a la salinidad y al exceso de Boro, o la accesión Michoacan 21 de México que presenta alelos latentes permitiendo a la planta sobrevivir periodos muy largos sin agua poniéndose en un estado de latencia caracterizado por un metabolismo muy lento. Sin embargo, esta información se encuentra muchas veces contenida en los programas nacionales y a veces no es publicada o difundida, representando a veces un cuello de botella dentro del proceso de exploración y búsqueda de opciones para el aprovechamiento de la diversidad de los maíces nativos. Por esta razón, es necesario el intercambio de información relacionada con los atributos y cualidades especiales de las razas de maíz como resistencia a enfermedades, plagas y calidad nutricional. En este aspecto, crear y/o fomentar una red de intercambio sólida de la información y germoplasma basada en un conocimiento integral de los materiales nativos podría generar la rápida identificación de materiales que pueden servir como base genética en el desarrollo de variedades que enfrenten de manera contundente retos debidos a sus ambientes, en particular en países en vías de desarrollo, los posibles escenarios de cambio climático como altas temperaturas, sequía, enfermedades, y sean eficientes en la utilización de nutrientes como Nitrógeno y Fósforo. Con el conocimiento integrado de las razas de maíz se podrían conducir diversos estudios dirigidos en materiales nativos y no al azar, agrupando material diverso, como se hace en la mayoría de los casos. En este contexto, los índices de diferenciación genética podrían ser de gran utilidad y aplicarse desde diferentes puntos de vista, por ejemplo, podrían estudiarse los índices de diferenciación entre materiales que corresponderían a diferentes adaptabilidades: trópico-subtrópico, trópico-valles altos, subtrópico-valles altos, que ligados a una caracterización fenotípica y genotípica robusta además de un sistema de información geográfica podrían revelar *loci* relacionados con la adaptabilidad de estos materiales.

Una caracterización fenotípica y genotípica más amplia de los recursos genéticos de maíz y el entendimiento de la estructura de la diversidad presente podrían mejorar la eficiencia de la utilización de estos recursos para el mejoramiento del cultivo. Con respecto a lo planteado, la caracterización genotípica cuenta con avances progresivos en nuevas herramientas y accesibilidad a la tecnología que puede ser aplicada a los materiales nativos. Sin embargo, aunque en el aspecto de la caracterización fenotípica también hay avances en el descubrimiento de caracteres secundarios e instrumentos para medirlos, los materiales nativos cuentan a nivel general, solo con datos de pasaporte, es decir, datos como lugar de colecta, color y tipo de grano, raza primaria, datos necesarios para ser clasificados en un banco de germoplasma. En este contexto, es necesario que la caracterización fenotípica sea dirigida a la evaluación de los materiales nativos en condiciones de estrés abiótico o biótico que enfrenta el material elite, que ligada a la caracterización genotípica permitiría la identificación de germoplasma útil poseedor de alelos favorables para condiciones adversas.

A pesar de que se ha avanzado en tecnologías computacionales, basadas en caracterizaciones fenotípicas, genotípicas, o ambas para clasificar los recursos genéticos, definir colecciones núcleo, e identificar alelos favorables para parámetros agronómicos, falta mucho para convertir este conocimiento en herramientas tangibles que puedan ser aplicadas rutinariamente en los programas de mejoramiento.

Aunque estudios de diversidad específica (ligados a caracteres de interés) podrían llevarse a cabo, incluso con marcadores SSRs como los utilizados en este estudio si se considera que más de 240 SSRs se han derivado de genes identificados previamente (<http://www.maizegdb.org/>) y la información se encuentra disponible para ser utilizada, es imperativo considerar nuevas herramientas para generar una cantidad mas grande de información genotípica y a menor costo. Lo anterior, no sólo con el fin de caracterizar el germoplasma de maíz disponible, sino también de hacer más eficiente el uso de la diversidad genética que no ha sido explotada y se encuentra en los maíces nativos como se ha puesto de manifiesto en este estudio y que también poseen sus parientes silvestres. Sin embargo, alcanzar estos objetivos representan retos importantes, entre ellos: i) una caracterización fenotípica amplia y de calidad para caracteres de interés, incluyendo evaluación multilocalidades en campo, e ii) inversiones substanciales en genómica y

disciplinas asociadas se requerirán para la aplicación de las nuevas tecnologías especialmente en el sector público y en los diferentes programas nacionales de investigación en los países latinoamericanos. En estos días, estas inversiones al nivel institucional no deben ser tanto para el desarrollo de las tecnologías por si mismas, ya que consorcios e iniciativas grandes se encargan de esto, pero más en el sector del manejo de los datos genotípicos a saber: almacenamiento, acceso e intercambio, documentacion y control de calidad.

De manera general, en base a los resultados presentados en este estudio, en la exploración de poblaciones heterogéneas de maíz para análisis de mapeo por asociación se deduce que esta metodología representa una alternativa para el aprovechamiento de la diversidad genética presente en los materiales heterogéneos de maíz. Una vez que los datos genotípicos están disponibles, el análisis de asociación se puede realizar utilizando los datos fenotípicos obtenidos bajo condiciones de sequía y de regadío. Combinando los resultados de los estudios de asociación en condiciones de agua limitada y regadío normal permitirá la identificación de los *loci* implicados en la tolerancia a la sequía y en vigor de la planta bajo este estrés. También se demostró el interés de la co-localización de genes candidatos, disponibles en diferente bases de datos (Hao et al., 2010), con *loci* que presentan nivel de asociación significativas para los parametros de interés, o que presentan cambios de frecuencias alélicas importantes a través de selección (ver resultados con Phi 063 y 076).

Aunque la metodología de los análisis de ligamiento ha permitido la identificación de muchos QTLs en caracteres de interés de importancia agronómica, varios desarrollos metodológicos recientes están permitiendo promover esta tecnología para ser todavía más potente. Por ejemplo el uso de marcadores como SNPs permite incrementar el número de datos genotípicos de manera significativa, reduciendo el costo por unidad de dato por un factor de aproximadamente al mínimo de 20 veces en comparación de SSRs. Estos SNPs pueden ser usados de manera individual, ya que hay miles de estos tipos de marcadores disponibles para la mayoría de los cultivos de interés incluyendo el maíz, o en tandas o grupos a través del uso de microarreglos que son la plataforma de marcadores de alta densidad para los datos genotípicos como por ejemplo: Veracode, GoldenGate, Infinium (McCouch et al., 2010). Gracias a estas nuevas herramientas genómicas, para maíz, el



último chip agrupa 50.000 SNPs, se sugiere que un estudio de asociación podría ser conducido combinando una caracterización fenotípica de calidad, como la realizada en este estudio, y una caracterización genotípica de alta densidad permitirán un estudio de asociación a través de todo el genoma (GWAS) en comparación de una asociación simple. Este nuevo método permitiría tener una visión simultánea de la variación genética de un gran número de *loci* que asociada a su caracterización fenotípica permitiría la disección de caracteres complejos relacionados con adaptabilidad, resistencia a estreses bióticos y abióticos, etc. Sin ninguna duda, el acceso a una plataforma de genotipificación de alta densidad incrementa de manera significativa el poder de los análisis de ligamiento y además, esta tecnología es accesible a todos, incluyendo programas nacionales que no cuentan con la infraestructura para la caracterización genotípica, siendo esto posible a través de laboratorios de servicios para caracterización molecular como se ha promovido a través de la Plataforma Integrada de Mejoramiento (Integrated Breeding Platform, por sus siglas en inglés; <https://www.integratedbreeding.net/> ) promovida por el GCP (Generation Challenge Programme, por sus siglas en inglés; <http://www.generationcp.org/>).

Concretamente, para la continuación del trabajo de investigación presentado en esta Tesis, una alternativa que actualmente se ha revisado es la inclusión del material aquí evaluado dentro de un mega proyecto que se está llevando a cabo en el CIMMYT (Seed of Discovery) y en el cual se discute el uso de GBS para la caracterización genotípica de una gran parte de las regiones codificadoras del genoma aprovechando las ventajas tecnológicas que este método representa. Esta metodología de secuenciación es basada en la reducción de la complejidad del genoma con enzimas de restricción por la creación de “librerías de GBSs”, donde no se necesita un genoma de referencia. Otro punto a favor de esta metodología es el soporte bioinformático disponible y el costo efectivo que representa pues se indica que con una inversión de 8.000 dólares pueden identificarse y mapearse hasta 200.000 marcadores en maíz, representando menos de 20 dólares por muestra y las estimaciones para un futuro cercano es la reducción de hasta 4-5 veces (Elshire et al., 2011). Con este precio para la caracterización genotípica, esta tecnología abre nuevas perspectivas para los estudios de asociación y la secuenciación, aunque parcial, de las 27.000 accesiones disponibles en el Banco de Germoplasma del CIMMYT o de otros

repositorios de germoplasma públicos se convierte en una realidad. Sin embargo, hay que acordarse de que, sin datos fenotípicos de calidad no hay interpretación posible para determinar el valor alélico de estas accessiones y el cuello de botella quedaría en la caracterización fenotípica.

Los resultados de esta Tesis abren la puerta a otra clase de enfoque, aparte del análisis de asociación que podría ser considerado. La identificación de *loci*/alelos de interés podría llevarse a cabo de manera satisfactoria analizando los cambios de las frecuencias alélicas a través de ciclos de selección, por ejemplo, trabajando por ejemplo con OPVs y *pooles* vs poblaciones nativas. El resultado de este enfoque, presentado aquí usando SSRs son alentadores, pero obviamente considerando el muy pequeño Desequilibrio de Ligamiento en maíz en este tipo de material (Flint et al., 2003) una alta densidad de información genotípica es requerida para identificar cambios de frecuencia alélica en una amplia gama de genes que influyen en los caracteres de interés dentro y a través de las poblaciones.

La comparación de los cambios de frecuencia alélica entre diferentes clases de germoplasma, como se reveló en este estudio, mostró gran potencial para materiales heterogéneos, como poblaciones de mejoramiento. De acuerdo con esta afirmación, un estudio que puede plantearse es la evaluación de la diversidad genética dentro de cada población y la estimación de los cambios de frecuencia alélica a través de ciclos de selección, esto no sólo permitiría la identificación de *loci* de interés sino también su co-localización con otros *loci* relacionados, o la identificación de *loci* con efectos fenotípicos desconocidos hasta ahora. Específicamente, esta idea se empezó a madurar y se está desarrollando hoy día en poblaciones de mejoramiento para sequía aprovechando los datos disponibles de las múltiples evaluaciones en campo a las que han estado sujetas. Las poblaciones seleccionadas por su gran importancia en mejoramiento para sequía fueron: LaPosta-sequía, Tuxpeño-sequía, Poblaciones tolerantes a la sequía de maíces dentados blancos y amarillos (DTPW y DTPY, Drought Tolerant Population White y Drought Tolerant Population Yellow, por sus siglas en inglés) de los ciclos 0 y más avanzados. En este caso, la caracterización genotípica se efectuará en 250 individuos por población y con una plataforma de marcadores de alta densidad.

En este mismo orden de ideas, incluso, la evaluación de las frecuencias alélicas en materiales diversos podrían aportar más información relacionada con la evolución de caracteres de interés, aunque de una manera más ambiciosa, esta clase de conocimiento entre materiales nativos y materiales mejorados también podría generar esfuerzos dirigidos para la restauración de la diversidad en *loci* que han sufrido una fuerte presión de selección o introgresiones históricas, generando LD no deseable, y estas actividades representarían las nuevas aplicaciones potenciales del GWAS cuando se enfatice estudios de diversidad diversidad (Morrell et al., 2012).

## 5. REFERENCIAS

- Anghiera, P. (1907). *De Orbe Novo* (1st complete ed. 1530). Leroux, Paris.
- Anderson, E., y Cutler, H. (1942). Races of *Zea mays*: I. Their recognition and classification. *Annals of the Missouri Botanical Garden*, 29:69–88.
- Aranguren-Méndez, J.A., Román-Bravo, R., Isea, W., y Villasmil, J.J. (2004). Los microsatélites (STR's), marcadores moleculares de ADN por excelencia para programas de conservación: una revisión. *Archivos Latinoamericanos de Produccion Animal*, 13:1-14.
- Araus, J. L. (2002). Plant Breeding and Drought in C3 Cereals: What Should We Breed For?. *Annals of Botany*, 89:925-940.
- Araus, J.L., Slafer, G.A., Royo, C., and Serret M.D. (2008). Breeding for yield potential and stress adaptation in cereals. *Critical Reviews in Plant Science*, 27:1-36.
- Araus, J.L., Sanchez, C., y Edmeades, G.O. (2010). Maíz tropical: Fisiología del rendimiento y adaptación a sequía. En C. de Leon y R. Rodriguez (Eds). *El cultivo del maíz-Temas Selectos II* (pp. 61-76). México: Mundi-Prensa.
- Araus, J.L., Sanchez, C. y Edmeades, G.O. (2011). Phenotyping maize for adaptation to drought. In J-M. Ribaut and P. Monneveux (Eds), *Drought phenotyping in crops: from theory to practice* (pp. 263-283). Texcoco, Mexico: CGIAR Generation Challenge Programme.
- Bänziger, M., Edmeades, G.O., Beck, D., y Bellon, M. (2000). *Breeding for drought and nitrogen stress tolerance in maize: From theory to practice*. Mexico D.F.: CIMMYT.
- Barros, C., y Buenrostro, M. (1997). El maíz nuestro sustento, *Arqueología Mexicana*, 5:6-15.
- Barua, U.M., Chalmers, K.J., Hackett, C.A., Thomas, W.T., Powell, W., y Waugh, R. (1993). Identification of RAPD markers linked to a *Rhynchosporium secalis* resistance locus in barley using near-isogenic lines and bulked segregant analysis. *Heredity*, 71: 177-184.
- Beaumont, M.A., Ibrahim, K.M., Boursot, P., y Bruford, M.W. (1998). Measuring genetic distance. In A. Karp et al. (Eds.). *Molecular tools for screening biodiversity* (pp. 315–325). London: Chapman and Hall.

