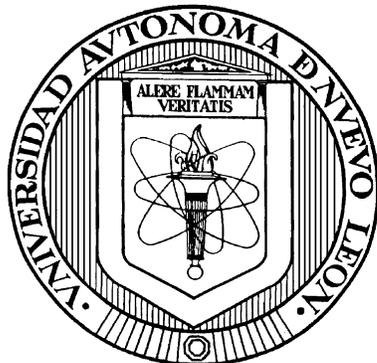


UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

FACULTAD DE AGRONOMÍA



**IDENTIFICACIÓN DE MARCADORES MOLECULARES LIGADOS A
CARACTERES DE INTERÉS EN MATERIALES ÉLITE DE SORGO DULCE
Sorghum bicolor (L.) Moench DE LA FA-UANL**

TESIS

QUE PARA OBTENER EL GRADO DE
MAESTRÍA EN CIENCIAS EN PRODUCCIÓN AGRÍCOLA
PRESENTA

IVON MONTSERRAT CERDA HURTADO

GENERAL ESCOBEDO, N.L.

DICIEMBRE 2014

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

FACULTAD DE AGRONOMÍA



**IDENTIFICACIÓN DE MARCADORES MOLECULARES LIGADOS A
CARACTERES DE INTERÉS EN MATERIALES ÉLITE DE SORGO DULCE
Sorghum bicolor (L.) Moench DE LA FA-UANL**

TESIS

QUE PARA OBTENER EL GRADO DE
MAESTRÍA EN CIENCIAS EN PRODUCCIÓN AGRÍCOLA
PRESENTA

IVON MONTSERRAT CERDA HURTADO

GENERAL ESCOBEDO, N.L.

DICIEMBRE 2014

**IDENTIFICACIÓN DE MARCADORES MOLECULARES LIGADOS A
CARACTERES DE INTERÉS EN MATERIALES ÉLITE DE SORGO DULCE
Sorghum bicolor (L.) Moench DE LA FA-UANL**

ESTA TESIS FUE REVISADA Y APROBADA POR EL COMITÉ PARTICULAR
COMO REQUISITO PARCIAL PARA OBTENER EL GRADO DE

MAESTRA EN CIENCIAS EN PRODUCCIÓN AGRÍCOLA

COMITÉ PARTICULAR

Ph. D. Francisco Zavala García
Asesor Principal

D. C. Adriana Gutiérrez Díez
Asesor Auxiliar

M. C. Eduardo Alejandro García Zambrano
Asesor Auxiliar

D. C. Ernesto Javier Sánchez Alejo
Subdirector de Estudios de Posgrado e Investigación

DEDICATORIA

Para mi familia, mis padres, mis hermanas Lala, Patty, Vero y Moni.

A mis amigos y compañeros, es especial a los que ahora pueden ser libres.

AGRADECIMIENTO

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por el apoyo económico otorgado para realizar mis estudios de posgrado.

A la Comunidad Europea por el apoyo al proyecto SWEET SORGHUM: AN ALTERNATIVE ENERGY CROP con que se realizó el presente trabajo.

A la Subdirección de Estudios de Posgrado e Investigación y personal administrativo de la Facultad de Agronomía de la Universidad Autónoma de Nuevo León.

Al Ph. D. Francisco Zavala García, por su confianza, por su paciencia y la revisión del escrito y por todos los conocimientos que me heredó como maestro y director principal.

A la D. C. Adriana Gutiérrez Díez, por su amistad, disponibilidad, observaciones y sugerencias, que me ayudaron a realizar un mejor trabajo de tesis, por sus conocimientos como maestra y asesora.

Al M. C. Eduardo A. García Zambrano, por sus consejos, su ánimo y optimismo así como sus conocimientos como maestro y asesor.

Al M. C. Héctor Williams Alanís por sus consejos, conocimientos y apoyo desinteresado que me brindó durante el periodo de mi formación.

A los maestros del Posgrado en Ciencias Agrícolas de la FA-UANL, por la formación, el apoyo y los consejos.

Al equipo del Campus de Marín, Javier Sánchez, Eulalio Benavides, la M. C. María del Carmen Rodríguez Vázquez, Ing. Héctor Reyes por todo su apoyo en el trabajo de campo para llevar a cabo ésta investigación.

Al Q.B.P. Josué Martínez Morales por su asistencia técnica y apoyo en el Laboratorio de Biotecnología.

A mis compañeros: Iván Pequeño, Alan Rodríguez, Fernando Villela, Félix Varela, Manuel Quintero, Javier Ayala y Eleazar Lugo con quienes compartí el periodo de la maestría.

ÍNDICE DE CONTENIDO

	Página
ÍNDICE DE CUADROS.....	viii
ÍNDICE DE FIGURAS.....	ix
NOMENCLATURA.....	xii
RESUMEN.....	xiii
SUMMARY.....	xv
1. INTRODUCCIÓN.....	1
1.1. OBJETIVO GENERAL.....	5
1.1.1. Objetivos específicos.....	5
1.2. HIPÓTESIS.....	5
2. REVISIÓN DE LITERATURA.....	7
2.1. BIOCOMBUSTIBLES.....	7
2.1.1. Bioetanol.....	8
2.2. SORGO DULCE.....	8
2.2.1. Colección de Sorgo Dulce de la FA-UANL.....	10
2.3. DIVERSIDAD GENÉTICA Y MEJORAMIENTO GENÉTICO.....	10
2.4. MARCADORES MOLECULARES.....	12
2.4.1. Marcadores tipo AFLP (Amplified Fragments Length Polymorphism).....	12

2.4.2. Marcadores tipo RAPD (Random Amplified Polymorphic DNA)	13
2.4.3. Marcadores tipo SCAR (Sequence Characterized Amplified Regions).....	14
2.5. ESTIMACIÓN DE LA DIVERSIDAD GENÉTICA	15
3. MATERIALES Y MÉTODOS	17
3.1. MATERIAL GENÉTICO.....	17
3.2. CARACTERIZACIÓN MORFOLÓGICA	19
3.3. DESARROLLO DE MARCADORES MOLECULARES	19
3.3.1. Generación de Marcadores AFLP	20
3.3.2. Generación de Marcadores RAPD	22
3.3.3. Selección de Marcadores RAPD Candidatos a Secuenciación	24
3.4. DESCRIPTIVA DE MARCADORES.....	25
3.5. ANÁLISIS ESTADÍSTICO.....	26
3.6. GENERACIÓN DE MARCADORES SCAR.....	29
4. RESULTADOS Y DISCUSIÓN	32
4.1. EXTRACCIÓN DE ADN	32
4.2. GENERACIÓN DE MARCADORES AFLP.....	34
4.3. GENERACIÓN DE MARCADORES RAPD.....	38
4.3.1. Reproducibilidad de los resultados en marcadores RAPD	41
4.4. SELECCIÓN DE MARCADORES RAPD CANDIDATOS A SECUENCIACIÓN	42
4.5 ANÁLISIS ESTADÍSTICO.....	43
4.5.1. Análisis de Agrupamiento	47
4.6. GENERACIÓN DE MARCADORES SCAR.....	50

5. CONCLUSIONES	64
6. LITERATURA CITADA.....	66
7. ANEXOS.....	74

ÍNDICE DE CUADROS

Cuadro		Página
1	Características generales del material genético utilizado en el trabajo de investigación.....	18
2	Componentes utilizados en la mezcla de reacción de pre amplificación del ADN genómico digerido de sorgo dulce de la FA-UANL.....	21
3	Descriptiva de datos marcadores AFLP.....	35
4	Coefficientes de correlación lineal entre las matrices.....	44
5	Correlaciones cofenética del análisis de conglomerados.....	50
6	Identificación de fragmentos RAPD secuenciados.....	51
7	Resultados del BLAST y descripción de las alineaciones con <i>Sorghum bicolor</i> sp. y <i>Zea mays</i> , secuencias agrupadas por iniciador común.....	52
8	Diseño de iniciadores <i>Forward</i> y <i>Reverse</i> derivados de fragmentos alineados con secuencias de <i>Sorghum bicolor</i> sp.....	54

ÍNDICE DE FIGURAS

Figura		Página
1	Electroforesis en gel de agarosa, extracciones ADN de tejido foliar de plantas de sorgo dulce FA-UANL.....	33
2	Electroforesis en gel de poliacrilamida 6.5% de marcadores AFLP obtenidos con las combinaciones de iniciadores M-CTA x E-ACA y M-CTC x E-AAC.....	37
3	Electroforesis en gel de poliacrilamida 6.5% de marcadores AFLP obtenidos con las combinaciones de iniciadores M-CTA x E-ACG y M-CTC x E-ACT.....	38
4	Electroforesis en gel de agarosa al 3% para productos de amplificación para los iniciadores OP-A 2, OP-C 14, OP-C 19 y OP-K 15.....	40
5	Electroforesis en gel de agarosa al 3% para productos de amplificación para el iniciador OP-E 20 con el genotipo FA-UANL 33A x FA-UANL 5.....	41
6	Electroforesis en gel de agarosa al 3% para productos de amplificación para el iniciador OP-I 14 con el genotipo FA-UANL 35A x FA-UANL 39.....	42

Figura		Página
7	Correlaciones entre matrices 1.-Distancia Euclidiana para Azúcares vs. 4.-Distancia Dice (1-S) para AFLP y RAPD.....	46
8	Correlaciones entre matrices 2.-Distancia Dice (1-S) para AFLP vs. 3.-Distancia Dice (1-S) para RAPD.....	46
9	Dendograma UPGMA con datos morfológicos de azúcares reductores (Glucosa y Fructosa). Distancia Euclidiana. Correlación cofenética=0.76.....	47
10	Dendograma UPGMA con datos moleculares AFLP y RAPD Distancia Dice ($\sqrt{(S-1)}$). Correlación cofenética =0.938.....	48
11	Dendograma UPGMA con datos morfológicos y moleculares AFLP y RAPD; utilizando distancia Gower ($\sqrt{(1-S)}$). Correlación cofenética=0.952.....	49
12	Ejemplo de cromatogramas obtenidos de la secuenciación de fragmentos. A) Fragmento 60, ejemplo de un cromatograma correcto: picos uniformemente espaciados, cada uno con un solo color, equidistantes. B) Fragmento 21, ejemplo de una secuenciación con poca definición.....	56
13	Electroforesis en gel de agarosa al 3% para productos de amplificación de los iniciadores <i>forward</i> y <i>reverse</i> 1, 3, 4, 7, 8 y 10.....	58