- 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.
- Bedoya, C.A, y Chávez, V. (2010). Teocintle: el ancestro del Maíz. *Claridades Agropecuarias*, 201:32-41.
- Belo, A., Zheng, S., Luck, S., Meyer, D.J, Li, B, Tingey, S. y Rafalsky A. (2008). Whole genome scan detects an allelic variant of *fad2* associated with increased oleic acid levels in maize. *Molecular Genetics and Genomics*, 279: 1-10.
- Benz, B.F (1997). Diversidad y distribución prehispánica del maíz mexicano. *Arqueología Mexicana*, 5(25):16-23.
- Bolaños, J., y Edmeades, G.O. (1996). The importance of the anthesis silking interval in breeding for drought tolerance in tropical maize. *Field Crops Research*, 48: 65-80.
- Blake, M. (2006). Dating the Initial Spread of *Zea mays*. In J. E. Staller, R. H. Tykot and B. F. Benz (Eds). *Histories of maize: multidisciplinary approaches to the prehistory, biogeography, domestication, and evolution of maize* (pp. 55-71). Amsterdam: Academic Press.
- Blum, A., Klueva, N., y Nguyen, H.T. (2001). Wheat cellular thermotolerance is related to yield under heat stress. *Euphytica*, 117:117-123.
- Blum, A. (2005). Drought resistance, water efficiency, and yield potential: are they compatible, dissonant, or mutually exclusive? *Australian Journal of Agricultural Research*, 56:1103-1115.
- Bolaños, J. y Edmeades, G.O. (1996). The importance of the anthesis-silking interval in breeding for drought tolerance in tropical maize. *Field Crop Research*, 48:65-80.
- Bracco, M., Lia, V. V., Gottlieb, a M., Cámara Hernández, J., y Poggio, L. (2009). Genetic diversity in maize landraces from indigenous settlements of Northeastern Argentina. *Genetics*, 135:39-49.
- Bradbury, P.J., Zhang, Z., Kroon, D.E., Casstevens, T.M., Ramdoss, Y., y Buckler, E.S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics*, 23:2633–2635.
- Breseghello, F., and M.E. Sorrells. (2006). Association analysis as a strategy for improvement of quantitative traits in plants. *Crop Science*, 46:1323–1330.
- Buckler, E.S., Gaut, B.S., y McMullen, M.D. (2006) Molecular and functional diversity of maize. *Current Opinions in Plant Biology*, 9:172-176.

- Campos, H., Cooper, M., Habben, J.E., Edmeades, G.O., y. Schussler, J.R. (2004). Improving drought tolerance in maize: a view from industry. *Field Crops Research*, 90:19–34.
- Camus-Kulandaivelu, L., Veyrieras, J.-B., Madur, D., Combes, V., Fourmann, M., Barraud, S., Dubreuil, P., et al. (2006). Maize adaptation to temperate climate: relationship between population structure and polymorphism in the Dwarf8 gene. *Genetics*, 172:2449-63.
- Cardona, J.O. (2010). Análisis de diversidad genética de las razas colombianas de maíz a partir de datos Roberts et al. (1957) usando la estrategia Ward-MLM. *CienciAgro*, 2:199-207.
- Cardoza-Arangón, L. (2006). Corazón de maíz. Mitos del maíz. *Artes de México*, 79:11-14.
- Castillo, N. (2009). El maíz y la arqueología. Desgranado una mazorca, orígenes y etnografía de los maíces nativos. En *Diario de Campo*, 52:40-65. México, D.F.: Publicación interna de la Coordinación Nacional de Antropología del Instituto Nacional de Antropología e Historia.
- Cavalli-Sforza, L.L., Menozzi, P., y Piazza, A. (1994). The history and geography of Human Genes. Princenton, New Jersey: Princeton University Press.
- Ceron-Rojas, J., y Sahaún-Castellanos, J. (2005). Un índice de selección basado en Componentes Principales. *Agrociencia*, 39:667-377.
- Ceron-Rojas, J., Crossa, J., Sahaún-Castellanos, J., Castillo-Gonzales, F. y Santacruz-Varela, A. (2006). A Selection Index Method Based on Eigenanalysis. *Crop Science*, 46:1711-1721.
- Cheng, H.H. y Crittenden, L.B. (1994). Microsatellite markers for genetic mapping in the chicken. *Poultry Science Association*. 73:539-546.
- Ciofi, C., Funk, S., Coote, T., Cheesman, D., Hammond, R., Saccheri, I. y Bruford, W. (1988). Genotyping with microsatellite markers. In A. Karp, PG Isaac, and DS Ingram (Eds.) *Molecular tools for screening biodiversity: plants and animals* (pp.195-201). London: Chapman and Hall.
- Collins, N.C., François, T. y Tuberosa, R. (2008). Quantitative trait loci and crop performance under abiotic stress: where do we stand? *Plant Physiology*. 147:469–486.
- Crow, J. F. y Kimura, M. (1970). *An introduction to population genetics theory*. New York: Evanston and London. Harper and Row Publishers.

- Dekkers, J.C.M. y Hospital, F. (2002). The use of Molecular Genetics in the Improvement of Agricultural Populations. *Nature*, 3:22-32.
- Delannay X., McLaren G. y Ribaut J.-M. (2012). Fostering molecular breeding in developing countries. *Molecular Breeding*, 29:857-873.
- DIVERSITY ARRAYS TECHNOLOGY (2011). Diversity Arrays Technology Pty Ltd (DART P/L). <http://www.diversityarrays.com/>, citado Enero 18 de 2012.
- Doebley, J.F., Goodman, M.M. y Stuber C.W. (1984). Isoenzymatic variation in *Zea* (Gramineae). *Systematic Botany*, 9:203-218.
- Doebley, J.F. (2004). The genetics of maize evolution. *Annual Reviews Genetics*, 38:37-59
- Dowswell, C., Paliwal, R. y Cantrell, R. (1996). *Maize in the third world*. Boulder, Colorado, USA: Westview Press.
- Dubreuil, P., Warburton, M., Chastanet, M., Hoisington, D., y Charcosset, A. (2006). More on the Introduction of Temperate Maize into Europe: Large-Scale Bulk SSR Gnotyping and New Historical Elements. *Maydica*, 51:281-291.
- Duvick, D.N. (1990). Genetic enhancement and plant breeding. In: J. Janick and J.E. Simon (Eds.). *Advances in new crops*, Portland, Oregon: Timber Press.
- Eding, H. y Laval, G. (1999). Measuring genetic uniqueness in livestock. In: Oldenbroek, K. (Ed.). *Genebanks and the management of farm animal genetic resources* (pp. 33-58). The Netherlands: IDO-DL press.
- Elshire, R.J., Glaubitz, J.C., Sun, Q., Poland, J.A., Kawamoto, K., Buckler, E.S., y Mitchell, S.E. (2011) A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. *Public Library of Science ONE*, 6:e19379.
- Excoffier, L., Smouse, P.E., y Quattro, J.M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, 131:479-91.
- Evanno, G., Regnaut, S., y Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular ecology*, 14:2611-20.
- Falconer, D.S. (1981). *Introduction to quantitative genetics*. London: Longman.
- Fischer, R.A. y Maurer, R. (1978). Drought resistance in spring wheat cultivars. I. Grain yield responses. *Australian Journal of Agricultural Research*, 29: 897-912.

- Flint-Garcia, S. A., Thuillet, A-C., Yu, J., Pressoir, G., Romero, S. M., Mitchell, S. E., Doebley, J., Kresovich, S., Goodman, M.M., and Buckler E.S. (2005). Maize association population: a high-resolution platform for quantitative trait locus dissection. *The Plant Journal*, 44: 1054-1064.
- FONTAGRO. (2008). Fondo Regional de Tecnología Agropecuaria. Proyecto: Generación y validación de variedades de maíz tolerantes a sequía como medio de estabilizar la productividad y disminuir el daño por micotoxinas como consecuencia del cambio climático. [http://www.fontagro.org/sites/default/files/prop\\_08\\_28.pdf](http://www.fontagro.org/sites/default/files/prop_08_28.pdf), citado 19 de Febrero de 2012.
- Fooland, M.R., y Jones, R.A. (1993). Mapping salt-tolerance genes in tomato. (*Lycopersicon esculentum*) using trait-based marker analysis. *Theoretical and Applied Genetics*, 87:184-192.
- Franco, J., Warburton, M., Dubreuil, P., y Dreisigacker, S. (2005). *User's manual for the FREQS-R Program for estimating allele frequencies for fingerprinting and genetic diversity studies using bulked heterogeneous populations*. Mexico, D.F.: CIMMYT.
- Franco, J., Warburton, M. y Dreisigacker, S. (2007). *User's manual for the FtoL-R Program for generating a "dummy" data set consisting of allele lengths for hypothetical individuals*. Mexico, D.F.: CIMMYT.
- Frankel, O.H. and Brown, A.H.D. (1984). *Plant genetic resources today: A critical appraisal*. In J.H.W. Holden and J.T. Williams (Eds.). *Crop genetic resources: Conservation and evaluation* (pp. 249-257). London: George Allen and Unwin Press.
- Fukunaga, K., Hill, J., Vigouroux, Y., Matsuoka, Y., Sanchez G, J., Liu, K., Buckler, E. S., y Doebley, J. (2005). Genetic diversity and population structure of teosinte. *Genetics*, 169: 2241-54.
- Gallais, A., Moreau, L., y Charcosset, A. (2007). Detection of marker-QTL associations by studying change in marker frequencies with selection. *Theoretical and Applied Genetics*, 114:669-81.
- González, A. (2009). El maíz como producto cultural desde tiempos antiguos. En *Diario de Campo- Desgranado una mazorca, orígenes y etnografía de los maíces nativos*. (52,40-65). México, D.F.: Publicación interna de la Coordinación Nacional de Antropología del Instituto Nacional de Antropología e Historia.
- Goodman, M.M, y Bird, R.McK. (1977). The races of Maize IV: tentative grouping of 219 Latin American Races. *Economy Botanic*. 31:204–221



- Goodman, M.M., and Brown, W.L. (1988). Races of corn. p. 33–79. In G.F. Sprague and J.W. Dudley (Eds.) *Corn and corn improvement*. 3rd ed. Madison, Wisconsin: Agronomy Monographs 18. ASA, CSSA, and SSSA.
- Hair, J.R., Anderson, R.E., Tatham, R.L., and Black, W.C. (1995). *Multi-variate data analysis with readings*. 4th ed. Englewood Cliffs, New Jersey: Prentice Hall.
- Hansey, C. N., Johnson, J. M., Sekhon, R. S., Kaeppler, S. M., y Leon, N. D. (2011). Genetic Diversity of a Maize Association Population with Restricted Phenology. *Crop Science*, 51:704-715.
- Heisey, P.W. y Morris, M.L. (2006) Economic Impact of water-limited conditions on cereal grain production. In J-M. Ribaut (Ed), *Drought Adaptation in Cereals* (pp. 259-299). New York: The Haworth Press
- Harjes, C.E, Rocheford, T.R., Bai, L., Brutnell, T.P., Bermudez, C., Sowwinski, S.G., Stapleton, A. E., Vallabhaneni, R., Williams, M., Wurtzel, E.T., Yan, J., y Buckler, E.S. (2008). Natural Genetic Variation in *Lycopene Epsilon Cyclase* Tapped for Maize Biofortification. *Science*, 319:330-333.
- Hastorf, C.A, Whiteheat, W.T., Bruno, M.C., y Wright, M. (2006). The Movements of Maize into Middle Horizon Tiwanaco, Bolivia. In J. E. Staller, R. H. Tykot and B. F. Benz (Eds). *Histories of maize: multidisciplinary approaches to the prehistory, biogeography, domestication, and evolution of maize* (pp. 429-443). Amsterdam: Academic Press.
- Hernández, X. (1987). Razas de maíz en México: Su origen, características y distribución. *Geografía Agrícola*; 2:609-732.
- Johnson, A.R., y Wichern, D.W. (1992). *Applied multivariate statistical analysis*. 3rd ed. Englewood Cliffs, New Jersey: Prentice-Hall,
- Joshi, S.P., Ranjekar, P.K., y Gupta, V.S. (1999). Molecular markers in plant genome analysis. *Current Science*, 77:230-240.
- Kato, T.A. (2005). Cómo y dónde se originó el maíz, *Investigación y Ciencia*, 34:68-72.
- Kato, T.A., Mapes, C., Mera, L.M., Serratos, J.A. y Bye, R.A. (2009). *Origen y diversificación del maíz: una revisión analítica*. México, D.F.: Universidad Nacional Autónoma de México, Comisión para el Conocimiento y Uso de la Biodiversidad.
- Lanaud, C. y Lebot, V. (1995). Molecular techniques for increased use of genetic resources. In: W.G. Ayad, T. Hodgkin, A. Jaradat, y V.R. Rao (Eds.). *Molecular genetic*