Figura		Página
14	Electroforesis en gel de agarosa al 3% para productos de amplificación de los iniciadores <i>forward</i> y <i>reverse</i> 1, 3, 4, 7, 8 y 10 utilizando el ADN genómico total de los genotipos en el estudio	59
15	Electroforesis en gel de acrilamida: bis-acrilamida (40:1) al 8% para los productos de PCR con el juego de iniciadores 4, <i>Forward</i> : 5'-CGCCCGTATGTGTTTCTCCT-3' y <i>Reverse</i> : 5'-GGAGTTGTAAATGGCGGTGC-3'.....	61
16	Secuencia de 469 pb de la región polimórfica de <i>Sorghum bicolor</i> sp., mostrando los sitios de unión del iniciador RAPD (OP-K 20) y los marcadores SCAR (<i>Forward</i> 4 y <i>Reverse</i> 4).....	61
17	Electroforesis en gel de acrilamida: bis-acrilamida (40:1) al 8% para los productos de PCR con el juego de iniciadores 7, <i>Forward</i> : 5'-GAGTGCTCCTACTCTTGTCATCC-3' y <i>Reverse</i> : 5'-AGCTAATGTACATCAGCGGGT-3'.....	62
18	Secuencia de 414 pb de la región polimórfica de <i>Sorghum bicolor</i> sp., mostrando los sitios de unión del iniciador RAPD (OP-C 14) y los marcadores SCAR (<i>Forward</i> 7 y <i>Reverse</i> 7).....	63

NOMENCLATURA

A	Adenina
Abs	Absorbancia
ADN	Ácido Desoxirribonucleico
AFLP	Amplified Fragments Length Polymorphism
BLAST	Basic Local Alignment Search Tool
C	Citosina
C ₄	Metabolismo de síntesis de ácidos de cuatro Carbonos
CTAB	Bromuro de hexadeciltrimetilamonio
dds	Días después de siembra
dNTPs	Deoxinucleótidos 5'-trifosfatados
<i>EcoRI</i>	<i>Escherichia coli</i> RI (enzima de restricción)
EDTA	Ethylenediaminetetra acetato
F1	Primera generación Filial
G	Guanina
g	gramos(s)
g ²	gravedad(es)
h	hora(s)
HPLC	High Performance Liquid Chromatography
Kb	Kilobase(s)
M	Concentración Molar
mA	mili amperio(s)
ml	mililitro(s)
mM	mili Molar
MPM	Marcador de Peso Molecular
<i>Mse</i>	<i>Micrococcus</i> sp. (R. Morgan) (enzima de restricción)
min	minuto(s)
NCBI	National Center for Biological Information
NH ₄ Ac	Acetato de amonio
pb	pares de bases
PCR	Polimerase Chain Reaction
PVC	Policloruro de vinilo
RAPD	Random Amplified Polymorphic DNA
RNasa	Ribonucleic acid nuclease (enzima)
T	Timina
<i>Taq</i>	<i>Thermophilus aquaticus</i>
TE	Tris-EDTA (buffer)
TES	Tris-EDTA-SDS (buffer)
T _m	Temperatura de alineación
Tris	Tris(hydroxymetil)amino-metano
s	segundo(s)
SCAR	Sequence Characterized Amplified Regions
SDS	Sulfato Dodecyl de Sodio
SSR	Simple Sequence Repeats
U	Unidad(es) de enzima
UPGMA	Unweighted Pair Group Method with Arithmetic Averages
UV	Ultravioleta
V	Voltios
W	Watts
°Bx	grados Brix
°C	grados Celsius
μg	microgramo(s)= 10 ⁻⁶ gramo
ng	nanogramo(s)=10 ⁻⁹ gramo
μl	microlitro(s)= 10 ⁻⁶ litro
nm	nanómetro(s)=10 ⁻⁹ metro

RESUMEN

El uso de marcadores moleculares en programas de mejoramiento ha demostrado ser un instrumento eficaz en los procedimientos de selección. Los marcadores AFLP y RAPD podrían ser efectivos para la identificación de polimorfismos y la construcción de diagramas de agrupamiento por distancias genéticas. El desarrollo de SCAR permite obtener resultados reproducibles basados en la PCR de los marcadores RAPD polimórficos, por lo tanto pueden utilizarse en la identificación de genotipos de sorgo altos en el contenido de azúcares en el jugo del tallo.

El objetivo del estudio fue desarrollar un marcador molecular fiable y reproducible que pueda utilizarse en programas de selección de sorgo dulce enfocados a la producción de bioetanol.

Se utilizaron catorce genotipos de sorgo: FA-UANL 37A, FA-UANL 33A, FA-UANL 35A, AN601, FA-UANL 39, FA-UANL 5, KANSAS COLLIER, FA-UANL 33A x KANSAS COLLIER, AN601 x FA-UANL 39, FA-UANL 35A x FA-UANL 39, FA-UANL 33A x FA-UANL 5 y FA-UANL 37A x FA-UANL5 y FORTUNA. Se realizaron reacciones dúplex para 30 combinaciones de iniciadores *EcoRI/MseI* para marcadores AFLP y se probaron 180 iniciadores RAPD. Se identificaron 53 fragmentos RAPD presentes en los genotipos altos en azúcares y ausentes en los genotipos más bajos en azúcares, éstos se purificaron, clonaron y secuenciaron.

Se obtuvieron 549 fragmentos AFLP y 256 RAPD. Se diseñaron dos pares de iniciadores SCAR: el juego de iniciadores 4 amplificó una banda de 231 pb presente en los genotipos altos en azúcares (A601 x FA-UANL 39, FA-UANL 35A x FA-UANL 39, FA-UANL 33A x FA-UANL 5 y FA-UANL 37A x FA-UANL 5), dicha banda no amplificó en ningún otro genotipo, pero tampoco se encontró en ningún progenitor; el juego de iniciadores 7 amplificó una banda de 341 pb presente en los genotipos altos en azúcares (A601 x FA-UANL 39 y FA-UANL 33A x FA-UANL 5), esta banda no se amplificó en ningún otro genotipo, pero sí se encontró en los progenitores B de éstas mismas líneas (FA-UANL 39 y FA-UANL 5), además de estar presente en el progenitor KANSAS COLLIER.

Los marcadores AFLP y RAPD fueron herramientas eficientes para la estimación de la variación genética en sorgo al presentar un coeficiente de

distancia genética desde 0.42 (Dice $\sqrt{S-1}$) en el agrupamiento. La técnica de AFLP resultó ser ventajosa para evaluar la variación entre genotipos, al tener un mayor rendimiento de bandas por juego de iniciadores; sin embargo, no son adecuados para el diseño de iniciadores SCAR por su tamaño. La utilización de marcadores RAPD permitió el diseño de dos juegos de iniciadores SCAR provenientes de fragmentos que alinearon con secuencias genómicas de sorgo reportadas en el NCBI.

SUMMARY

The use of molecular markers in breeding programs has proven to be an effective tool in selection procedures. AFLP and RAPD markers could be effective for identification of polymorphisms and the construction of grouping by genetic distances. SCAR development allows for reproducible results of PCR-based polymorphic RAPD markers therefore be used in identifying high sorghum genotypes in the content of sugars in the juice from the stem.

The objective of the study was to develop a reliable and reproducible molecular marker that can be used in selection programs focused on sweet sorghum bioethanol production.

Fourteen sorghum genotypes were used: FA-UANL 37A, FA-UANL 33A, FA-UANL 35A, AN601, FA-UANL 39, FA-UANL 5, KANSAS COLLIER, FA-UANL 33A x KANSAS COLLIER, AB601 x FA-UANL 39, FA-UANL 35A x FA-UANL 39, FA-UANL 33A x FA-UANL 5 and FA-UANL 37A x FA-UANL5 and FORTUNA. Duplex reactions for 30 primer combinations *EcoRI/MseI* for AFLP markers were conducted, along with the test of 180 RAPD primers. 53 RAPD fragments present in the genotypes higher in sugars and absent at lower sugar genotypes were identified, purified, cloned and sequenced.

There were obtained 549 AFLP and 256 RAPD fragments. Two pairs of SCAR primers were designed: the set of primers 4 amplified a band of 231 bp present in the high sugar genotypes (A601 x FA-UANL 39, FA-UANL 35A x FA-UANL 39, FA-UANL 33A x FA-UANL 5 and FA-UANL 37A x FA-UANL 5). This band did not amplified in any other genotype, but neither was present in any parent; the set of primers 7 amplified a band of 341 bp present in the high sugar genotypes (A601 x FA-UANL 39 and FA-UANL 33A x FA-UANL 5), this band was not amplified in any other genotype, but was found in the B lines of the parents (FA-UANL 39 and FA-UANL 5) and it was also present in the parent KANSAS COLLIER.

The AFLP and RAPD markers were efficient for estimating genetic variation in sorghum by presenting a genetic distance coefficient from 0.42 (Dice $\sqrt{S-1}$) in the cluster tools. AFLP technique proved to be advantageous to evaluate the variation between genotypes, having a higher yield of bands by primers set; however, they are not suitable for designing SCAR primers by size. Using RAPD markers allowed the design of two SCAR primer sets from fragments that aligned with sorghum reported genomic sequences at NCBI.

1. INTRODUCCIÓN

El agotamiento de las reservas petroleras ha obligado a proponer fuentes alternativas de energía (Sperling, 1989); una de ellas es el bioetanol basado en la utilización de biomasa vegetal, cuya ventaja radica en la habilidad fotosintética de las plantas y su adaptabilidad para crecer en diversos climas y bajo diferentes condiciones ambientales (Shoemaker y Bransby, 2010; Calviño y Messing, 2012).

Actualmente el sorgo dulce (*Sorghum bicolor* (L.) Moench) es una alternativa para la producción de bioetanol; cuenta con características agronómicas que lo convierten en una opción viable como fuente de materia prima tales como una elevada actividad fotosintética para convertir la energía solar en biomasa, alta eficiencia en el uso de agua y tolerancia al estrés por salinidad. Estas características le proporcionan adaptabilidad, obteniendo rendimientos rentables con mayor calidad así como bajo costo de producción cuando se establece en zonas áridas, tropicales y subtropicales (Almodares y Hadi, 2009; Shen *et al.*, 2011).

Normalmente, la selección de genotipos de sorgo altos en azúcares se realiza dependiendo de ciertas características morfológicas (peso y altura de

planta, área foliar, diámetro de tallo, número de nodos, longitud de panícula, etc.), fisiológicas (días a floración, tasa de respiración, rendimiento fotosintético, tasa de crecimiento) y bioquímicas (concentración de azúcares y almidón en el jugo, caracterización de azúcares) bien definidas en condiciones naturales de campo, siendo este proceso de selección, un proceso que consume tiempo y es relativamente laborioso (Zou *et al.*, 2011). Debido a la dificultad en los procedimientos de selección en el sorgo dulce, se requiere el desarrollo de técnicas moleculares para identificar genotipos altos en azúcares necesarios en la realización de cruas dirigidas de forma más eficiente que combinen el alto contenido de azúcar, alta producción de jugo y biomasa (Raghuwanshi y Birch, 2010; Calviño y Messing, 2012).

El uso de marcadores moleculares en programas de mejoramiento vegetal ha demostrado ser un instrumento eficaz. Una técnica disponible para generar marcadores moleculares es la de los fragmentos amplificados de longitud polimórfica (AFLP: *Amplified Fragments Length Polymorphism*), ésta se basa en la amplificación de fragmentos de restricción del ADN (ácido desoxirribonucleico) genómico total, usando iniciadores específicos como adaptadores, son marcadores dominantes y permite la identificación de un gran número de polimorfismos por *loci* analizado, aunque su costo puede ser una limitante (Vos *et al.*, 1995; Morris *et al.*, 2013). Se ha reportado que los AFLP son eficientes para distinguir diferencias entre genotipos de sorgo dulce detectando hasta un 69% de diversidad genética, con rangos de polimorfismo

desde 0.3 hasta 0.93 (Pecina *et al.*, 2012); además, existen estudios para evaluar el nivel de diversidad genética en líneas mantenedoras de esterilidad (Líneas B) restauradoras de fertilidad (Líneas R) de sorgo en comparación con un grupo de especies exóticas y domesticadas (Menz *et al.*, 2004); los AFLP sirvieron para sugerir que el sorgo dulce es de origen polifilético dentro de *S. bicolor* sp. *bicolor* (Ritter *et al.*, 2007).

Con el desarrollo del método de ADN polimórfico amplificado al azar (RAPD: *Random Amplified Polymorphic DNA*) (Williams *et al.*, 1990), se generan fragmentos al azar de un genoma que pueden muestrearse y utilizarse para detectar diferencias entre genotipos; además, en combinación con la evaluación de datos morfológicos pueden utilizarse en los programas de mejoramiento genético. El procedimiento RAPD tiene varias desventajas, la principal es la sensibilidad a las condiciones de reacción que reduce la reproducibilidad de los resultados obtenidos por diferentes laboratorios. No obstante, los resultados se pueden utilizar para identificar el polimorfismo, y una vez que se han asegurado un conjunto de regiones amplificadas, se secuencian para caracterizar la región amplificada (SCAR: *Sequence Characterized Amplified Regions*), éstos marcadores se pueden desarrollar para producir resultados únicos fiables basados en la PCR (*Polymerase Chain Reaction*) de los marcadores RAPD polimórficos (Kiran *et al.*, 2010).

Los marcadores SCAR son ventajosos con respecto a los marcadores RAPD, ya que son codominantes, detectan un *loci* único definido genéticamente, identificado como bandas distintas mediante electroforesis, son más fáciles de anotar, menos sensible a las condiciones de reacción y son más reproducibles, lo que aumentaría la fiabilidad y reproducibilidad de los ensayos de PCR (Paran y Michelmore, 1993; Singh y Lohithaswa, 2006).

Los marcadores tipo SCAR se han utilizado ampliamente en la identificación de varias especies de plantas (Choi *et al.*, 2008; Devaiah y Venkatasubramanian, 2008; Dhanya *et al.*, 2009; Ladhalakshmi *et al.*, 2009; Ray y Roy, 2009); mediante ésta técnica se han identificado marcadores asociados a la concentración de azúcar y tolerancia a sequía en caña de azúcar (*Saccharum officinarum*) (Srivastava *et al.*, 2012) y la identificación de híbridos intergenéricos de caña de azúcar (Wang *et al.*, 2009).

En sorgo, los trabajos realizados mediante esta técnica se han enfocado a identificar un marcador RAPD estrechamente ligado a un gen para la resistencia a la antracnosis *Colletotrichum graminicola* (Ces.) (Singh *et al.*, 2006), resistencia al tizón de la hoja *Exserohilum turcicum* (Pass.) (Boora *et al.*, 1999). Hasta ahora, la aplicación y la autenticidad de los marcadores SCAR está limitada en el caso de estrés biótico o para la identificación de un rasgo particular correspondiente a un solo gen (Kasai *et al.*, 2000; Kiran *et al.*, 2010).

1.1. Objetivo General

Desarrollar al menos un marcador molecular fiable y reproducible que pueda utilizarse como herramienta para facilitar el programa de selección en poblaciones de sorgo dulce enfocados a la producción de bioetanol.

1.1.1. Objetivos específicos

1. Evaluar mediante marcadores moleculares tipo AFLP y RAPD la diversidad genética y estructura de un subconjunto de accesiones de la colección de sorgo dulce de la FA-UANL.
2. Identificar marcadores RAPD polimórficos en genotipos de sorgo dulce de la FA-UANL.
3. Diseñar iniciadores SCAR que pudieran estar asociados a la concentración de azúcar en la colección de sorgos dulces de la FA-UANL.

1.2. Hipótesis

1. El uso de marcadores moleculares en programas de mejoramiento vegetal ha demostrado ser un instrumento eficaz en los procedimientos de selección; por lo tanto, los marcadores AFLP y RAPD podrían ser efectivos y brindar mayores ventajas para la

identificación de polimorfismos y la construcción de diagramas de agrupamiento por distancias genéticas.

2. El desarrollo de Secuencias Caracterizadas de Regiones Amplificadas (SCAR) permite obtener resultados reproducibles basados en la PCR de los marcadores RAPD polimórficos, por lo tanto pueden utilizarse en la identificación de genotipos altos en azúcares.

2. REVISIÓN DE LITERATURA

2.1. Biocombustibles

Biocombustible se define como un combustible sólido, líquido o gas derivado de biomasa. Se consideran amigables con el medio ambiente porque teóricamente las emisiones de CO₂ que se producen durante su combustión son balanceadas por la absorción del CO₂ en el crecimiento de las plantas. Para que sea sustituto viable a los combustibles fósiles, un biocombustible debe tener beneficio ambiental superior al combustible fósil al que está sustituyendo (Elbehri *et al.*, 2013).

Los futuros biocombustibles líquidos dependerán tanto de las fuentes de materia prima y de las tecnologías para su conversión, así como de la viabilidad de éstos para la infraestructura existente de transporte. Los biocombustibles pueden clasificarse en primera y segunda generación, siendo el bioetanol un derivado de ambas tecnologías (Kosaric *et al.*, 2001).

2.1.1. Bioetanol

El bioetanol es un tipo de biocombustible líquido obtenido mediante la fermentación de azúcares. Las diversas materias primas utilizadas para su producción pueden clasificarse en tres tipos principales: azúcares, almidones y materiales lignocelulósicos (Shoemaker y Bransby, 2010). En la actualidad, el bioetanol se utiliza como combustible alternativo para vehículos en E.U.A., Brasil y China, originados principalmente a partir de almidón de maíz o caña de azúcar, lo que ha sido discutido desde el punto de vista ético por utilizarse como materia prima para un combustible en lugar de ser destinados a la alimentación humana (Kosaric *et al.*, 2001). El sorgo (*Sorghum bicolor* (L.) Moench) cuenta con características agronómicas que lo convierten en una opción viable como fuente de materia prima para la producción de biocombustibles (Almodares y Hadi, 2009; Shen *et al.*, 2011).

2.2. Sorgo Dulce

Clasificación botánica

División: Magnoliophyta

Clase: Liliopsida

Subclase: Commenlinidae

Orden: Cyperales

Familia: Poaceae

Subfamilia: Panicoideae

Tribu: Andropogoneae

Subtribu: Sorghinae

Género: *Sorghum*

Especie: *Sorghum bicolor* (L.) Moench

Subespecie: *Sorghum bicolor* (L.) Moench ssp. *bicolor* (Clayton, 1961)

Todos los sorgos botánicamente identificados como *Sorghum bicolor* ssp. *bicolor* tienen $2n=20$ cromosomas y pueden categorizarse siguiendo variantes agronómicas como: sorgo de grano, sorgo para forraje, sorgo de fibra, sorgo escobero, sorgo dulce y sorgo para biomasa (Murray *et al.*, 2009).

En el sorgo dulce se incluyen aquellos cultivares con alto contenido de azúcar en el jugo en tallos y hojas; es un cultivo de ciclo de vida corto (alrededor de 120 dds), se establece por semilla, tiene alta adaptabilidad al ambiente, posee elevada actividad fotosintética para convertir la energía solar en biomasa, es un cultivo C_4 , es eficiente en el uso de agua para el crecimiento en regiones con altas temperaturas y sequías; tolerancia al estrés por salinidad, es un cultivo que permite obtener alimento, forrajes y combustible a bajo costo (Almodares y Hadi, 2009; Srinivasa *et al.*, 2009; Shen *et al.*, 2011).