- techniques for plant genetic resources. Report of an IPGRI workshop, October 9-11 1995, Rome, Italy. International Plant Genetic Resources Institute (pp 92-97).*
- Lee, M. (1998). Genome projects and gene pools: new germplasm for plant breeding? *Proceedings of the National Academy of Sciences of the United States of America*, 95:2001-2014.
- Levitt, J. (1972). *Responses of plants to environmental stresses*. New York: Academy Press.
- Lebowitz, R. J., Soller, M., y Beckmann, J. S. (1987). Trait-based analyses for the detection of linkage between marker loci and quantitative trait loci in crosses between inbred lines. *Theoretical and Applied Genetics*, 73:556-562.
- Lia, V.V, Poggio, L., y Confalonieri, V.A. (2009). Microsatellite variation in maize landraces from Northwestern Argentina: genetic diversity, population structure and racial affiliations. *Theoretical and Applied Genetics*, 119:1053-1067.
- Liu, K., Goodman, M., Muse, S., Smith, J. S., Buckler, E., y Doebley, J. (2003). Genetic structure and diversity among maize inbred lines as inferred from DNA microsatellites. *Genetics*, 165:2117-28.
- Liu, K., y Muse, S. V. (2005). PowerMarker: an integrated analysis environment for genetic marker analysis. *Bioinformatics*, 21:2128-9.
- Lu, Y., Zhang, S., Shah, T., Xie, C., Hao, Z., Li, X., Farkhari, M., Ribaut, J-M., Cao, M., Rong, T., y Xu, Y. (2011). Joint linkage-linkage disequilibrium mapping is a powerful approach to detecting quantitative trait loci underlying drought tolerance in maize. *Proceedings of the National Academy of Sciences of the United States of America*, 107:19585-19590.
- McClung de Tapia, E. (1997). La domesticación del maíz. El Maíz. *Arqueología Mexicana*, 5:34-39.
- McClintock, B., Kato, Y., y Blumenschein, A. (1981). *Constitución Cromosómica de las Razas de Maíz*. Texcoco, México: Colegio de Postgraduados, Chapingo,.
- McCouch, S.R., Zhao, K., Wright, M., Tung, C-W., Ebana, K., Thomson, M., Reynolds, A., Wang, D., DeClerck, G., Ali, M.L., McClung, A., Eizenga, G., y Bustamante, C. (2010). Development of genome-wide SNP assays for rice. *Breeding Science*, 60:524-535.
- Manrique, P.A. (1997). *El maíz en Perú*. Lima: Consejo Nacional de Ciencia y Tecnología – CONCYTEC.

- Matsuoka, Y., Mitchell, S.E., Kresovich, S., Goodman, M., y Doebley, J. (2002a). Microsatellites in *Zea* - variability, patterns of mutations, and use for evolutionary studies. *Theoretical and Applied Genetics*, 104:436-450.
- Matsuoka, Y., Vigouroux, Y., Goodman, M. M., Sanchez G, J., Buckler, E., y Doebley, J. (2002b). A single domestication for maize shown by multilocus microsatellite genotyping. *Proceedings of the National Academy of Sciences of the United States of America*, 99:6080-4.
- Melgar-Tísoc, E (2000). Navegación prehispánica y la ruta del maíz *Amerística*, Año 3, 5:101-127.
- Melchinger, A.E., Utz, H.F., y Schön, C.C. (2004). QTL analyses of complex traits with cross validation, bootstrapping and other biometric methods. *Euphytica*, 137:1-11.
- Mohammadi, S.A. y Prasanna, B.M. (2003). Analysis of genetic diversity in crop plants – Salient Statistical tools and considerations. *Crop Science*. 43:1235-1248.
- Monneveux, P., y Ribaut, J-M. (2006). Secondary Traits for Drought Tolerance Improvement in Cereals. In J-M. Ribaut (Ed). *Drought Adaptation in Cereals* (pp. 97-143). New York: The Haworth Press.
- Mullis, K. B., Faloona, F., Scharf, S., Saiki, R., Horn, G. y Erlich, H. (1986). Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harbor Symposia on Quantitative Biology*. 51:263-273.
- Muñoz, A., Cuevas, J.A., Miranda, S., Molina, J., y Santacruz, A. (2003). Las grandes comarcas culturales y la diversidad de razas de maíz. En A. Muñoz (Ed), *Centli-Maíz: Prehistoria e Historia, Diversidad, Potencial, Origen Genético y Geográfico, Glosario Centli-Maíz* (pp. 19-31). México, Texcoco: Colegio de Postgraduados.
- Moose, S.P., y Mumm, R.H. (2008): Molecular plant breeding as the foundation for 21st century crop improvement. *Plant Physiology*, 147:969-977.
- MUSEO CHILENO DE ARTE PRECOLOMBINO (2011) Culturas precolombinas <http://www.precolombino.cl/>, citado Noviembre de 2011.
- Nagamine, Y. y Higuchi, M. (2001). Genetic distance and classification of domestic animals using genetic markers. *Journal of Animal Breeding and Genetics*. 118:101-109.
- Nei, M. (1973). Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences of the United States of America*, 70:3321–3323.

- Newson, L.A. (2006). Caribbean Maize, First Farmers to Columbus. In J. E. Staller, R. H. Tykot and B. F. Benz (Eds). *Histories of maize: multidisciplinary approaches to the prehistory, biogeography, domestication, and evolution of maize* (pp. 325-326). Academic Press, Amsterdam.
- Oliveira-Freitas, F., Bendel, G., Allaby, R.G., y Brown, T.A. (2003). DNA from primitive maize landraces and archaeological remains: implications for the domestication of maize and its expansion into South America. *Journal of Archaeological Science*, 30:901–908
- Oraguzie, N.C. y Wilcox, P.L. (2007). An overview of association mapping. In N.C. Oraguzie, E.H.A. Rikkerink, S.E. Garduner, and H.N. De Silva (Eds). *Association mapping in plants* (pp.1-9). New York: Springer.
- Ortega-Paczka, R. (2003). La diversidad del maíz en México. En C. Esteva, G., Marielle. (Eds.), *Sin Maiz no hay País* (pp. 123-154). México, D.F.: Consejo Nacional para la cultura y las artes, Dirección General de Culturas Populares e Indígenas.
- Pandey, S. y Gardner, C.O. (1992). Recurrent selection for population, variety, and hybrid improvement in tropical maize. *Advances in Agronomy*, 48:2-79.
- Paterniani, E. y Goodman, M.M. (1977). Races of Maize in Brazil and Adjacent Areas. México, D.F.: Centro Internacional de Mejoramiento de Maiz y Trigo.
- Patterson, N., Price, A.L., y Reich, D. (2007). Population structure and eigenanalysis. *Public Library of Science Genetics*, 2:2074-2093e90.
- Perrier, X., y Jacquemoud-Collet, J. (2006). DARwin software. Available online at: <http://Darwin.cirad.fr/darwin>
- Piperno, D. y Pearsall, D. (1998). *The origins of agriculture in the lowland Neotropics*. San Diego, CA: Academic Press.
- Poland, J.A., Bradbury, P.J., Buckler, E.S., y Nelson, R.J. (2011). Genome-wide nested association mapping of quantitative resistance to northern leaf blight in maize. *Proceedings of the National Academy of Sciences of the United States of America*. 108:6893–6898.
- Pressoir, G., y Berthaud, J. (2004). Patterns of population structure in maize landraces from the Central Valleys of Oaxaca in Mexico. *Heredity*, 92:88-94.
- Pritchard, J. K., Stephens, M., y Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155:945-59

- Pritsch, C., (2001). El pre-mejoramiento y la utilización de los recursos fitogenéticos 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/~fitotecnia/docencia/materiales%20apoyo/Premejoramiento.pdf>
- Pye, M.E. y Deamrest, A.A. (1993). El Mesak y los inicios de culturas complejas en Guatemala. En J.P. Laporte, H. Escobedo y S. Villagrán (Eds), *III Simposio de Investigaciones Arqueológicas en Guatemala, 1989* (pp.225-238). Guatemala. Museo Nacional de Arqueología y Etnología.
- Ramirez, R., Timothy, E.D., Diaz, E., y Grant, U.J. (1961). *Razas de Maíz en Bolivia*. Bogotá, República de Colombia: Ministerio de Agricultura, Oficina de Investigaciones Especiales.
- Rebourg, C., Chastanet, M., Gouesnard, B., Welcker, C., Dubreuil, P., y Charcosset, A. (2003). Maize introduction into Europe: the history reviewed in the light of molecular data. *Theoretical and Applied Genetics, 106*:895-903.
- Reif, J.C., Xia, X.C., Melchinger, A., Warburton, M.L., Hoisington, D, Beck, D., Bohn y Frish, M.M. (2004). Genetic Diversity determined within and among CIMMYT maize population of tropical, subtropical and temperate germplasm by SSR markers. *Crop Science. 44*:326-334.
- Reif, J. C., Warburton, M.L., Xia, X.C., Hoisington, D., Crossa, J, Taba, S., Muminović, J., Bohn, M., Frisch, M., y Melchinger, A.E. (2006). Grouping of accessions of Mexican races of maize revisited with SSR markers. *Theoretical and Applied Genetics, 113*:177-85.
- Reyna, R.M., y Alvarez, C. (2009). El maíz arqueológico, evidencia ancestral de las razas de maíz contemporáneas: el caso de La Organela Xochipala, Guerrero. En *Diario de Campo. Desgranado una mazorca, orígenes y etnografía de los maíces nativos* (52, 28-39). México, D.F.: Publicación interna de la Coordinación Nacional de Antropología del Instituto Nacional de Antropología e Historia.
- Ribaut, J.-M., Gonzalez de Leon, D., Huerta, E., Deutsch, J., y Hoisington, D. (1996). Mapping of factors involved in the expression of anthesis silking interval (ASI) in tropical maize under drought. In *Integrated Study of Drought Tolerance of Higher Plants. Proceedings of the First International Interdrought Meeting*, Montpellier, France.
- Ribaut, J.-M., y Hoisington, D. (1998). Marker-assisted selection new tools and strategies. *Trends Plant Science, 3*:236-239

- Ribaut, J.-M., Bänziger, M., Betran, J., Jiang, C., Edmeades, G.O., Dreher, K., y Hoisington, D. (2002). Use of molecular markers in plant breeding: Drought tolerance improvement in tropical maize. In: S. Manjit and S. Kang (Eds.). *Quantitative genetics, genomics, and plant breeding* (pp. 85-99). Wallingford, UK: CABI.
- Ribaut, J.-M., Hoisington, D., Banziger, M., Setter, T.L., y Edmeades, G.O. (2004) Genetic dissection of drought tolerance in maize: A case study. In: H.T. Nguyen, A. Blum (Eds.). *Physiology and Biotechnology Integration for Plant Breeding* (pp. 571-611). New York: Marcel Dekker Inc.
- Ribaut, J.-M., Betran, J., Monneveux, P., y Setter, T. (2008). Drought tolerance in maize. In J. L. Bennetzen and S. C. Hake (Eds.) *Handbook of Maize* (pp. 311-344). New York: Springer New York.
- Ribaut, J.-M., de Vicente, M.C. y Delannay, X. (2010). Molecular breeding in developing countries: challenges and perspectives. *Current Opinion in Plant Biology*, 13:1-6.
- Risch, N., y Merikangas, K. (1996). The future of genetic studies of complex human diseases. *Science*, 273:1516-1517.
- Roberts, L.M, Grant, U.J., Ramirez, R., Hatheway, W.H. y Smith, D.H., en colaboración con P.C. Mangelsdorf. (1957). *Razas de Maíz en Colombia*. Bogotá. República de Colombia, Ministerio de Agricultura, Oficina de Investigaciones Especiales.
- Rohlf, F.J. (1972). An empirical comparison of three ordination techniques in numerical taxonomy. *Systematic Zoology*. 21:271-280.
- Rodríguez, M., y Arencibia, A. (2002). Principales tipos de marcadores del polimorfismo de los ácidos nucleídos. En: M.T. Cornide-Hernandez (Ed.). *Marcadores Moleculares: Nuevos horizontes en la genética y la selección de plantas* (pp. 13-35). La Habana: Félix Varela Ediciones.
- Ron, J., Sánchez, J.J., Jimenez, A.A., Carrera, J.A, Martín, J.G, Morales, M.M, Larios, L., Hurtado de la Peña, S.A., Mena, S., y Rodríguez, J.G. (2006). *Maíces nativos del Occidente de México I. Coletas 2004*. México: Universidad de Guadalajara, Edición del Centro Universitario de Ciencias Biológicas y Agropecuarias.
- Schlötterer, C. (2004). The evolution of molecular markers-just a matter of fashion? *Nature Reviews Genetics*, 5:63-69.
- Sahagún-Castellanos, J. J. y C.-Rojas J. (2005). A Selection index based on Principal Components, *Agrociencia*, 39:667-677.
- Saitou, N., and Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, 4: 406-425.