2.2.1. Colección de Sorgo Dulce de la FA-UANL

Actualmente la Facultad de Agronomía de la Universidad Autónoma de Nuevo León participa en el consorcio *SWEETFUEL: An Alternative energy crop*, representando a México, que, junto con los demás miembros: Francia (líder), Italia, Alemania, Brasil, India y Sudáfrica tiene el objetivo de mejorar variedades de sorgo dulce para su adaptación a regiones templadas, semi tropicales y áridas, con tolerancia al frío o sequía dependiendo de la región, la producción de biomasa y granos con fácil digestión y alta producción de azúcares en el tallo (www.sweetfuel-project.eu).

2.3. Diversidad Genética y Mejoramiento Genético

El género *Sorghum* comprende un alta diversidad genética y por tanto, potencial para mejoramiento genético y la posibilidad de incrementar su productividad. La diversidad de colores y formas muestran la cantidad de variación en la especie. No existen barreras biológicas o taxonómicas entre cultivares lo que permite su cruzamiento e hibridación. La primera domesticación del género estuvo asociada con el cambio de semillas pequeñas, frágiles y panículas abiertas hacia semillas más grandes, firmes y panículas compactas. Recientemente el mejoramiento se ha enfocado para servir como alimento, forraje y combustible, dependiendo de las necesidades del país,

haciendo énfasis al rendimiento de grano, estabilidad, resistencia a plagas y patógenos, calidad en grano y biomasa (Elbehri *et al.*, 2013).

Los fitomejoradores han explotado genes para acumulación de azúcares (Zheng *et al.*, 2011), alto rendimiento en grano (Qazi *et al.*, 2012) y alto rendimiento de biomasa para su uso como forraje o generación de biocombustibles de segunda generación (Vinutha *et al.*, 2014), entre otros (Elbehri *et al.*, 2013). A causa de la escases de germoplasma especializado en este cultivo para la producción de bioetanol, se requiere la realización de cruza dirigidas de forma más eficiente para combinar el alto contenido de azúcar, producción de jugo y biomasa (Raghuwanshi y Birch, 2010; Calviño y Messing, 2012).

Los rasgos deseables para la producción de etanol son complejos e involucran muchos genes de diferentes rutas, es necesario identificar los recursos genéticos disponibles que puedan utilizarse en programas de mejoramiento genético. La selección de progenitores es un paso esencial para un mejoramiento genético eficaz y requiere de un conocimiento adecuado de la diversidad genética existente y de las relaciones entre los genotipos particulares (Schut *et al.*, 1997).

2.4. Marcadores Moleculares

El conocimiento y entendimiento de la distribución genética de la población de origen se puede lograr con el desarrollo de marcadores moleculares (Calviño *et al.*, 2009; Wang *et al.*, 2009; Alvarado *et al.*, 2011); son además una herramienta disponible para el análisis de genomas de plantas y la asociación de rasgos hereditarios de la variación genética (Wang *et al.*, 2009).

2.4.1. Marcadores tipo AFLP (*Amplified Fragments Length Polymorphism*)

Son una técnica disponible para generar marcadores basados en la amplificación de fragmentos de restricción del ADN (ácido desoxirribonucleico) genómico total, los cuales son rápidos de generar y altamente reproducibles, usan iniciadores específicos como adaptadores, requiere pequeñas cantidades de ADN y no se necesita de conocimiento previo de la secuencia del genoma, son además marcadores dominantes y permiten la identificación de un gran número de polimorfismos por *loci* analizado. Su costo puede ser una limitante y su identificación depende de una electroforesis en gel lo que disminuye su rendimiento por ensayo, además de que existen inconvenientes al precisar la correlación de la separación de bandas con variantes alélicas (Vos *et al.*, 1995; Kesawat y Das, 2009).

Se han reportado que los AFLP son eficientes para distinguir diferencias entre genotipos de sorgo dulce detectando hasta un 69% de diversidad genética, con rangos de polimorfismo desde 0.3 hasta 0.93 (Pecina *et al.*, 2012); además de existir estudios para evaluar el nivel de diversidad genética en líneas mantenedoras de esterilidad (Líneas B) restauradoras de fertilidad (Líneas R) de sorgo en comparación con un grupo de especies exóticas y convertidas (Menz *et al.*, 2004); los AFLP sirvieron para sugerir que el sorgo dulce es de origen polifilético dentro de *S. bicolor* ssp. *bicolor* (Ritter *et al.*, 2007).

2.4.2. Marcadores tipo RAPD (*Random Amplified Polymorphic DNA*)

Por otro lado la técnica de ADN polimórfico amplificado al azar ó RAPD tiene como ventaja ser una técnica simple, rápida y sólo se requiere una pequeña cantidad de ADN, los iniciadores tienen entre 10 y 15 nucleótidos de longitud y una composición en G+C entre 50% y 80%, no requiere el conocimiento previo del genoma y es relativamente económica; sin embargo es muy sensible a cambios pequeños en las condiciones de amplificación, lo que ocasiona baja reproducibilidad, son marcadores dominantes y permite detectar un número medio de polimorfismos (Williams *et al.*, 1990; Kesawat y Das, 2009; Kumar *et al.*, 2009; Jiang, 2010).

Existen reportes que apoyan que los RAPD puede utilizarse eficazmente para el estudio de la diversidad genética del sorgo (Menkir *et al.*, 1997; Ayana *et al.*, 2000; Iqbal *et al.*, 2010); además de demostrar que su uso es más eficiente que los descriptores agronómicos en la formación de grupos en una colección de germoplasma (Dahlberg *et al.*, 2002).

2.4.3. Marcadores tipo SCAR (*Sequence Characterized Amplified Regions*)

Michelmore *et al.* (1991) describieron por primera vez la técnica de SCAR. En esta técnica los extremos de los marcadores RAPD son secuenciados permitiendo diseñar cebadores más largos (22-24 nucleótidos) a partir de fragmentos RAPD ligados a un rasgo de interés para un locus en particular (Paran y Michelmore, 1993; Kumar *et al.*, 2009). Mediante ésta técnica se han podido identificar marcadores asociados a la concentración de azúcar y tolerancia a sequía en caña de azúcar (*Saccharum officinarum*) (Srivastava *et al.*, 2012) y la identificación de híbridos intergenéricos de caña de azúcar (Wang *et al.*, 2009). En sorgo, los trabajos realizados mediante esta técnica han estado enfocados a identificar un marcador RAPD estrechamente ligado a un gen para la resistencia a la antracnosis *Colletotrichum graminicola* (Ces.) (Singh *et al.*, 2006) y resistencia al tizón de la hoja *Exserohilum turcicum* (Pass.) (Boora *et al.*, 1999).

2.5. Estimación de la Diversidad Genética

Los marcadores moleculares son descriptores del tipo de caracterización porque permiten la discriminación fácil entre fenotipos, con métodos numéricos y a partir de datos moleculares es posible obtener la relación entre individuos (Franco e Hidalgo, 2003).

El análisis multivariante refiere a todos aquellos métodos estadísticos que analizan simultáneamente (más de dos variables) de cada individuo. El análisis de conglomerados es un método analítico que se puede aplicar para clasificar las accesiones de un germoplasma en grupos altamente homogéneos con base en alguna similitud existente entre ellas (Franco e Hidalgo, 2003). Un dendograma es una representación gráfica del proceso de agrupamiento para mostrar las relaciones entre los miembros de un grupo; es considerado como una buena representación de una matriz de asociaciones si el coeficiente de correlación es 0.85 o mayor (Rincón *et al.*, 1996).

El agrupamiento jerárquico UPGMA (*Unweighted Pair Group Method with Arithmetic Averages*) define la distancia entre dos grupos como la media de sus distancias entre todos los casos de los cuales un miembro de cada uno proviene de un agrupamiento distinto. Este método usa información acerca de todos los pares de distancia no sólo del más cercano o el más lejano, considera

sólo las distancias entre pares de casos en diferentes grupos (Leps y Smilauer, 2003; Norusis, 2008).

El dendograma resultante presenta el coeficiente de correlación cofenética que mide la correlación entre las interdistancias en el dendograma y las interdistancias en la matriz de distancia sobre la que se aplicó el análisis. La correlación cofenética es la medida del grado de buena clasificación (Cuadras, 1981).

3. MATERIALES Y MÉTODOS

La presente investigación se llevó a cabo en el Campo Experimental de la Facultad de Agronomía de la UANL, ubicado en Marín, Nuevo León, con coordenadas geográficas de 25°52'24.36''N y -100°03'28.97''W y una altitud de 375 msnm. El trabajo se realizó en tres ciclos agrícolas (primavera/verano 2013, otoño/invierno 2013-2014 y primavera/verano 2014).

El análisis molecular se realizó en el Laboratorio de Biotecnología de la Facultad de Agronomía de la UANL, en el Campus de Ciencias Agropecuarias, ubicado en el municipio de General Escobedo, N.L.

3.1. Material Genético

Los catorce genotipos utilizados provienen del Programa de Sorgo de la Facultad de Agronomía de la Universidad Autónoma de Nuevo León, del proyecto *SWEETFUEL* financiado por la Comunidad Europea y consisten en cinco híbridos de sorgo dulce, desarrollados dentro del programa, los progenitores de estos híbridos y dos testigos comerciales aparecen en el Cuadro 1.

Cuadro 1. Características generales del material genético utilizado en el trabajo de investigación.

Nombre	Tipo de material	Altura de planta (cm)	Producción de jugo (ml)	°Bx
FA-UANL33A x KANSAS COLLIER	Híbrido experimental	159	388	6.3
A601 x FA-UANL39	Híbrido experimental	228	679	7.7
FA-UANL 35A x FA-UANL39	Híbrido experimental	181	313	8.5
FA-UANL33A x FA-UANL5	Híbrido experimental	166	419	5.7
FA-UANL37A x FA-UANL5	Híbrido experimental	153	300	6.6
FA-UANL37A	Progenitor Hembra	91	60	9.0
FA-UANL33A	Progenitor Hembra	66	134	6.3
FA-UANL35A	Progenitor Hembra	65	6	4.9
AN601	Progenitor Hembra	93	84	10.5
FA-UANL39	Progenitor Macho	159	242	9.1
FA-UANL5	Progenitor Macho	133	163	8.5
KANSAS COLLIER	Progenitor Macho	147	202	9.4
FORTUNA	Variedad comercial	167	497	9.5
KING GOLD	Híbrido comercial	50	38	4.3

3.2. Caracterización Morfológica

La caracterización morfológica fue previamente realizada por Lazarín *et al.* (2013), se midió glucosa y fructosa con la técnica de cromatografía líquida de alta resolución (HPLC, *High Performance Liquid Chromatography*) por sus siglas en inglés.

3.3. Desarrollo de Marcadores Moleculares

Para el desarrollo de los marcadores moleculares se sembraron 20 semillas por genotipo en dos macetas de dos litros de PVC (Policloruro de vinilo) con *peatmoss* como substrato y se seleccionaron cinco plantas al azar de una edad de 15 días después de la siembra y una altura de 15 cm, medido desde la base del tallo hasta la curva de la última hoja ligulada; se cortaron a nivel de la base del tallo, se colocaron en bolsas individuales y se almacenaron a -70 °C.

La extracción de ADN se realizó con el método de Möller *et al.* (1992) modificado (Chen *et al.*, 2009) descrito en el Anexo 1. La calidad del ADN se determinó por electroforesis en gel de agarosa (INVITROGEN® grado molecular) al 0.8%: se cargaron 10 µl de muestra teñida con SYBR Green® 1:10000, con buffer de carga 1%, con un tiempo de corrida de 120 min a 80 V;

cada una de la muestras se estandarizó a $25 \text{ ng } \mu\text{l}^{-1}$ utilizando buffer TE (100 mM Tris, pH 8.0; 1 mM EDTA).

3.3.1. Generación de Marcadores AFLP

La generación de marcadores AFLP se realizó de acuerdo al protocolo de Vos *et al.* (1995) modificado por Simpson y Gil (2004), mediante el uso del estuche comercial IRDye™ Flourescent AFLP® Kit for Large Plant Genome Analysis de LI-COR® (LI-COR Biosciencias, Lincoln, NE, USA).

Se efectuó la digestión de 100 ng de ADN genómico de las plantas de sorgo dulce utilizando una mezcla de enzimas de restricción *EcoRI/MseI*; se añadió en un tubo de PCR de 0.2 ml a un volumen final de reacción de 12.5 μl , los componentes se mezclaron cuidadosamente, se centrifugaron e incubaron a 37 °C por dos horas y 15 minutos a 70 °C en un termociclador PCR Express (Thermo Hybaid®). Al ADN digerido de cada muestra se le ligaron los adaptadores con 2.5 U de *T4* ADN ligasa para un volumen final de 25 μl , los componentes se mezclaron, centrifugaron brevemente e incubaron a 20 °C por dos horas en un termociclador PCR Express (Thermo Hybaid®). Posteriormente se realizó una dilución 1:10 de la mezcla de ADN digerido y ligado en agua libre de nucleasas.

La mezcla de reacción de pre-amplificación consistió en: MgCl_2 3 mM, 20 μl de Mezcla de pre amplificación AFLP, Solución amortiguadora de reacción 1X,

Taq ADN polimerasa 2.5 U, 2.5 μl de ADN digerido, ligado y diluido 1:10, en un volumen final de 27 μl , la reacción se incubó durante 20 ciclos a 94 °C por 30 s, 56 °C por 60 s y 72 °C por 60 s. Posterior a la preamplificación se llevó a cabo una dilución 1:40.

Para la amplificación selectiva se utilizó el protocolo de PCR *dúplex* (un iniciador de *Msel* con dos iniciadores *EcoRI* marcados con fluorescencia a 700 y 800 ηm) en un volumen total de reacción de 11 μl , los componentes de la reacción (Cuadro 2) se colocaron en tubos para PCR de 0.2 ml mezclándolos y centrifugándolos fugazmente, se evitó exponerlos a la luz por largos períodos de tiempo.

Cuadro 2. Componentes utilizados en la mezcla de reacción de pre amplificación del ADN genómico digerido de sorgo dulce de la FA-UANL.

Componentes	Volumen (μl)
Solución de trabajo de la <i>Taq</i> ADN polimerasa (MgCl_2 3 mM, Solución amortiguadora de amplificación 1X, <i>Taq</i> ADN polimerasa 0.1 U μl^{-1}).....	6.0
ADN digerido, pre amplificado y diluido 1:40.....	2.0
Iniciador <i>Msel</i> (contiene dNTPs).....	2.0
Iniciador A marcado a 700 ηm <i>EcoRI</i> IRDye.....	0.5
Iniciador B marcado a 800 ηm <i>EcoRI</i> IRDye.....	0.5
Volumen Total	11.0

Se realizaron reacciones *dúplex* para 30 combinaciones de iniciadores cuya información se encuentra descrita en el Anexo 2. El programa térmico utilizado para la amplificación selectiva consistió de un ciclo de 94 °C por 30 s, 65 °C por 30 s y 72 °C por 60 s; doce ciclos en donde subsecuentemente se disminuyó la

temperatura de hibridación (65 °C) por 0.7 °C por ciclo mientras se mantiene a 94 °C por 30 s para desnaturalizar y 60 s a 72 °C para extender las cadenas; 23 ciclos por 30 s, 56 °C por 30 segundos y 72 °C por 60 s. La reacción se mantuvo a 4 °C hasta que fue utilizada para la electroforesis.

Los productos de amplificación AFLP se separaron en un gel de poliacrilamida al 6.5% en un secuenciador LI-COR IR² 4200 (LI-COR Biosciences, Lincoln, NE); la electroforesis se corrió por 3 h a 1500 V, 40 W, con una corriente de 40 mA, 45 °C y una velocidad de escáner de cuatro. Para cargar los productos de amplificación se añadió 5 µl de la solución amortiguadora de carga azul de bromofenol y se desnaturalizaron durante tres minutos a 94 °C. Posteriormente se generaron las matrices de datos analizando las imágenes obtenidas con el programa SAGA^{MX}, se seleccionaron bandas definidas registrándose como datos binarios, utilizando el número uno para indicar la presencia de una banda y cero para indicar su ausencia.

3.3.2. Generación de Marcadores RAPD

Para la realización de los marcadores RAPD se probaron 180 iniciadores pertenecientes al juego de OP-A, OP-B, OP-C, OP-D, OP-E, OP-F, OP-G, OP-I y OP-K. Dichos iniciadores fueron sintetizados por Eurofins MWG Operon® (<http://www.operon.com/products/downloads/OperonsRAPD10merSequences.xls>).

Se siguió el protocolo de Williams *et al.* (1990) modificado por Prakash *et al.* (2006); la reacción de PCR consistió en 25 ng de ADN genómico, iniciador OP-X 0.25 mM, 0.25 mM Mezcla de deoxinucleótidos dNTPs Invitrogen[®], 0.02 U de enzima *Taq* Polimerasa Promega[®], 2.5 mM MgCl₂ y buffer para *Taq* 1X en un volumen final de reacción de 15 µl. El programa térmico de reacción utilizado fue de un ciclo inicial de 94 °C; 45 ciclos de 94 °C por 60 s, Tm* (Temperatura de alineación) por 60 s, 72 °C por 2 min, y un ciclo final de 72 °C por siete min. La reacción se mantuvo a 4 °C hasta que fue cargada para electroforesis. La Tm* se ajustó para cada iniciador específico y se calculó mediante la fórmula: $Tm^* = (wA + xT) \cdot 2 + (yG + zC) \cdot 4$, donde w, x, y, z es el número de bases A, T, G, C en la secuencia, respectivamente.

Los productos de amplificación se separaron y visualizaron por medio de electroforesis en gel de agarosa (INVITROGEN[®] grado molecular) al 3%, se cargaron el total de cada muestra teñida con SYBR Green[®] 1:10000, con un tiempo de corrida de 180 min a 90 V. Los geles se fotodocumentaron con el equipo BioDoc-It[™] UVP[®] para su posterior análisis; se construyó una matriz de datos binarios donde cero representa la ausencia de la banda y uno representa la presencia de la banda.

Para la pruebas de reproducibilidad se seleccionaron cinco genotipos contrastantes en el contenido de glucosa y fructosa FA-UANL 35A x FA-UANL 39, FA-UANL 33A x FA-UANL 5, FA-UANL 37A, FA-UANL 39 y FORTUNA. Se

realizaron cuatro extracciones de plantas diferentes independientes, así como cinco PCR en tiempos diferentes. Se realizaron con los iniciadores que mayor bandeo obtuvieron: OP-A 3 (T_m 32 °C); OP-F 3 (T_m 39.5 °C); OP-E 20 (T_m 39.5 °C); OP-I 4 (T_m 43.6 °C) y OP-I 14 (T_m 43.6 °C). Se calculó el porcentaje de bandas presentes en la reacción original con respecto a los individuos (extracciones diferentes) y las repeticiones de la reacción de PCR.

3.3.3. Selección de Marcadores RAPD Candidatos a Secuenciación

Se alinearon los genotipos altos en azúcares (AN601 x FA-UANL 39, FA-UANL 35A x FA-UANL 39 y FA-UANL 33A x FA-UANL 5) y los genotipos más bajos en azúcares (FA-UANL 37A, FA-UANL 33A y FA-UANL 35A) para la identificación de los marcadores RAPD en común.

Las bandas polimórficas RAPD seleccionadas se extrajeron del gel de agarosa y se purificaron mediante el protocolo de extracción de aislamiento de fragmentos de ADN utilizando geles de agarosa de bajo punto de fusión (Moore *et al.*, 2002). El ADN aislado se reamplificó por PCR y se cuantificó, se realizó una electroforesis en agarosa para comparar peso y calidad respecto a la banda previamente seleccionada.