- Salhuana, W. y Machado, V. (1999). *Races of maize in Paraguay: considerations in Organization and Utilization of Maize Genetic Resources*. Publication 025, United States: Department of Agriculture, Agricultural Research Service and The Maize Program of the Paraguayan Ministry of Agriculture and Livestock.
- 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.
- Salvi, S. y Tuberosa, R. (2005). To clone or not to clone plant QTLs: present and future challenges. *Trends in Plant Science*, 10:297-304.
- Sánchez, J.J, Goodman, M. M., y Stuber, C.W. (2000) Isozymatic and morphological diversity in the races of maize of Mexico. *Economic Botany*, 54:43-59.
- Sánchez, J., Goodman, M., Bird, R., y Stuber, C. (2006). Isozyme and morphological variation in maize of five Andean countries. *Maydica*, 51: 25-42.
- Sawkins, M.C., DeMeyer, J., y Ribaut, J-M. (2006) Drought Adaptation in Maize. In J-M. Ribaut (Ed), *Drought Adaptation in Cereals* (pp. 259-299). New York: The Haworth Press
- Serratos, A., Arnason, J.T., Nozzolillo, C., Lambert, J.D.H., Philogène B. J. R., Fulcher, G., Davidson, K., Peacock, L., Atkinson, J., y Morand, P. (1987). Factors contributing to resistance of exotic maize populations to maize weevil, *Sitophilus zeamais*. *Journal of Chemical Ecology*, 13:751-762.
- Setter T.L., Yan J., Warburton, M., Ribaut, J-M., Xu Y., Sawkins M., Buckler, E.S., Zhang Z. and Gore, M.A (2011). Genetic association mapping identifies single nucleotide polymorphisms in genes that affect abscisic acid levels in maize floral tissues during drought. *Journal of Experimental Botany*, 62:701-716.
- Smalley, J., y Blake, M. (2003). Sweet Beginnings: Stalk Sugar and the Domestication of Maize. *Current Anthropology*, 44:675-703.
- Smith, H.F. (1936). A discriminant function for plant selection. *Annals Eugenics*, 7, 240-250.
- Staller, J.E, y Thompson, R.G (2002). A multidisciplinary approach to understanding the initial introduction of maize into coastal Ecuador. *Journal of Archaeological Science*. 29:33-50.
- Stuber, C. W., Moll, R. H., Goodman, M M, Schaffer, H. E., y Weir, B. S. (1980). Allozyme Frequency Changes Associated with Selection for Increased Grain Yield in Maize (*Zea Mays L.*). *Genetics*, 95:225-36.

- Taba, S., (Ed.). (2005). Latin American Maize Germplasm Conservation: Regeneration, in situ, Conservation, Core Subsets, and Prebreeding; *Proceedings of a Workshop held at CIMMYT*, April 7-10, 2003. Mexico, D.F.: CIMMYT.
- Takezaki, N., and Nei, M. (1996). Genetic distance and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics*. 144: 389-399.
- Thornsberry, J. M., Goodman, M M, Doebley, J, Kresovich, S, Nielsen, D., y Buckler, E S. (2001). Dwarf8 polymorphisms associate with variation in flowering time. *Nature genetics*, 28:286-289.
- Timothy, D.H., Peña, B., y Ramirez, R., en colaboración con W.L. Brown, y E. Anderson. (1961). *Razas de maíz en Chile*. Publication 847. Washington, D.C.: National Academy of Science National Research Council.
- Torres, C.M. (1988). *Taino: Los descubridores de Colón*. Santiago de Chile: Museo de Arte Precolombino de Chile.
- Tuberosa, R (2011). Phenotyping drought-stressed crops: key concepts, issues and approaches. In J-M. Ribaut and P. Monneveux (Eds), *Drought phenotyping in crops: from theory to practice* (pp. 3-35). Texcoco, México: CGIAR Generation Challenge Programme.
- Varshney, R.K., Ribaut, J-M., Buckler, E.S., Tuberosa, R., Rafalski, J.A., y Langridge, P. (2012). Bajo revisión.
- Vela, E. (2011). El maíz – Catálogo Visual: Historia, Simbolismo, Botánica y Gastronomía. *Arqueología Mexicana*. Especial 38:10-26.
- Vigourox, Y., Jaqueth, J., Matsuoka, Y., Smith, O., Beavis, W., Smith, S. y Doebley, J. (2002). Rate and pattern of mutation at microsatellite loci in maize. *Molecular Biology Evolution*. 19: 1251-1260.
- Vigouroux, Y., Glaubitz, J. C., Matsuoka, Y., Goodman, M.M., Sánchez, G.J., y Doebley, J. (2008). Population structure and genetic diversity of New World maize races assessed by DNA microsatellites. *American Journal of Botany*, 95:1240-53.
- Wang, S., Lewis, C. M., Jakobsson, M., Ramachandran, S., Ray, N., Bedoya, G., Rojas, W., Parra, M.V., Molina, J.A., Gallo, C., Mazzotti, G., Poletti, G., Hill, K., Hurtado, A.M., Labuda, D., Klitz, W., Barrantes, R., Bortolini, A.C., Salzano, F.M., Petzl-Erler, M.L., Tsuneto, L.T., Llop, E., Rothhammer, F., Excoffier, L., Feldman, M.W., Rosenberg, N.A., y Ruiz-Linares, A. (2007). Genetic variation and population structure in Native Americans. *Public Library of Science Genetics*, 3:2049-2067.



- Wang, R.H., Yu, Y.T, Zhao, J.R., Shi, Y.S., Song, Y.C., Wang, T.Y., y Li, Y. (2008). Population structure and linkage disequilibrium of a mini core set of maize inbred lines in China. *Theoretical and Applied Genetics*, 117:1141-1153.
- Warburton, M.L., Wilkes, G., Taba, S., Charcosset, A., Mir, C., Dumas, F., Madur, D., Dreisigacker, S., Bedoya, C., Prasanna, B.M., Xie, X., Herane, S., y Franco, J. (2011). Gene flow among different teosinte taxa and into the domesticated maize gene pool. *Genetic Resources and Crop Evolution*, 58:1243-1261.
- Warburton, M.L., Xianchun, X., Crossa, J., Franco, J., Melchinger, A.E., Frisch, M., Bohn, M., y Hoisington, D. (2002). Genetic Characterization of CIMMYT Inbred Maize Lines and Open Pollinated Populations Using Large Scale Fingerprinting Methods, *Crop Science*, 42:1832-1840.
- Warburton, M. L., Reif, J. C., Frisch, M., Bohn, M., Bedoya, C., Xia, X. C., Crossa, J., Franco, J., Hoisington, D., Pixley, K., Taba, S., y Melchinger, A.E (2008). Genetic Diversity in CIMMYT Nontemperate Maize Germplasm: Landraces, Open Pollinated Varieties, and Inbred Lines. *Crop Science*, 48, 617-624.
- Wellhausen, E.J., Roberts, L.M, y Hernández X., en colaboración con P.C. Mangelsdorf. (1951). *Razas de maíz en México, su origen, características y distribución*. Folleto Técnico no.5, México, D.F.: Oficina de Estudios Especiales, Secretaria de Agricultura y Ganadería.
- Wellhausen, E.J., Fuentes, A., Hernandez, A. in collaboration with P.C. Mangelsdorf. (1957). *Races of Maize in Central America*. Publication 511, Washington, D.C.: National Academy of Sciences-National Research Council.
- Wiley, E.O. (1981). *Phylogenetics: The theory and practice of phylogenetics and systematics*. New York: John Wiley and Sons.
- Wilson, L. M., Whitt, S. R., Ibañez, A. M., Rocheford, T.R., Goodman, M.M. y Buckler, E.S. (2004). Dissection of Maize Kernel Composition and Starch Production by Candidate Gene Association. *The Plant Cell*, 16:2719-2733.
- WORLD BANK, (2007). World development report 2008. Agriculture for development. Washington DC, USA.  
[http://siteresources.worldbank.org/INTWDR2008/Resources/WDR\\_00\\_book.pdf](http://siteresources.worldbank.org/INTWDR2008/Resources/WDR_00_book.pdf)
- Wright, S. (1951). The genetical structure of populations. *Annals. Eugenics*, 15: 323-54.
- Wright, S. (1965). The interpretation of population structure by F-statistics with special regard to system of mating. *Evolution* 19: 395-420.

- Xu, S. (2003). Advanced Statistical methods for estimating genetic variances in plants. *Plant Breeding Reviews*, 113-163.
- Xu, Y., Wang, J., y Crouch, J. (2008). Selective Genotyping and Pooled DNA Analysis : An Innovative Use of an Old Concept. Proceedings of the 5<sup>th</sup> International Crop Science Congress. Publish in CDROM.
- Yang, X., Li, Q., Xu, S., Li, L., Yan, J., Gao, Y., Bai, G., Shah, T., Chai, Y., Meng, Y., Warburton, M. L., Fu, Z., Zhou, Y., Zheng, Y., y Li, J. (2010). Genetic analysis and characterization of a new maize association mapping panel for quantitative trait loci dissection. *Theoretical and Applied Genetics*, 121:417-31.
- Yu, J., y E.S. Buckler. (2006). Genetic association mapping and genome organization of maize. *Current Opinion Biotechnology*, 17:155–160.
- Yu, J., Holland, J.B., McMullen, M.D., y Buckler, E.S. (2008). Genetic design and statistical power of nested association mapping in maize, *Genetics* 178:539–551.
- Zapata, C. (1987). La variabilidad genética de las poblaciones. En: Espinosa de los Monteros, J., Labarta, U. (Eds.). *Genética de la acuicultura* (pp. 33-57). Madrid: Caicyt-Feuga.
- Zhang, Z., Buckler, E.S., Casstevens, T.M., y Bradbury, P.J. (2009). Software engineering the mixed model for genome-wide association studies on large samples. *Briefings in Bioinformatics*, 10:664–675.
- Zhu, C., Gore, M., Buckler, Edward S., y Yu, J. (2008). Status and Prospects of Association Mapping in Plants. *The Plant Genome Journal*, 1:-20.

## 7. ANEXOS

### 7.1 Datos de pasaporte para las 194 poblaciones de maíz lationamericanas

ID	Accesión	Raza	Pais de Origen	Altura de Colecta (msnm)	Latitud	Longitud	Tipo de grano
1	<i>ANTI GP2</i>	Cristalino Costeño Tropical	Antigua	40	17.1	-61.77	Dentadooado>Cristalino
2	<i>ARGE 486</i>	Cuarento Cateto	Argentina	300	-34.9	-57.95	Cristalino
3	<i>ARGE 564</i>	Cateto Amarillo	Argentina	1700	-26.05	-65.83	Cristalino
4	<i>ARGE GP8</i>	Cristal Sulino	Argentina	1554	-29	-62	Cristalino
5	<i>PPS920</i>		Argentina	81	-34.8	-61.667	Cristalino
6	<i>ARGENT 24</i>	Canario de Ocho	Argentina	117	-36.156	-63.506	Cristalino
7	<i>ARGENT 306</i>	Cateto	Argentina	1006	-34.5	-68.5	Cristalino
8	<i>ARGENT TUC43</i>	Culli	Argentina	2500	-23.2	-65.35	Dentado
9	<i>ARZM 03 013</i>	Camelia	Argentina	62	-30.82	-57.98	Cristalino
10	<i>ARGENT TUC16</i>	Capia Blanca	Argentina	2500	-23.2	-65.35	Harinoso
11	<i>BARB 5</i>	Cristalino Costeño Tropical	Barbados	100	13.26	-59.61	Cristalino
12	<i>BARB GP2</i>	Tusón	Barbados	70	13.12	-59.6	Dentado
13	<i>BARB GP1</i>	Cristalino Costeño Tropical	Barbados	100	13.18	-59.58	Cristalino>Dentado
14	<i>BOLI 905</i>	Altiplano	Bolivia	1976	-20	-64.2	Harinoso
15	<i>BOLI 929</i>	Perola	Bolivia	1828	-17.8	-63.17	Dentado
16	<i>BOZMO 214</i>	Uchuquilla	Bolivia	2700	-17.6	-66.03	Cristalino>Harinoso
17	<i>BOZM 1609</i>	Blando Amazónico	Bolivia	537	-18.139	-60.019	
18	<i>BOLI 711</i>	Perola	Bolivia	475	-21.52	-64.75	Cristalino
19	<i>BOLIVI 957</i>	Chusilla, Chuspillu	Bolivia	1947	-20	-64.2	Dentado
20	<i>BOLIVI 1037</i>	Coroico	Bolivia	156	-11.33	-67.67	Cristalino>Harinoso
21	<i>BOLIVI 928</i>	Checchi	Bolivia	1976	-20	-64.2	Harinoso
22	<i>BOLIVI 968</i>	Aysuma	Bolivia	1947	-20	-64.2	Dentado>Harinoso
23	<i>BOZM 1791</i>	Hualtaco	Bolivia	242	-16.992	-65.152	
24	<i>BOZM 988</i>	Hualtaco Colorado	Bolivia	2900	-21.47	-65.35	Harinoso
25	<i>BOLIVI 351</i>	Pojoso Chico, Ecuatoriana	Bolivia	1050	-19.65	-63.68	Harinoso
26	<i>BOZM 1681</i>	Duro Amazónico	Bolivia	418	-17.775	-63.194	
27	<i>BOLIVI 90</i>	Acre Interlock	Bolivia	500	-17.85	-60.783	Harinoso