3.4. Descriptiva de Marcadores

La descripción de datos de marcadores permite determinar el número de Muestras que se refiere a la cantidad de observaciones, Muestras duplicadas determina si existen observaciones que tengan el mismo perfil a través de todos los marcadores o bandas. El patrón de bandas duplicadas determina si existen marcadores (bandas) que presenten el mismo perfil a través de todas las observaciones. Las Bandas Monomórficas establece el número de marcadores que no varían a través de todo el perfil de observaciones, es decir aquellos marcadores que, por ejemplo, son siempre cero o son siempre uno. Las Bandas Polimórficas (%) calcula el porcentaje de marcadores polimórficos para cada iniciador (Balzarini *et al.*, 2006).

Para cuantificar el polimorfismo también es común utilizar el contenido de información polimórfica (PIC). Un valor de PIC alto indica mayor contenido de información de ligamiento. El PIC se calculó mediante la metodología de Botstein *et al.* (1980), mediante la ecuación:

$$PIC = 1 - \left(\sum_{i=1}^n p_i^2 \right) - 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^n p_i^2 p_j^2$$

donde p_i y p_j son las frecuencias de los n alelos.

Cuando el número de alelos es grande, el contenido de información polimórfica se aproxima a la heterocigosis. Cuando todos los alelos tienen igual frecuencia el PIC es menor que la heterocigosis (Balzarini *et al.*, 2006).

La probabilidad de que dos individuos (muestras) compartan el mismo alelo (PDICMA) es calculada como $P_i(X_D)^n$, donde (X_D) representa el índice de similitud promedio para todos los pares de comparaciones, calculado como $X_D=(2N_{AB})/(N_A+N_B)$, donde N_{AB} representa el número total de bandas en las muestras A y B respectivamente y n representa el número promedio de productos de amplificación por muestra (Wetton *et al.*, 1987; Balzarini *et al.*, 2006)

3.5. Análisis estadístico

Con las matrices de datos binarios obtenidas se hizo la caracterización genética de las muestras utilizando el software para análisis estadístico de datos genéticos *Info-Gen* (Balzarini *et al.*, 2008).

Se calcularon las matrices de distancia con lo datos morfológicos (glucosa y fructuosa en diferentes estados fenológicos de la planta) por medio de distancia Euclidiana:

$$d_{ij} = \left[\sum_{k=1}^m |x_{ik} - x_{jk}|^2 \right]^{1/2}$$

donde:

m = número de variables o características observadas

x_{ik} = es la i -ésimo observación del j -ésimo objeto

x_{jk} = ponderación para la k -ésima variable entre el j -ésimo objeto

así como las matrices para los datos genéticos provistos por los AFLP y RAPD por medio de la distancia de Dice (1945):

$$d_{ij} = \left[1 - \left(\frac{2a}{2a + b + c} \right)_{ij} \right]^{1/2}$$

donde a , b y c son frecuencias absolutas para eventos (1,1), (1,0) y (0,1), respectivamente; considerando que no tiene en cuenta el evento de ausencia simultánea (0,0).

Se llevó a cabo una prueba de Mantel (1967) para obtener la correlación que existía entre las matrices y distancias y probar, también, si una pudiese estar afectando a la otra.

El agrupamiento jerárquico se realizó por un análisis de conglomerados utilizando el algoritmo UPGMA (encadenamiento promedio) que define las distancias entre dos grupos o conglomerados como el promedio de todas las distancias de pares entre elementos de un grupo y elementos del otro grupo. Se calcula la media aritmética de las similitudes entre todos los elementos de los

dos grupo que se van a fusionar. Hay que promediar todas las distancias entre los elementos de AB y de C:

$$d_{(AB,C)} = \frac{\sum_{ij} d_{ij}}{n_{AB}n_C}$$

donde d_{ij} es la distancia entre el elemento i del grupo AB y el j de C; n_{AB} y n_C son los números de elementos en el grupo AB y C, respectivamente.

La correlación cofenética se calcula sea $D=(d_{ij})$ la matriz de disimilitudes y h la jerarquía indexada (dendograma) construida por algún método de clasificación. El índice d de h define la distancia ultramétrica

$$u_{ij} = d(h) \text{ si } h \text{ es la menor clase que contiene } i, j$$

se verifica $d_{ij} = u_{ij}$ si y sólo si la disimilaridad inicial es también ultramétrica, por lo tanto, el grado de distorsión entre la matriz d_{ij} y u_{ij} nos puede dar a medida de la calidad de la clasificación jerárquica obtenida. El coeficiente de correlación r_c entre los $n(n-1)/2$ pares de distancias (d_{ij}, u_{ij}) verifica $0 < r_c < 1$. Cuando r_c es próximo a 1, existe una clara estructura jerárquica entre los objetos. Valores bajos de r_c indican una distorsión notable entre las disimilaridades iniciales y las que resultan en el dendograma (Cuadras, 1981).

Para la generación de la matriz de distancias incluyendo datos morfológicos y moleculares se empleó el coeficiente de similaridad propuesto por Gower (1971) que permite analizar datos continuos y datos discretos (Yan, 2003; Levitus *et al.*, 2010).

$$S_{ij} = \frac{\sum_{c=1}^m W_{ijc} S_{ijc}}{\sum_{K=1}^p W_{ijc}}$$

donde:

S_{ij} = similaridad entre el i -ésimo y j -ésimo objeto

m = número de variables o características observadas

W_{ijc} = ponderación para la c -ésima variable entre el i -ésimo y j -ésimo objeto

Con esta nueva matriz de distancias se construyó un dendograma con el fin de identificar aquellos genotipos agrupados en función de caracteres morfológicos y datos moleculares. Se identificaron los marcadores AFLP y RAPD en común que coincidieran con los genotipos altos en azúcares.

3.6. Generación de Marcadores SCAR

El producto final de PCR fue secuenciado por el servicio SimpleSeq™ de Eurofins MWG Operon®.

Las secuencias se analizaron con el programa CodonCode Aligner® Versión 5.0.1 (LI-COR Inc., 2014), con el cual se definieron cuales secuencias estaban leídas correctamente y se realizaron alineaciones entre las secuencias del mismo iniciador para comparar entre genotipos, obteniéndose 10 agrupamientos. Las secuencias también se utilizaron para buscar en una base de datos de nucleótidos utilizando una consulta de nucleótidos con la

herramienta en línea *nucleotide* BLAST

(http://blast.ncbi.nlm.nih.gov/Blast.cgi?CMD=Web&PAGE_TYPE=BlastHome)

para corroborar que dichas secuencias pertenecieran a sorgo.

Con base en la secuencia de ADN de los agrupamientos, se diseñaron un oligonucleótido *forward* y uno *reverse* con el fin de tener: a) una T_m experimental igual o superior a 50 °C, b) longitud de entre 20 y 22 pares de bases, c) un extremo 3' rico en bases GC y d) una autocomplementariedad menor a 3 bases (Kalendar *et al.*, 2011). Se seleccionaron iniciadores que alinearan sobre los amplicones correspondientes a los iniciadores RAPD y la secuencia interna de cada fragmento RAPD. El diseño se realizó mediante la herramienta en línea *Primer* BLAST (<http://www.ncbi.nlm.nih.gov/tools/primer-blast>). El par de iniciadores seleccionados se sintetizaron en Eurofins MWG Operon®.

Con la finalidad de repetir la secuenciación y garantizar que la información fuera correcta se realizaron seis amplificaciones por iniciador *forward* en 20 µl de mezcla de reacción que contenía 20 ηg de DNA molde, 1X PCR buffer, 3.5 mM MgCl₂, 0.2 µM de iniciador, 200 µM dNTPs Mezcla de deoxinucleótidos Invitrogen®, 0.02 U de enzima *Taq* Polimerasa Promega®. Las condiciones del programa térmico fueron de 94 °C durante 3 min, seguido de 45 ciclos a 94 °C durante 60 s, temperatura de hibridación durante 45 s y 72 °C durante 90 s; y una extensión final a 72 °C durante 10 min.

Para la electroforesis de las reacción con los iniciadores SCAR se optó por gel de acrilamida: bis-acrilamida (40:1) al 8% para obtener una separación de las bandas con mayor definición. Se cargaron 7 μ l de muestra, con buffer de carga de glicerol al 1%, con un tiempo de corrida de 240 min a 1500 V para las PCR del ADN genómico total con el juego de iniciadores. Se realizó la tinción del gel con AgNO_3 (0.02%) según el protocolo descrito por Sambrook *et al.* (1989).

4. RESULTADOS Y DISCUSIÓN

4.1. Extracción de ADN

La extracción de ADN permitió obtener concentraciones desde $95.9 \text{ ng } \mu\text{l}^{-1}$ a $1568 \text{ ng } \mu\text{l}^{-1}$, con un promedio de $401 \text{ ng } \mu\text{l}^{-1}$ y un rendimiento promedio de 8.03 ng por mg de tejido fresco vegetal. La relación $260/280 \text{ nm}$ fue de 1.63 para la más baja y 1.92 para la más alta, con una media de 1.82 , por lo que el ADN obtenido fue puro. La calidad del ADN se puede observar en la Figura 1.

Weising *et al.* (2005) establecen que el rango de rendimiento obtenidos con protocolos estándar para plantas está entre menos de $1 \text{ ng } \text{mg}^{-1}$ a $200 \text{ ng } \text{mg}^{-1}$, lo cual muestra que los rendimientos obtenidos en este estudio se encontraron dentro de lo normal. La contaminación con otros componentes celulares como proteínas, RNA, membranas, polisacáridos y polifenoles es definida con la absorbancia de luz en un rango de longitud de onda de 260 a 280 nm para el ADN, una preparación de ADN pura muestra una relación $260/280 \text{ nm}$ de 1.8 a 2.0 . La calidad determina que el aislamiento de ADN fue intacto y compacto si se encuentra en un rango de 100 a $10,000 \text{ Kb}$ (llamado ADN megabase) y ésta es esencial para estudios de huella genética.

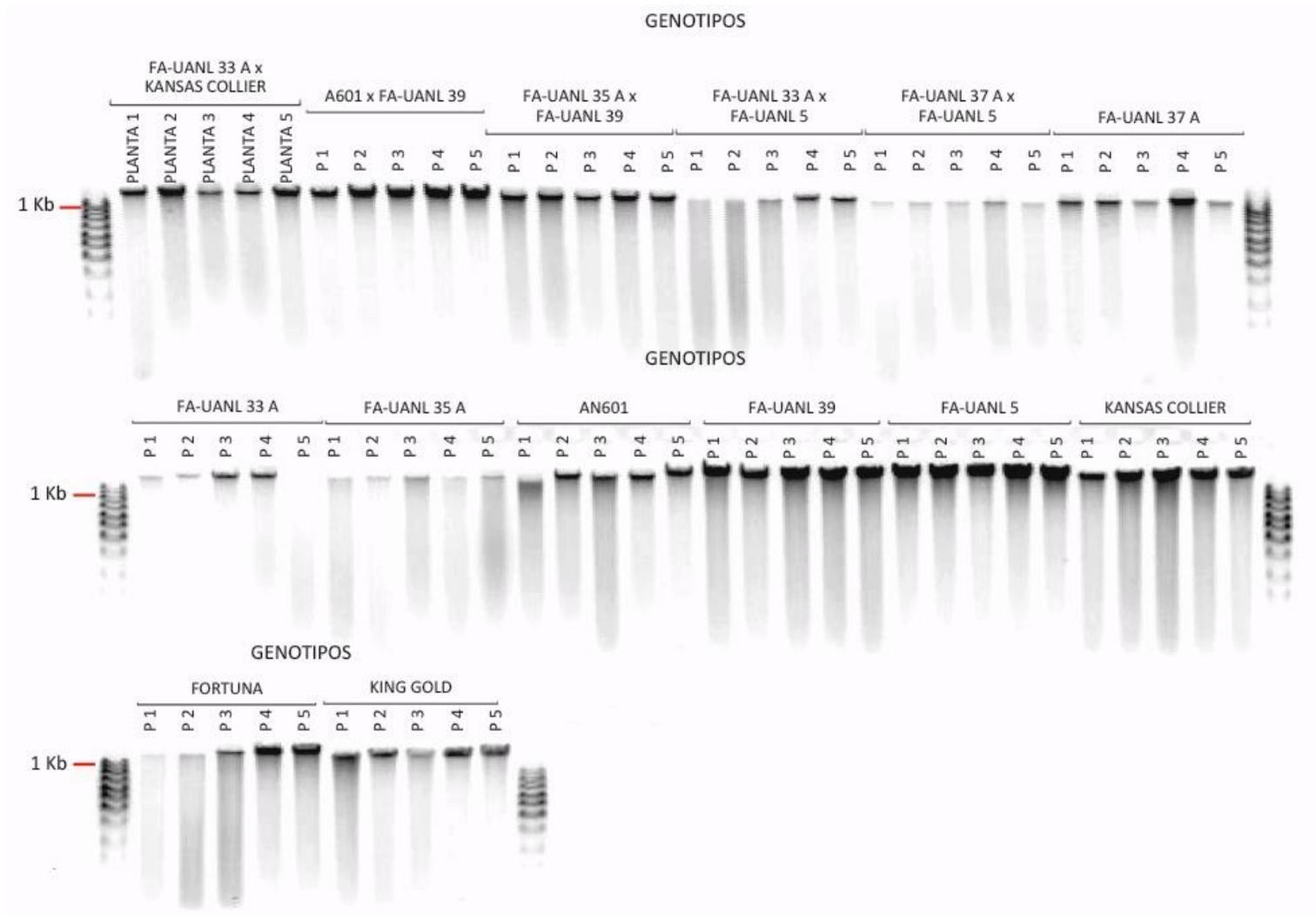


Figura 1. Electroforesis en gel de agarosa, extracciones ADN de tejido foliar de plantas de sorgo dulce FA-UANL.

4.2. Generación de Marcadores AFLP

Las 30 combinaciones de iniciadores *MseI/EcoRI* utilizadas produjeron fragmentos de amplificación AFLP con un rango de bandeo promedio de 60 pb a 352 pb; se seleccionó un total de 549 bandas, dando un promedio de 18.3 bandas por iniciador, de las cuales 75 fueron monomórficas y 474 polimórficas lo que representa un 86.34%. Se encontraron 334 patrones de banda duplicados.

Ritter *et al.* (2007) generaron la huella genética de 95 líneas de sorgo mediante AFLP con un promedio de 37 bandas por par de iniciadores encontrando 46.3% de bandas polimórficas entre dos o más líneas; éstos datos fueron muy similares a los registrados por Shegro *et al.* (2013) con 31 bandas por par de iniciadores y 43.1% de bandas polimórficas.

El Cuadro 3 muestra la información descriptiva para datos AFLP. Los valores obtenidos de la información polimórfica oscilaron entre 0.18 a 0.33, lo que indica que todas las combinaciones son útiles para estudios de huella genética (Schut *et al.*, 1997; Ritter *et al.*, 2007). Estos resultados fueron menores a Geleta *et al.* (2006) quiénes encontraron un valor PIC medio de 0.65 en un estudio realizado a 45 genotipos de sorgo utilizando AFLP y 0.46 con marcadores SSR (*Sequence Simple Repetition*). Por su parte Ritter *et al.* (2007) y Shegro *et al.* (2013) reportaron un PIC de 0.48 y 0.178, respectivamente.

Cuadro 3. Descriptiva de datos marcadores AFLP.

Iniciador	BP	BM	BT	PIC	AMP	PDICMA
*M-CAA x E-AAG	16	0	16	0.24	78.85	3.20E-12
*M-CAA x E-AGG	25	0	25	0.22	74.77	1.10E-10
*M-CAC x E-AAC	16	2	18	0.27	59.83	3.80E-10
*M-CAC x E-ACT	10	1	11	0.27	67.13	2.10E-10
*M-CAG x E-ACA	22	3	25	0.21	49.85	4.30E-08
*M-CAG x E-ACG	16	0	16	0.25	38.94	5.90E-11
*M-CAT x E-ACC	9	7	16	0.30	75.48	1.70E-06
*M-CAT x E-AGC	15	9	24	0.28	67.95	1.00E-06
*M-CTA x E-ACA	21	2	23	0.25	59.20	8.30E-11
*M-CTA x E-ACG	22	0	22	0.24	54.20	2.50E-10
*M-CTC x E-AAC	15	4	19	0.26	59.51	7.10E-08
*M-CTC x E-ACT	24	4	28	0.24	53.02	1.30E-07
*M-CTG x E-AAG	12	6	18	0.26	56.41	4.90E-07
*M-CTG x E-AGC	19	4	23	0.28	50.17	3.30E-09
*M-CTT x E-ACC	21	5	26	0.25	42.01	7.90E-08
*M-CTT x E-AGG	18	5	23	0.25	56.19	1.90E-07
*M-CTA x E-AAG	11	0	11	0.22	78.32	1.50E-09
*M-CTA x E-AGG	18	0	18	0.24	55.56	4.50E-11
*M-CAG x E-AAG	15	2	17	0.30	56.56	4.90E-12
*M-CAG x E-AGG	16	4	20	0.26	68.85	4.30E-08
*M-CAC x E-ACA	15	4	19	0.25	63.16	2.10E-07
*M-CAC x E-AGC	6	7	13	0.33	77.51	8.50E-06
*M-CAT x E-AAC	11	3	14	0.29	54.40	2.00E-11
*M-CAT x E-AGG	17	1	18	0.29	49.15	7.90E-12
*M-CAA x E-ACA	8	1	9	0.30	48.72	4.80E-11
*M-CAA x E-ACG	12	0	12	0.18	64.74	1.40E-08
*M-CTC x E-ACC	23	0	23	0.24	43.81	4.50E-11
*M-CTC x E-AGC	11	0	11	0.27	61.54	6.20E-13
*M-CTG x E-AAC	16	0	16	0.22	49.04	7.60E-10
*M-CTG x E-ACG	14	1	15	0.22	50.26	7.00E-09
Total	474	75	549	0.26	58.05	4.16E-07

Nota: BM=Bandas monomórficas; BP= Bandas polimórficas; BT=Bandas totales; PIC= Índice de contenido polimórfico; AMP=Porcentaje de amplificación; PDICMA= Probabilidad de que dos individuos compartan el mismo alelo; E=Exponencial.

La mayor amplificación se obtuvo con la combinación M-CAA x E-AAG que amplificó en un 78.85% de las muestras, en contraste con el 38.94% de la combinación M-CAG x E-ACG, que fue la de menor amplificación. La menor probabilidad de que dos individuos compartan el mismo alelo por iniciador fue

encontrado en la combinación M-CTC x E-AGG, siendo el de mayor confianza en la identificación de individuos.

El conocimiento de la diversidad genética de las colecciones de germoplasma de sorgo es importante para el aprovechamiento eficiente de su potencial genético, para su uso en mejoramiento genético, mantenimiento y para la conservación. El análisis de los marcadores AFLP para la diversidad genética ha sido utilizado para complementar los datos agronómicos cuando la información es limitada o inexistente (Ritter *et al.*, 2007; Iqbal *et al.*, 2010; Pecina *et al.*, 2012; Elangovan *et al.*, 2013 y Lekgari y Dweikat, 2014) o bien, para el diseño de iniciadores SCAR en la identificación de resistencia genética al áfido del sorgo *Melanaphis sacchari* (Zehntner) (Chang *et al.*, 2012).

En el presente estudio los fragmentos de ADN de marcadores de AFLP fueron tan pequeños (60-352 pb) que no fueron apropiados para el diseño de iniciadores SCAR ya que en el proceso de conversión, muchos marcadores perderían su especificidad de secuencia o no podrían ser utilizados para la amplificación de ADN genómico (Horejsi *et al.*, 1999).

En la Figura 2 se muestra como ejemplo, el perfil de bandas amplificadas con las combinaciones de iniciadores M-CTA x E-ACA y M-CTC x E-AAC; mientras que en la Figura 3 se observan las combinaciones de iniciadores M-CTA x E-ACG y M-CTC x E-ACT.