ID	Accesión	Raza	Pais de Origen	Altura de Colecta (msnm)	Latitud	Longitud	Tipo de grano
28	<b>BOZM 694</b>	Uchuquilla	Bolivia	2600	-21.87	-64.15	Dentado>Cristalino
29	<b>BOZM 677</b>	Paru	Bolivia	3000	-17.083	-66.85	Harinoso
30	<b>BOZM 96</b>	Kulli	Bolivia	2150	-18.18	-65	Harinoso
31	<b>BRAZIL 2441</b>	Cateto Paulista Grosso (Cateto Grosso)	Brasil	1081	-15.78	-47.936	Cristalino
32	<b>BRAZIL 1546</b>	Cateto Assis Brasil	Brasil	274	-29.333	-54.167	Cristalino>Dentado
33	<b>BRAZIL 2305</b>	Dente Paulista	Brasil	770	-23.33	-46.38	Dentado
34	<b>BRAZIL 953</b>	Dente Riograndense Rugoso	Brasil	274	-28.183	-54.733	Dentado
35	<b>BRAZIL PE012</b>	Dentado	Brasil	553	-8.283	-35.967	Dentado
36	<b>26373</b>	Cateto Nortista	Brasil	801	-16.667	-49.255	Cristalino
37	<b>BRAZIL 60</b>	Cateto Grande	Brasil	500	-22.217	-54.8	Cristalino
38	<b>BRVI 104</b>	Cristalino Costeño Tropical	Islas Virgenes	16	18.443	-64.582	Dentado
39	<b>BRVI 117</b>	Cristalino Cubano	Islas Virgenes	190	18.419	-64.629	Cristalino>Dentado
40	<b>BRVI 139</b>	Caribeño Precoz	Islas Virgenes	63	18.395	-64.647	Dentado
41	<b>BRVI 142</b>	Caribeño Precoz	Islas Virgenes	92	18.419	-64.662	Dentado
42	<b>CHZM 13 080</b>	Camelia	Chile	625	-33.45	-70.667	Cristalino>Dentado
43	<b>PPS1066</b>		Chile	234	-40.917	-73.167	Cristalino
44	<b>PPS938</b>		Chile	458	-38.85	-72.367	Cristalino
45	<b>PPS941</b>		Chile	458	-39.517	-71.867	Semi-Cristalino
46	<b>PPS949</b>		Chile	338	-39.583	-72.033	Cristalino
47	<b>PPS961</b>		Chile	338	-40.233	-72.017	Cristalino
48	<b>PPS939</b>		Chile	458	-38.85	-71.7	Cristalino
49	<b>CHZM 08 060</b>	Pisankalla	Chile	450	-37.75	-72	Cristalino
50	<b>PPS55</b>		Chile	337	-40.633	-72.317	Cristalino
51	<b>CHZM 09 030</b>	Araucano	Chile	100	-38.733	-72.95	Cristalino
52	<b>CHZM 01 062</b>	Chulpi	Chile	2300	-19.3	-69.26	Dulce>Harinoso
53	<b>CHZM 08 049</b>	Ocho Corridas	Chile	250	-37.45	-72	Cristalino
54	<b>NARINO 369</b>	Pija naranja	Colombia	1554	1.51	-77.14	Cristalino
55	<b>NARINO 392</b>	Chococeno	Colombia	2000	1.3	-78.3	Cristalino
56	<b>TOLIMA 389</b>	Yucatan	Colombia	450	3.45	-75.15	Cristalino>Dentado
57	<b>MAGDAL 443</b>	Guirua	Colombia	1860	10	-74	Cristalino

<b>ID</b>	<b>Accesión</b>	<b>Raza</b>	<b>Pais de Origen</b>	<b>Altura de Colecta (msnm)</b>	<b>Latitud</b>	<b>Longitud</b>	<b>Tipo de grano</b>
58	<i>CAUCA 384</i>	Amagaceno	Colombia	1600	3.055	-76.211	Cristalino
59	<i>ANTI 392</i>	Amagaceño	Colombia	1371	6.55	-75.83	Cristalino>Dentado
60	<i>BOYACA 462</i>	Pira	Colombia	900	5.3	-73.3	Cristalino
61	<i>CUNDIN 327</i>	Pira	Colombia	2377	5.317	-73.817	Cristalino
62	<i>TOLIMA 390</i>	Pira	Colombia	312	4.033	-74.967	
63	<i>CAQUET 321</i>	Andaqui	Colombia	700	1	-74	Cristalino
64	<i>CUNDIN 428</i>	Sabanero	Colombia	1960	4.528	-73.929	
65	<i>CUNDIN 465</i>	Pollo	Colombia	1600	5	-74	Cristalino
66	<i>COLOMB 613</i>	Negrito	Colombia	540	11.25	-74.18	Cristalino
67	<i>SANTAN 332</i>	Puya grande	Colombia	978	7.61	-72.65	Dentado
68	<i>CUNDIN 480</i>	Pira	Colombia	1000	5	-74	Cristalino
69	<i>NARINO 534</i>	Montana	Colombia	1500	1.3	-78.3	Cristalino
70	<i>SANTAN 317</i>	Cabuya	Colombia	1920	7.383	-72.65	Harinoso
71	<i>CORDOB 342</i>	Cariaco	Colombia	229	8.717	-75.883	Harinoso
72	<i>MAGDAL 390</i>	Puya	Colombia	229	8.317	-73.633	Dentado
73	<i>SANTAS 340</i>	Cacao	Colombia	1252	6.283	-73.167	Harinoso
74	<i>CRIC 141</i>	Salvadoreño>Oloton	Costa Rica	50	10.89	-85.01	Cristalino>Dentado
75	<i>CRIC 166</i>	Salvadoreño	Costa Rica	80	10.25	-85.62	Dentado>Cristalino
76	<i>CUBA 44</i>	Cristalino Cubano	Cuba	244	20.383	-76.433	Cristalino>Dentado
77	<i>CUBA 54</i>	Canilla	Cuba	305	20.9	-76.25	Dentado>Cristalino
78	<i>CUBA 63</i>	Cristalino Cubano	Cuba	91	20.667	-75.667	Cristalino>Dentado
79	<i>CUBA 85</i>	Chandelle	Cuba	80	20.3	-76.25	Dentado>Cristalino
80	<i>CUBA 12</i>	Cristalino Cubano	Cuba	122	22.417	-83.733	Cristalino>Dentado>Harinoso
81	<i>RDOM 270</i>	Chandelle	República Dominicana	130	19.73	-70.92	Dentado
82	<i>RDOM GPI</i>	Chandelle	República Dominicana	170	19.11	-70.3	Dentado
83	<i>ECUADO 398</i>	Racimo de Uva	Ecuador	2423	-0.083	-78.417	Harinoso
84	<i>ECUA 476</i>	Morocho	Ecuador	2195	-0.05	-78.45	Cristalino
85	<i>ECUA 617</i>	Tusilla	Ecuador	549	-4.15	-78.92	Cristalino
86	<i>ECUA 500</i>	Canguil	Ecuador	2213	0.167	-78.417	Cristalino
87	<i>ECUA 696</i>	Canguil	Ecuador	1920	0.22	-78.25	Cristalino

<b>ID</b>	<b>Accesión</b>	<b>Raza</b>	<b>Pais de Origen</b>	<b>Altura de Colecta (msnm)</b>	<b>Latitud</b>	<b>Longitud</b>	<b>Tipo de grano</b>
88	<b>ECUADO 486</b>	Chillo	Ecuador	2560	-1.733	-78.583	Cristalino>Harinoso
89	<b>ECUADO 704</b>	Kcello Ecuatoriano	Ecuador	1800	-4.1	-79.3	Cristalino
90	<b>ECUA 418</b>	Morocho	Ecuador	2195	-0.317	-78.45	Cristalino
91	<b>ECUA 881</b>	Tusilla	Ecuador	1372	-3	-79	Dentado>Cristalino
92	<b>ECUADO 573</b>	Montana	Ecuador	2259	0.22	-78.25	Cristalino
93	<b>ECUADO 746</b>	Uchima	Ecuador	1737	-4.167	-79.5	Dentado
94	<b>ECUADO 979</b>	Chococeno	Ecuador	27	0.4	-79.3	Cristalino
95	<b>ECUADO 424</b>	Chulpi	Ecuador	2579	0.05	-78.2	Dulce>Harinoso
96	<b>ECUADO 418</b>	Morocho	Ecuador	2195	-0.317	-78.45	Cristalino
97	<b>ECUADO X14237</b>	Shima	Ecuador	2689	-3.09	-79.01	Harinoso>Cristalino
98	<b>ECUADO 459</b>	Mishca	Ecuador	2286	-0.111	-78.291	Harinoso
99	<b>DESI03</b>	Caribeño Precoz	Guadalupe	40	16.308	-61.073	Cristalino
100	<b>GFO053</b>	Chandelle	Guadalupe	40	16.226	-61.393	Cristalino, semi-Cristalino
101	<b>GFO059</b>	Cristalino Costeño Tropical	Guadalupe	40	16.251	-61.529	Cristalino
102	<b>GUAD 6</b>	Caribeño Precoz	Guadalupe	100	16.317	-61.367	Dentado
103	<b>STB073</b>	Chandelle	Guadalupe	286	16.026	-61.723	Cristalino>semi-Cristalino
104	<b>EGT014</b>	Cristalino Costeño Tropical	Guadalupe	40	16.275	-61.257	Cristalino, semi-Cristalino
105	<b>MGA201</b>	Early caribbean	Guadalupe	143	15.903	-61.249	Cristalino
106	<b>DESI06</b>	Caribeño Precoz	Guadalupe	40	16.334	-61.015	Cristalino
107	<b>MGA235</b>	Caribeño Precoz	Guadalupe	10	15.955	-61.312	Cristalino
108	<b>GUAT 162</b>	Quicheño early	Guatemala	1753	15.09	-91.15	Cristalino
109	<b>GUAT 36</b>	San Marceño	Guatemala	1829	15.41	-91.15	Cristalino
110	<b>GUAT 45</b>	Olotón	Guatemala	1749	14.633	-90.517	Cristalino>Dentado>Harinoso
111	<b>GUAT 527</b>	Quicheño early	Guatemala	1844	15.5	-91.77	Dentado
112	<b>GUAT 529</b>	Comiteco	Guatemala	1600	15.42	-91.77	Cristalino>Dentado
113	<b>GUAT 606</b>	Olotón > Nal-Tel	Guatemala	1890	15.28	-91.43	Cristalino
114	<b>GUAT 820</b>	Negro de Chimaltenango	Guatemala	1219	15.4	-90.38	Cristalino
115	<b>GUATEM 685</b>	Negro de Chimaltenango	Guatemala	1219	15.4	-90.38	Cristalino
116	<b>GUATEM SH22</b>	Olotón	Guatemala	1749	14.633	-90.517	Cristalino>Dentado>Harinoso
117	<b>GUATEM 213</b>	Quicheno Early	Guatemala	610	15.01	-91.11	Cristalino>Dentado
118	<b>GUATEM 390</b>	Negro de Altura (Negro de Tierra Fria)	Guatemala	2316	14.8	-91.52	Dentado

<b>ID</b>	<b>Accesión</b>	<b>Raza</b>	<b>Pais de Origen</b>	<b>Altura de Colecta (msnm)</b>	<b>Latitud</b>	<b>Longitud</b>	<b>Tipo de grano</b>
119	<b>GUATEM 404</b>	Nal-Tel Blanco Tierra Alta	Guatemala	2774	15.05	-91.41	Cristalino
120	<b>GUATEM 155</b>	Nal-Tel Blanco Tierra Baja	Guatemala	1006	15.09	-90.49	Dentado>Cristalino
121	<b>GUATEM 548</b>	Salpor	Guatemala	688	14.62	-91.6	Dentado
122	<b>GUATEM 10</b>	San Marceno	Guatemala	206	14.85	-92.08	Dentado
123	<b>GUATEM 93</b>	Nal-Tel Amarillo de Tierra Baja	Guatemala	914	14.783	-89.55	Dentado>Harinoso
124	<b>MART 4</b>	Caribeño Precoz	Martinica	213	14.75	-61.18	Cristalino>Dentado
125	<b>CHIH 131</b>	Gordo	Mexico	2095	29.2	-108.12	Harinoso
126	<b>CHIH 133</b>	Azul > Cristalino de Chihuahua	Mexico	2095	29.2	-108.15	Dentado
127	<b>CHIH 150</b>	Palomero Toluqueño > Cristalino de Chihuahua	Mexico	2140	29.37	-107.73	Cristalino
128	<b>CHIH 207</b>	Cristalino de Chihuahua>Apachito	Mexico	2510	26.8	-107.25	Cristalino>Dentado
129	<b>CHIH 218</b>	Cristalino de Chihuahua>Azul	Mexico	1964	28.85	-107.75	Harinoso
130	<b>CHIH 38</b>	Apachito	Mexico	1800	27.8	-107.48	Cristalino>Harinoso
131	<b>CHIS 104</b>	Zapalote Grande	Mexico	100	15.33	-92.67	Dentado
132	<b>CHIS 94</b>	Comiteco	Mexico	1800	16.25	-92.13	Cristalino>Dentado
133	<b>MEXI5/CHIH13 5</b>	Palomero Toluqueño	Mexico	2652	19.283	-99.667	Cristalino
134	<b>MEXI 7</b>	Cacahuacintle	Mexico	2652	19.283	-99.633	Harinoso
135	<b>MEXI 726</b>	Chalqueño	Mexico	2700	19.283	-99.65	Dentado
136	<b>MICH 362</b>	Mushito	Mexico	2340	19.667	-102.338	Harinoso
137	<b>NAYA 24</b>	Harinoso de Ocho	Mexico	100	21.97	-105.28	Dentado>Cristalino
138	<b>OAXA 50</b>	Zapalote Chico	Mexico	100	16.58	-94.6	Dentado
139	<b>PI 217413</b>	Zapalote Chico	Mexico	2220	19.383	-99.15	Dentado, Cristalino, semi-Dentado
140	<b>PUEB 70</b>	Cónico	Mexico	2469	18.87	-97.4	Dentado
141	<b>PUEB 91</b>	Arrocillo Amarillo	Mexico	2060	19.75	-97.55	Dentado
142	<b>SINA 6</b>	Chapalote	Mexico	75	26.433	-108.633	Cristalino
143	<b>SONO 24</b>	Onaveño	Mexico	1640	30.9	-110.61	Cristalino
144	<b>TLAX 151</b>	Conico	Mexico	2463	19.3	-97.77	Dentado
145	<b>VERA 359</b>	Arrocillo Amarillo	Mexico	2200	19.77	-97.23	Dentado
146	<b>YUCA GP2</b>	Nal-Tel>Harinoso	Mexico	30	20.25	-89.65	Dentado
147	<b>CHIS 662</b>	Zapalote Chico	Mexico	100	16.217	-93.889	Dentado

<b>ID</b>	<b>Accesión</b>	<b>Raza</b>	<b>Pais de Origen</b>	<b>Altura de Colecta (msnm)</b>	<b>Latitud</b>	<b>Longitud</b>	<b>Tipo de grano</b>
148	<i>SINA 2</i>	Chapalote	Mexico	61	24.83	-107.36	Cristalino
149	<i>OAXA 223</i>	Bolita	Mexico	200	15.75	-96.517	Dentado
150	<i>MEXI 3</i>	Conico	Mexico	2652	19.283	-99.65	Cristalino
151	<i>VERA 39</i>	Tuxpeño	Mexico	106	20.47	-97.08	Dentado
152	<i>CHIS 63</i>	Tepecintle	Mexico	760	16.767	-93.38	Dentado
153	<i>GUER GP25</i>	Vandeno	Mexico	1000	17.52	-101.28	Cristalino>Dentado
154	<i>JALI 71</i>	Elotes Occidentales	Mexico	1555	20.17	-103.03	Harinoso>Cristalino
155	<i>NAYA 337K</i>	Jala	Mexico	1080	21.09	-104.43	
156	<i>NAYA 39</i>	Reventador	Mexico	100	21.95	-105.28	Cristalino
157	<i>OAXA GPI</i>	Olotillo	Mexico	120	16.33	-98.03	Dentado
158	<i>CHIH 160</i>	Gordo>Cristalino de Chihuahua	Mexico	2000	28.55	-107.47	Dentado>Harinoso
159	<i>ZACA 12</i>	Cónico Norteño	Mexico	1950	21.35	-102.8	Dentado
160	<i>JALI 102</i>	Tabloncillo	Mexico	1280	19.45	-103.28	Dentado
161	<i>JALI 78</i>	Maíz Dulce	Mexico	1890	21.18	-102.85	Dulce>Dentado
162	<i>MEXI 212</i>	Cacahuacintle	Mexico	2800	19.07	-99.45	Harinoso
163	<i>PANA 168</i>	Salvadoreño>Clavillo	Panama	213	8.22	-81.7	Cristalino>Dentado
164	<i>PAZM 13041</i>	Avati' Moroti' Ti'	Paraguay	230	-23	-56	Harinoso
165	<i>PARAGU CB6-62</i>	Avati' Pichinga	Paraguay	300	-27	-57	Cristalino
166	<i>PARAGU PG6-42</i>	Avati' Moroti'	Paraguay	400	-26.1	-55.933	Harinoso
167	<i>PAZM 6060</i>	Avati	Paraguay	240	-26.267	-56.117	
168	<i>ANC. 212</i>	Cuzco	Peru	3100	-9.33	-77.23	Harinoso
169	<i>ANC 393</i>	San Jeronimo>Huancavelicano	Peru	2600	-9.29	-77.64	Harinoso
170	<i>APUC 140</i>	Confite Puneño	Peru	2500	-13.62	-72.88	Cristalino
171	<i>APUC 171</i>	San Jeronimo	Peru	2300	-13.62	-72.88	Cristalino
172	<i>PERU 674</i>	Uchuquilla	Peru	2025	-15.83	-70.03	Harinoso
173	<i>PERU 1303</i>	Confite Morocho	Peru	2744	-12.77	-75.03	Cristalino
174	<i>PERU 1283</i>	Huancavelicano	Peru	2698	-12.77	-75.03	Harinoso
175	<i>CUZ. 363</i>	Cuzcri	Peru	3200	-13.46	-72.16	Cristalino, Harinoso, Dentado
176	<i>CUZ. 56</i>	Cuzco Cristalino Amarillo	Peru	3268	-13.55	-71.89	Cristalino
177	<i>ANC. 186</i>	Ancashino	Peru	2900	-9.43	-77.25	



ID	<i>Accesión</i>	Raza	Pais de Origen	Altura de Colecta (msnm)	Latitud	Longitud	Tipo de grano
178	<i>LBQUE. 7</i>	Mochero	Peru	23	-6.7	-79.917	Harinoso
179	<i>HVCA. 69</i>	Chulpi	Peru	2800	-12.4	-74.9	Harinoso
180	<i>LIM. 47</i>	Perla	Peru	253	-12.083	-76.95	Cristalino
181	<i>CAJ. 24</i>	Kkulli	Peru	2300	-7.167	-78.517	Harinoso>Dulce
182	<i>SVIN 5</i>	Cristalino Costeño Tropical	San Vicente	80	13.28	-61.12	Cristalino>Dentado
183	<i>SVIN GP2A</i>	Cristalino Costeño Tropical	San Vicente	100	13.28	-61.24	Dentado>Cristalino
184	<i>TRIN 6</i>	Tusón	Trinidad	31	10.651	-61.4	Dentado>Cristalino
185	<i>URUG 697</i>	Cuarenton >Colorado Cateto	Uruguay	183	-33.8	-56.833	Cristalino
186	<i>URUG 8A</i>	Cateto Sulino	Uruguay	20	-34.455	-56.617	semi-Cristalino
187	<i>URUGUA 1131A</i>	Cateto Sulino	Uruguay	108	-34.167	-57.5	Cristalino
188	<i>URUGUA 1187A</i>	Dente Branco	Uruguay	5	-32.683	-58.133	Dentado
189	<i>URUGUA 637</i>	Cateto Sulino	Uruguay	80	-34.34	-57.73	Cristalino
190	<i>VEN 736</i>	Sabanero	Venezuela	2468	8.75	-71.33	Cristalino
191	<i>VENEZU 631</i>	Cariaco	Venezuela	365	6.333	-63.5	Harinoso
192	<i>VEN 650</i>	Cristalino Cubano	Venezuela	137	8.9	-64.22	Cristalino
193	<i>VEN 405</i>	Tusón	Venezuela	425	7.58	-71.97	Dentado
194	<i>VEN 442</i>	Tusón	Venezuela	996	10.267	-67.083	Dentado

## 7.2 Aislamiento del ADN genómico

1. Coloque entre 300 y 400 mg de tejido liofilizado y molido en un tubo de polipropileno de 15 ml para centrifugación. Se obtienen de 50 a más de 100  $\mu\text{g}$  de ADN / 100 mg de tejido seco.
2. Agregue 9.0 ml de solución amortiguadora CTAB para extracción, calentada ( $65^{\circ}\text{C}$ ), a los 300-400 mg de tejido liofilizado y molido. Es mejor distribuir el tejido sobre las paredes laterales del tubo antes de agregar la solución amortiguadora, con el fin de evitar que se apelmace el tejido seco en el fondo. Mezcle invirtiendo el tubo varias veces con suavidad.
3. Incube durante 60 a 90 minutos en un horno a  $65^{\circ}\text{C}$  , agitando los tubos continuamente con suavidad.
4. Retire los tubos del horno. Espere de 4 a 5 minutos para que se enfríen los tubos y agregue 4.5 ml de cloroformo/octanol (24:1). Mezcle agitando suavemente los tubos durante 10 minutos
5. Centrifugue en una centrifugadora de mesa durante 10 minutos a  $1300\text{-}1500 \times g^2$  a temperatura ambiente.
6. Vierta la capa acuosa superior en otros tubos de 15 ml. Agregue 4.5 ml de cloroformo/octanol y agite suavemente durante 5 a 10 minutos.
7. Centrifugue en una centrifugadora de mesa durante 10 minutos a  $1300\text{-}1500 \times g^2$  a temperatura ambiente.
8. Con una pipeta transfiera la capa acuosa superior a otros tubos de 15 ml que contengan 30  $\mu\text{l}$  de 10 mg/ml de ARNasa (previamente hervida). Mezcle invirtiendo con suavidad los tubos e incube durante 30 minutos a temperatura ambiente.
9. Agregue 6.0 ml de isopropanol (2-propanol). Mezcle invirtiendo con suavidad los tubos.
10. Retire el ADN precipitado con un gancho de vidrio y proceda a realizar lavados de ADN (Lavado 1: 76% de EtOH, 0.2 M NaOAc, Lavado 2: (76% EtOH, 10 mM  $\text{NH}_4\text{OAc}$ ) y resuspenda.

### 7.3 Lista de marcadores moleculares tipo SSR utilizados

SSR	Diversidad Latinoamérica	Exploración Análisis Asociación	Posición en el genoma	Unidad de repetición	Gen	Rango en pares de bases	Programa de PCR	stutter ratio
<i>phi109275</i>	*	*	1	AGCT		108-134	SSR56	0
<i>phi056</i>		*	1.01	CCG	<a href="#">tub1 beta tubulin1</a>	241-258	SSR56	0
<i>umc1917</i>		*	1.04	CTG		130-160	SSR52	0
<i>umc2047</i>	*	*	1.09	GACT		119-133	Q58	0
<i>phi308707</i>	*		1.1	AGC		118-125	SSR54	0
<i>phi227562</i>	*		1.12	ACC		309-325	SSR56	0.07
<i>phi083</i>	*		2.04	AGCT	<a href="#">prp2 pathogenesis-related protein2</a>	125-139	SSR54	0.48
<i>umc2250</i>		*	2.04	ACG		150-160	Q58	0
<i>phi090</i>		*	2.08	ATATC		137-146	Q60	0
<i>phi127</i>		*	2.08	AGAC		110-127	Q62	0
<i>phi374118</i>		*	3.02	ACC		214-238	SSR54	0
<i>phi029</i>	*	*	3.04	AG/AGC G	<a href="#">tpi4 triose phosphate isomerase4</a>	150-161	SSR56	0.26
<i>phi102228</i>	*	*	3.04-05	AAGC		122-131	SSR54	0
<i>umc1266</i>	*	*	3.06	CAG	<a href="#">psbs1 photosystem II subunit PsbS1</a>	120-147	Q52	0
<i>phi046</i>	*		3.08	ACGC			Q60	0
<i>phi072</i>	*		4.01	AAAC		142-162	SSR56	0.03
<i>phi079</i>		*	4.05	AGATG	<a href="#">gpc1 glyceraldehyde-3-phosphate dehydrogenase1</a>	180-200	SSR62	0
<i>phi076</i>	*	*	4.11	GAGCG G	<a href="#">cat3 catalase3</a>	156-177	Q60/SSR60	0
<i>phi109188</i>	*	*	5	AAAG		145-181	Q56/SSR54	0
<i>phi008</i>		*	5.03	GGC	<a href="#">rab15 responsive to abscisic acid15</a>	100-110	Q60	0
<i>umc1447</i>		*	5.03	CTT		115-120	Q58	0
<i>umc1332</i>	*	*	5.04	CTA		126-147	SSR60	0
<i>phi331888</i>	*		5.04	AAG		129-136	Q56	0
<i>phi085</i>	*		5.07	AACGC		243-267	Q60	0
<i>phi075</i>	*	*	6	CT	<a href="#">fdx1 ferredoxin1</a>	203-235	Q54	0
<i>phi031</i>	*		6.04	GTAC	<a href="#">pl1 purple plant1</a>	186-225	SSR60	0
<i>phi299852</i>	*	*	6.08	AGC		102-147	Q60	0
<i>phi034</i>	*	*	7.02	CCT	<a href="#">cyp6 cytochrome P450</a>	125-143	Q60	0

SSR	Diversidad Latinoamérica	Exploración Análisis Asociación	Posición en el genoma	Unidad de Repetición	Gen	Rango en pares de bases	Programa de PCR	stutter ratio
<i>umc1545</i>		*	7	AAGA	<a href="#">hsp3 heat shock protein3</a>	70-98	Q60	0
<i>phi114</i>		*	7.03	GCCT	<a href="#">oec6 oxygen evolving complex17 candidate</a>	135-166	SSR60	0
<i>phi069</i>	*		7.05	GAC		187-207	Q58	0.07
<i>umc1304</i>		*	8.02	TCGA		121-141	SSR54	0
<i>phi115</i>	*		8.03	AT/ATAC	<a href="#">act1 actin1</a>	242-258	SSR54	0.12
<i>phi014</i>	*		8.04	GGC	<a href="#">rip1 ribosome-inactivating protein1</a>	431-435	SSR64	0
<i>umc1161</i>		*	8.06	GCTGGG		140-160	SSR54	0
<i>phi065</i>		*	9.03	CACTT	<a href="#">pep1 phosphoenolpyruvate carboxylase1</a>	116-123	SSR54	0
<i>phi108411</i>	*	*	9.06	AGCT		116-123	Q60	0
<i>phi059</i>	*		10.02	ACC	<a href="#">aasr1 abscisic acid stress ripening1</a>	147-165	Q58	0
<i>phi063</i>	*	*	10.02	TATC		154-180	Q62	0
<i>umc1367</i>		*	10.03	CGA		130-160	Q60/Q62	0
<i>phi062</i>	*		10.04	ACG	<a href="#">mgs1 male-gametophyte specific1</a>	160-164	SSR56	0
<i>phi084</i>	*	*	10.04	GAA	<a href="#">nac1 NaCl stress protein1</a>	148-157	SSR54	0
<i>umc1196</i>	*	*	10.07	CACACG		137-173	SSR56	0

#### 7.4 Reacción en cadena de la Polimerasa (PCR)

Reactivo	10 µl de 1 RXN
Amortig Taq (10X)	1.0
dNTP (2.5 mM)	1.2
MgCl <sub>2</sub> (50 mM)	0.4
Iniciadores (2µM)	X
H <sub>2</sub> Odd	Y
Enzima <i>Taq</i>	0.15
ADN (5ng/µl)	1.5

X= la cantidad varía según iniciador

Y= Se ajusta para llevar a un volumen total de 10 µl

#### Programa de PCR estándar

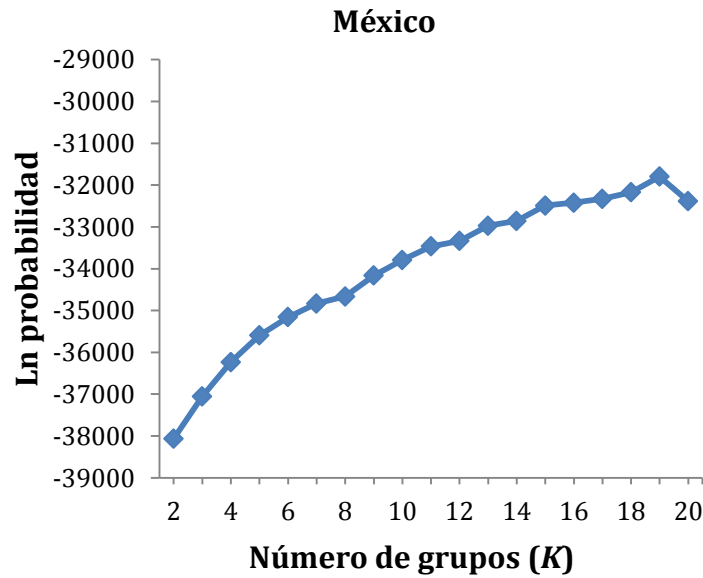
<b>1 ciclo de:</b> (93°C por 1 minuto)	<b>30 ciclos de:</b> 93°C por 30 seg X°C por 1 minuto (X° está entre 50 y 68°) 72°C por 1 minuto	<b>1 ciclo de:</b> 72°C por 5 minutos
---	---	--

#### Programa de PCR “touchdown”

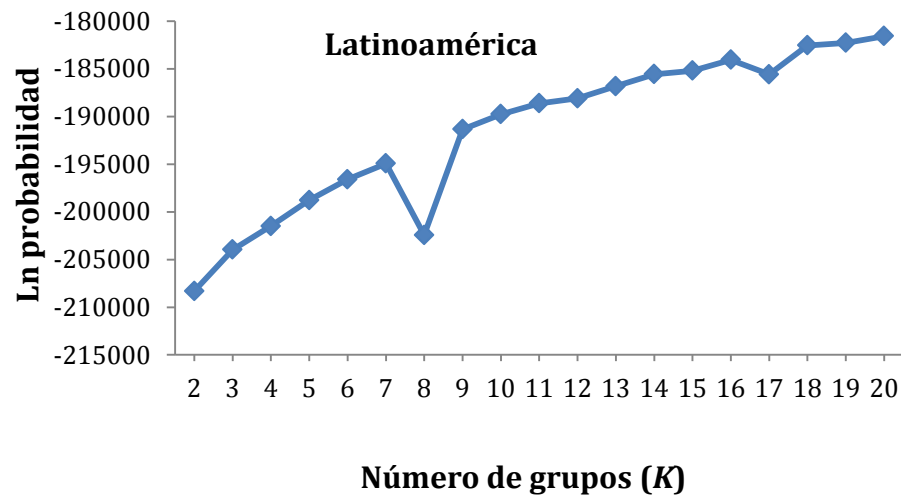
<b>1 ciclo de:</b> 94°C por 2 minutos	<b>7 ciclos de:</b> 94°C por un minuto Y°C por 1 minuto (reduciendo un grado por ciclo) 72°C por 1 minuto Y=69, 59 ó 54°	<b>35 ciclos de:</b> 94°C por 1 minuto Z°C por 1 minuto 72°C por 1 minuto Z= 62, 57, 52 ó 5 min	<b>1 ciclo de:</b> 72°C por 5 minutos
--	---	---	--

7.5 Gráficas del Ln de la probabilidad del análisis de estructura poblacional de 38 poblaciones de maíz de México (a), y (b) 194 poblaciones de Latinoamérica, se graficaron los valores promedios de  $\ln[\text{prob}(\text{dato})]$  de 2 corridas independientes para cada valor de K.

a.



b.



**7.6 Cálculo  $\Delta K$  basado en Evanno et al. 2005**

**Paso 1.** Media de los valores  $[\ln[\text{prob}(\text{dato})]]$  (m) y cálculo de la desviación estándar (s).

**Paso 2.** Cálculo  $L'(K)$ , donde,  $L'(k) = L(k) - L(k-1)$

**Paso 3.** Cálculo  $|L''(k)|$ , donde  $|L''(k)| = |L'(k+1) - L'(k)|$

De esta forma:

$$|L''(k)| = |L(k+1) - L(k) - L'(k)|$$

$$|L''(k)| = |L(k+1) - L(k) - [(L(k) - L(k-1))]|$$

$$|L''(k)| = |L(k+1) - L(k) - L(k) + L(k-1)|$$

$$|L''(k)| = |L(k+1) - 2L(k) + L(k-1)|$$

**Paso 4.** Cálculo  $\Delta k$ , donde,  $\Delta k = |m L''(k)| / s [L(k)]$

Ejemplo para México con 5 repeticiones por cada  $K$  (valores de las 5 repeticiones para cada  $k$  no mostrados)

<b><math>K</math></b>	<b>(m)</b>	<b>(s)</b>	<b><math> L''(K) </math></b>	<b><math>\Delta K</math></b>
1	-39165.6	0.1		
2	-37951.6	12.4	447.8	36.0
3	-37185.5	444.5	257.7	0.6
4	-36161.6	20.1	671.1	33.4
5	-35808.9	449.9	251.8	0.6
6	-35204.4	29.3	372.2	12.7
7	-34972.0	231.4	267.7	1.2
8	-34472.0	143.5	251.5	1.8
9	-34223.4	285.6		

## 7.7 Códigos de raza asignados a las poblaciones de maíz mexicanas

<b>Raza</b>	<b>Código</b>
Gordo	GO
Gordo>Cristalino de Chihuahua	GO-CCH
Azul > Cristalino de Chihuahua	AZ-CCH
Palomero Toluqueño > Cristalino de Chihuahua	PT-CCH
Cristalino de Chihuahua>Apachito	AP
Cristalino de Chihuahua>Azul	AZ
Zapalote Chico	ZC
Zapalote Grande	ZG
Comiteco	CM
Palomero Toluqueño	PT
Cacahuacintle	CA
Chalqueño	CQ
Mushito	MU
Harinoso de Ocho	HO
Cónico Norteño	CN
Cónico	CO
Arrocillo Amarillo	AA
Chapalote	CH
Onaveño	ON
Nal-Tel>Harinoso	NT
Chapalote	CH
Bolita	BO
Tuxpeño	TU
Tepecintle	TE
Vandeno	VE
Elotes Occidentales	EO
Jala	JA
Reventador	RE
Olotillo	OL
Tabloncillo	TA
Maíz Dulce	MD





**7.9** Frecuencias alélicas expresadas en el rango de 0-1, cambios de frecuencia alélica absolutas expresadas en porcentaje entre materiales de pre-mejoramiento del Banco de germoplasma, material mejorado del Programa de Fisiología del CIMMYT y poblaciones nativas, marcadas con color los cambios mayores a 10%.

<b>Marcador</b>	<b>Alelo</b>	<b>Banco de Germoplasma</b>	<b>P. Fisiología</b>	<b>Poblaciones nativas</b>	<b>Banco-Pop</b>	<b>Banco-Fisiologia</b>	<b>Fisiología-Pop</b>
phi008	100	0.83	0.94	0.87	3.59	11.11	7.52
phi008	104	0.15	0.06	0.10	4.75	9.26	4.51
phi008	89	0.00	0.00	0.01	0.65	0.00	0.65
phi008	98	0.02	0.00	0.02	0.50	1.85	2.35
phi029	140	0.00	0.00	0.00	0.12	0.00	0.12
phi029	144	0.00	0.19	0.02	2.22	18.89	16.67
phi029	146	0.74	0.68	0.48	25.29	5.97	19.32
phi029	148	0.00	0.00	0.08	7.59	0.00	7.59
phi029	151	0.25	0.13	0.39	13.64	11.67	25.31
phi029	153	0.01	0.00	0.01	0.57	1.25	0.68
phi029	155	0.00	0.00	0.02	1.67	0.00	1.67
phi029	158	0.00	0.00	0.01	0.62	0.00	0.62
phi034	117	0.09	0.03	0.07	2.12	5.33	3.21
phi034	120	0.32	0.64	0.40	8.27	32.50	24.23
phi034	123	0.06	0.00	0.01	4.71	5.67	0.95
phi034	126	0.03	0.00	0.03	0.25	2.67	2.92
phi034	129	0.11	0.02	0.07	4.19	9.67	5.48
phi034	132	0.06	0.00	0.05	1.69	6.33	4.64
phi034	135	0.08	0.02	0.07	0.88	6.00	5.12
phi034	137	0.15	0.13	0.18	2.17	2.83	5.00
phi034	141	0.11	0.17	0.14	2.90	6.00	3.10
phi056	234	0.11	0.00	0.03	8.77	11.33	2.56
phi056	237	0.13	0.20	0.16	3.05	7.33	4.29
phi056	240	0.00	0.00	0.00	0.06	0.00	0.06
phi056	243	0.02	0.00	0.08	5.50	2.00	7.50
phi056	246	0.39	0.47	0.38	1.18	7.33	8.51
phi056	249	0.29	0.18	0.25	3.58	11.50	7.92
phi056	252	0.03	0.16	0.07	4.05	12.50	8.45
phi056	255	0.02	0.00	0.03	0.88	2.33	3.21

<b>Marcador</b>	<b>Alelo</b>	<b>Banco de Germoplasma</b>	<b>P. Fisiología</b>	<b>Poblaciones nativas</b>	<b>Banco-Pop</b>	<b>Banco-Fisiología</b>	<b>Fisiología-Pop</b>
<b>phi063</b>	154	0.04	0.18	0.04	0.91	13.33	14.24
<b>phi063</b>	156	0.00	0.00	0.00	0.14	0.00	0.14
<b>phi063</b>	159	0.00	0.00	0.02	1.56	0.00	1.56
<b>phi063</b>	161	0.13	0.60	0.19	5.92	46.67	40.75
<b>phi063</b>	166	0.00	0.00	0.01	0.88	0.00	0.88
<b>phi063</b>	170	0.56	0.18	0.49	6.24	37.78	31.54
<b>phi063</b>	174	0.13	0.04	0.16	2.93	8.89	11.81
<b>phi063</b>	178	0.04	0.00	0.04	0.57	4.44	3.88
<b>phi063</b>	182	0.02	0.00	0.01	0.83	1.85	1.02
<b>phi063</b>	186	0.00	0.00	0.02	2.11	0.00	2.11
<b>phi063</b>	190	0.00	0.00	0.00	0.27	0.00	0.27
<b>phi063</b>	214	0.02	0.00	0.01	0.36	1.85	1.50
<b>phi063</b>	218	0.04	0.00	0.00	3.94	4.07	0.14
<b>phi063</b>	222	0.00	0.00	0.00	0.14	0.00	0.14
<b>phi063</b>	226	0.01	0.00	0.00	1.11	1.11	0.00
<b>phi065</b>	117	0.01	0.02	0.05	3.93	0.67	3.27
<b>phi065</b>	127	0.47	0.53	0.59	11.93	5.17	6.77
<b>phi065</b>	135	0.01	0.00	0.01	0.53	1.33	0.80
<b>phi065</b>	138	0.15	0.00	0.07	8.07	14.67	6.60
<b>phi065</b>	143	0.05	0.00	0.03	1.27	4.67	3.40
<b>phi065</b>	148	0.31	0.46	0.25	6.00	14.83	20.83
<b>phi075</b>	205	0.00	0.00	0.00	0.07	0.00	0.07
<b>phi075</b>	209	0.00	0.00	0.01	0.72	0.00	0.72
<b>phi075</b>	221	0.00	0.00	0.00	0.26	0.00	0.26
<b>phi075</b>	223	0.42	0.48	0.29	12.65	5.83	18.48
<b>phi075</b>	225	0.02	0.03	0.03	1.21	0.83	0.38
<b>phi075</b>	227	0.49	0.41	0.52	2.82	8.50	11.32
<b>phi075</b>	229	0.00	0.00	0.04	3.66	0.00	3.66
<b>phi075</b>	232	0.04	0.00	0.02	2.83	4.33	1.50
<b>phi075</b>	234	0.00	0.00	0.01	0.65	0.00	0.65
<b>phi075</b>	235	0.02	0.08	0.03	1.01	5.83	4.82
<b>phi075</b>	238	0.01	0.02	0.05	3.96	0.33	3.63
<b>phi075</b>	239	0.00	0.00	0.01	0.72	0.00	0.72
<b>phi075</b>	241	0.00	0.00	0.00	0.26	0.00	0.26
<b>phi075</b>	244	0.00	0.00	0.00	0.13	0.00	0.13
<b>phi076</b>	157	0.00	0.00	0.01	0.72	0.33	1.05
<b>phi076</b>	159	0.12	0.00	0.15	2.98	12.33	15.31
<b>phi076</b>	163	0.01	0.00	0.00	0.67	0.67	0.00
<b>phi076</b>	165	0.02	0.00	0.14	12.41	1.67	14.07
<b>phi076</b>	168	0.00	0.02	0.00	0.19	1.67	1.48
<b>phi076</b>	170	0.84	0.63	0.68	15.94	21.00	5.06
<b>phi076</b>	176	0.01	0.35	0.01	0.32	34.33	34.01

<b>Marcador</b>	<b>Alelo</b>	<b>Banco de Germoplasma</b>	<b>P. Fisiología</b>	<b>Poblaciones nativas</b>	<b>Banco-Pop</b>	<b>Banco-Fisiología</b>	<b>Fisiología-Pop</b>
<b>phi079</b>	177	0.17	0.20	0.13	3.52	3.33	6.85
<b>phi079</b>	182	0.24	0.23	0.09	15.45	0.67	14.79
<b>phi079</b>	185	0.26	0.14	0.31	5.27	11.83	17.11
<b>phi079</b>	187	0.30	0.34	0.38	7.73	3.83	3.89
<b>phi079</b>	192	0.00	0.00	0.06	5.30	0.33	5.64
<b>phi079</b>	195	0.00	0.00	0.01	1.03	0.00	1.03
<b>phi079</b>	197	0.03	0.08	0.02	0.36	5.67	6.03
<b>phi084</b>	149	0.01	0.00	0.04	2.90	1.00	3.90
<b>phi084</b>	154	0.31	0.47	0.30	1.38	15.67	17.04
<b>phi084</b>	157	0.62	0.50	0.60	1.29	11.67	10.38
<b>phi084</b>	161	0.06	0.03	0.06	0.48	3.00	2.52
<b>phi084</b>	163	0.00	0.00	0.00	0.25	0.00	0.25
<b>phi090</b>	126	0.00	0.00	0.00	0.12	0.00	0.12
<b>phi090</b>	137	1.00	0.80	0.98	2.10	20.00	17.90
<b>phi090</b>	145	0.00	0.20	0.02	1.98	20.00	18.02
<b>phi102228</b>	120	0.94	0.83	0.82	11.34	10.37	0.97
<b>phi102228</b>	125	0.06	0.13	0.17	10.25	7.04	3.21
<b>phi102228</b>	129	0.00	0.03	0.01	1.09	3.33	2.24
<b>phi108411</b>	110	0.00	0.00	0.00	0.42	0.00	0.42
<b>phi108411</b>	112	0.00	0.00	0.01	0.55	0.00	0.55
<b>phi108411</b>	119	0.76	0.86	0.84	8.30	9.83	1.53
<b>phi108411</b>	122	0.24	0.14	0.15	9.27	9.83	0.56
<b>phi109188</b>	144	0.01	0.02	0.00	0.55	0.93	1.48
<b>phi109188</b>	148	0.00	0.00	0.01	1.13	0.00	1.13
<b>phi109188</b>	153	0.06	0.00	0.02	3.92	5.56	1.64
<b>phi109188</b>	157	0.08	0.12	0.10	2.29	3.52	1.23
<b>phi109188</b>	161	0.42	0.64	0.47	5.07	22.31	17.25
<b>phi109188</b>	165	0.13	0.14	0.28	14.34	0.83	13.51
<b>phi109188</b>	168	0.30	0.08	0.11	18.62	21.30	2.67
<b>phi109188</b>	173	0.00	0.00	0.00	0.31	0.00	0.31
<b>phi109188</b>	177	0.01	0.00	0.01	0.05	0.74	0.69
<b>phi109275</b>	118	0.31	0.25	0.41	9.85	6.11	15.96
<b>phi109275</b>	121	0.00	0.00	0.00	0.26	0.00	0.26
<b>phi109275</b>	124	0.02	0.00	0.08	5.47	2.22	7.69
<b>phi109275</b>	125	0.08	0.00	0.02	6.24	7.78	1.54
<b>phi109275</b>	127	0.22	0.00	0.15	6.66	21.85	15.19
<b>phi109275</b>	130	0.00	0.00	0.01	0.64	0.00	0.64
<b>phi109275</b>	131	0.16	0.43	0.12	3.95	27.78	31.73
<b>phi109275</b>	134	0.18	0.32	0.22	3.89	13.89	10.00
<b>phi109275</b>	138	0.04	0.00	0.00	3.25	3.70	0.45

Marcador	Alelo	Banco de Germoplasma	P. Fisiología	Poblaciones nativas	Banco-Pop	Banco-Fisiología	Fisiología-Pop
phi114	133	0.67	0.50	0.55	11.70	16.67	4.97
phi114	159	0.00	0.11	0.05	4.96	10.46	5.50
phi114	163	0.04	0.21	0.11	6.95	16.39	9.44
phi114	166	0.29	0.18	0.28	0.70	10.19	9.48
phi114	169	0.00	0.00	0.00	0.48	0.00	0.48
phi127	101	0.00	0.00	0.00	0.42	0.00	0.42
phi127	109	0.49	0.70	0.65	15.98	20.74	4.76
phi127	114	0.00	0.00	0.01	0.60	0.00	0.60
phi127	117	0.00	0.00	0.00	0.36	0.00	0.36
phi127	122	0.34	0.22	0.31	3.61	12.78	9.17
phi127	126	0.13	0.08	0.02	10.51	4.26	6.25
phi127	93	0.04	0.00	0.00	3.23	3.70	0.48
phi299852	102	0.06	0.05	0.05	0.94	1.00	0.06
phi299852	105	0.01	0.00	0.01	0.05	1.00	1.05
phi299852	107	0.08	0.02	0.12	3.46	6.67	10.12
phi299852	110	0.39	0.56	0.37	1.53	16.83	18.36
phi299852	114	0.01	0.01	0.01	0.44	0.17	0.28
phi299852	117	0.17	0.02	0.10	7.04	15.00	7.96
phi299852	119	0.20	0.24	0.28	7.96	4.17	3.80
phi299852	125	0.00	0.00	0.01	0.41	0.33	0.74
phi299852	129	0.01	0.00	0.00	0.23	0.67	0.43
phi299852	132	0.00	0.07	0.01	0.80	6.67	5.86
phi299852	135	0.04	0.04	0.04	0.15	0.50	0.65
phi299852	141	0.03	0.00	0.01	2.78	3.33	0.56
phi299852	95	0.00	0.00	0.00	0.06	0.00	0.06
phi299852	99	0.00	0.00	0.00	0.37	0.00	0.37
phi374118	212	0.55	0.71	0.60	4.85	15.50	10.65
phi374118	215	0.02	0.00	0.00	1.64	2.00	0.36
phi374118	218	0.08	0.03	0.15	7.00	5.00	12.00
phi374118	220	0.00	0.00	0.03	2.82	0.33	3.15
phi374118	222	0.01	0.01	0.04	3.58	0.17	3.41
phi374118	225	0.33	0.25	0.14	19.27	8.33	10.94
phi374118	228	0.00	0.00	0.00	0.42	0.00	0.42
phi374118	232	0.00	0.00	0.02	1.58	0.00	1.58
phi374118	237	0.00	0.00	0.01	0.67	0.00	0.67
umc1161	130	0.04	0.00	0.03	1.08	3.70	2.63
umc1161	133	0.00	0.00	0.00	0.13	0.00	0.13
umc1161	135	0.04	0.00	0.08	3.70	4.44	8.14
umc1161	138	0.02	0.00	0.06	4.37	1.85	6.22
umc1161	142	0.77	0.93	0.63	14.01	15.93	29.94
umc1161	145	0.00	0.00	0.02	1.73	0.00	1.73
umc1161	149	0.09	0.07	0.13	4.94	1.85	6.79
umc1161	152	0.04	0.00	0.04	0.22	4.07	4.29

Marcador	Alelo	Banco de Germoplasma	P. Fisiología	Poblaciones nativas	Banco-Pop	Banco-Fisiología	Fisiología-pop
umc1196	120	0.00	0.01	0.00	0.00	0.83	0.83
umc1196	124	0.00	0.00	0.00	0.12	0.00	0.12
umc1196	127	0.00	0.00	0.00	0.12	0.00	0.12
umc1196	131	0.06	0.03	0.05	0.68	2.59	1.91
umc1196	132	0.03	0.06	0.00	2.41	3.24	5.65
umc1196	135	0.14	0.18	0.11	3.40	4.26	7.65
umc1196	138	0.00	0.00	0.00	0.06	0.00	0.06
umc1196	140	0.11	0.09	0.10	0.31	1.57	1.27
umc1196	144	0.54	0.33	0.51	3.33	21.94	18.61
umc1196	149	0.01	0.02	0.02	0.56	0.19	0.37
umc1196	151	0.02	0.02	0.03	1.54	0.19	1.73
umc1196	153	0.00	0.00	0.00	0.12	0.00	0.12
umc1196	156	0.09	0.27	0.16	7.59	17.78	10.19
umc1266	119	0.06	0.00	0.01	5.51	6.30	0.78
umc1266	130	0.94	1.00	0.98	4.79	6.30	1.50
umc1266	143	0.00	0.00	0.00	0.33	0.00	0.33
umc1266	147	0.00	0.00	0.00	0.39	0.00	0.39
umc1304	117	0.03	0.05	0.06	2.67	2.00	0.67
umc1304	120	0.00	0.00	0.00	0.47	0.00	0.47
umc1304	122	0.01	0.00	0.01	0.40	1.00	0.60
umc1304	127	0.72	0.84	0.77	4.93	12.17	7.23
umc1304	131	0.01	0.06	0.00	0.80	4.83	5.63
umc1304	133	0.23	0.05	0.16	6.87	18.00	11.13
umc1332	107	0.03	0.01	0.02	0.85	1.83	0.98
umc1332	114	0.00	0.00	0.01	1.21	0.00	1.21
umc1332	117	0.35	0.37	0.37	1.79	1.67	0.12
umc1332	121	0.02	0.00	0.00	1.97	2.33	0.36
umc1332	130	0.02	0.06	0.01	0.73	3.83	4.56
umc1332	132	0.01	0.00	0.01	0.06	1.33	1.27
umc1332	135	0.01	0.00	0.02	1.64	0.67	2.30
umc1332	139	0.22	0.22	0.20	2.58	0.67	1.91
umc1332	142	0.34	0.35	0.35	1.55	1.33	0.21
umc1367	144	0.05	0.00	0.04	1.03	5.33	4.30
umc1367	148	0.00	0.00	0.01	1.09	0.00	1.09
umc1367	151	0.03	0.02	0.03	0.06	1.00	0.94
umc1367	154	0.11	0.07	0.08	2.91	4.00	1.09
umc1367	157	0.81	0.92	0.82	0.48	10.33	9.85
umc1367	160	0.00	0.00	0.02	2.42	0.00	2.42

Marcador	Alelo	Banco de Germoplasma	P. Fisiología	Poblaciones nativas	Banco-Pop	Banco-Fisiología	Fisiología-pop
umc1447	109	0.20	0.40	0.21	0.91	20.00	19.09
umc1447	113	0.05	0.02	0.02	2.91	3.67	0.76
umc1447	115	0.35	0.40	0.38	2.91	4.67	1.76
umc1447	119	0.38	0.18	0.38	0.33	20.17	20.50
umc1447	122	0.02	0.00	0.00	1.24	1.67	0.42
umc1447	127	0.00	0.01	0.00	0.00	0.83	0.83
umc1545	63	0.18	0.05	0.12	5.43	12.78	7.35
umc1545	66	0.00	0.00	0.00	0.06	0.00	0.06
umc1545	71	0.14	0.19	0.17	2.78	4.72	1.94
umc1545	75	0.57	0.75	0.61	3.52	17.59	14.07
umc1545	79	0.10	0.01	0.08	2.53	9.54	7.01
umc1545	81	0.00	0.00	0.02	1.60	0.00	1.60
umc1917	128	0.68	0.79	0.77	9.42	11.50	2.08
umc1917	131	0.01	0.00	0.01	0.36	1.33	0.97
umc1917	134	0.09	0.10	0.05	3.27	1.33	4.61
umc1917	137	0.06	0.07	0.10	3.85	0.33	3.52
umc1917	140	0.02	0.00	0.01	0.85	2.00	1.15
umc1917	144	0.14	0.04	0.05	8.79	9.83	1.05
umc2047	113	0.00	0.00	0.00	0.48	0.00	0.48
umc2047	119	0.00	0.00	0.01	0.77	0.00	0.77
umc2047	123	0.48	0.34	0.40	7.96	13.50	5.54
umc2047	131	0.52	0.66	0.59	6.71	13.50	6.79
umc2250	133	0.13	0.03	0.11	1.28	9.26	7.98
umc2250	149	0.87	0.97	0.89	1.28	9.26	7.98