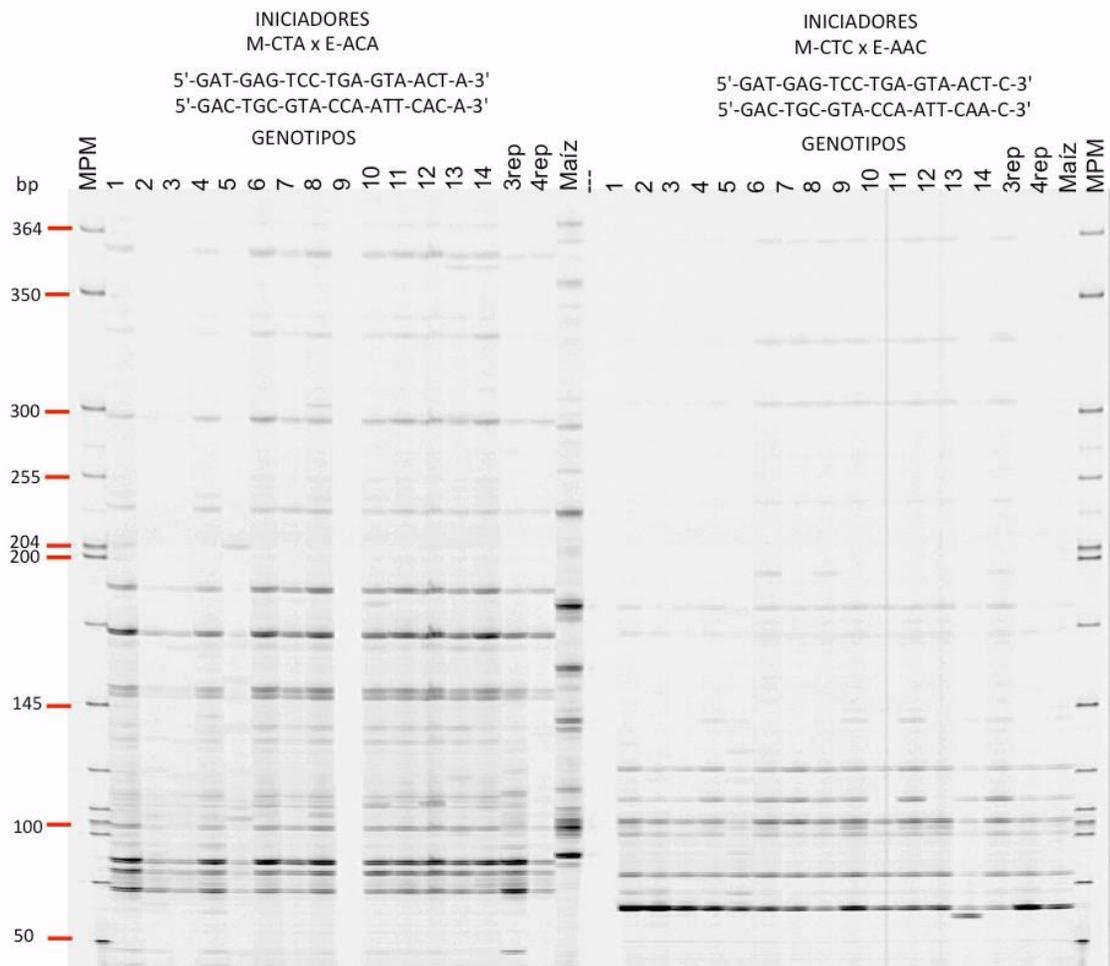


Figura 2. Electroforesis en gel de poliacrilamida 6.5% de marcadores AFLP obtenidos con las combinaciones de iniciadores M-CTA x E-ACA y M-CTC x E-AAC. El carril MPM corresponde al marcador de peso molecular; los genotipos son: 1) FA-UANL 33A x KANSAS COLLIER; 2) A601 x FA-UANL 39; 3) FA-UANL 35A x FA-UANL 39; 4) FA-UANL 33A x FA-UANL 5; 5) FA-UANL 37A x FA-UANL 5; 6) FA-UANL 37A; 7) FA-UANL 33A; 8) FA-UANL 35A; 9) AN601; 10) FA-UANL 39; 11) FA-UANL 5; 12) KANSAS COLLIER; 13) FORTUNA; 14) KING GOLD; 3rep) FA-UANL 35A x FA-UANL 39; 4rep) FA-UANL 33A x FA-UANL 5 y Maiz) *Zea mays*.

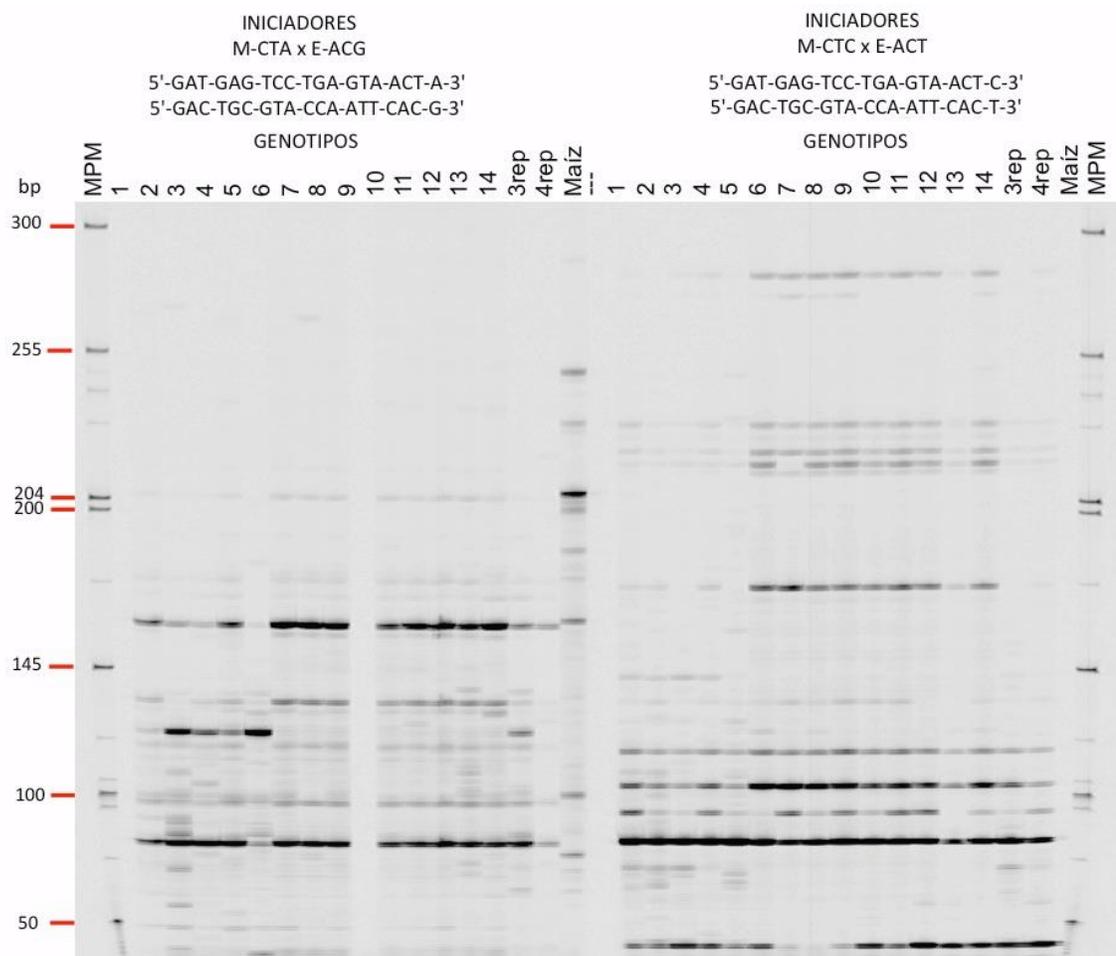


Figura 3. Electroforesis en gel de poliacrilamida 6.5% de marcadores AFLP obtenidos con las combinaciones de iniciadores M-CTA x E-ACG y M-CTC x E-ACT. El carril MPM corresponde al marcador de peso molecular; los genotipos son: 1) FA-UANL 33A x KANSAS COLLIER; 2) A601 x FA-UANL 39; 3) FA-UANL 35A x FA-UANL 39; 4) FA-UANL 33A x FA-UANL 5; 5) FA-UANL 37A x FA-UANL 5; 6) FA-UANL 37A; 7) FA-UANL 33A; 8) FA-UANL 35A; 9) AN601; 10) FA-UANL 39; 11) FA-UANL 5; 12) KANSAS COLLIER; 13) FORTUNA; 14) KING GOLD; 3rep) FA-UANL 35A x FA-UANL 39; 4rep) FA-UANL 33A x FA-UANL 5 y Maíz) *Zea mays*.

4.3. Generación de Marcadores RAPD

De los 180 iniciadores RAPD se encontraron 52 iniciadores que amplificaron un total de 256 bandas con un rango entre 220 y 2800 pb, una media de 4.92

bandas por iniciador, 6.5% de ellas monomórficas, dando un 93.5% de bandas polimórficas y 179 patrones de bandas duplicados. Se obtuvo una media de 0.22 de PIC y 17.34% de amplificación por iniciador; lo que indica que las condiciones de extracción de ADN y PCR fueron adecuadas (Weir *et al.*, 1996; Prakash *et al.*, 2006). En la Figura 4 se observan los productos de amplificación para los iniciadores OP-A02, OP-C 14, OP-C 19 y OP-K 15. En el Anexo 3 se encuentran la descriptiva de datos de los marcadores RAPD.

Anteriormente, Jaiswal *et al.* (1998) ya habían generado 254 bandas utilizando 38 iniciadores, con una media de 6.68 bandas por iniciador, de las cuáles sólo 65% fueron polimórficas y un PIC medio de 0.234, en un rango de amplificación de 400 a 2500 pb; siendo los resultados del presente estudio muy similares. Sin embargo, difieren de lo mencionado por Prakash *et al.* (2006) quienes presentaron 12.7 bandas 100% polimórficas por iniciador en un estudio entre 32 genotipos de sorgo usando marcadores RAPD. Posteriormente, Iqbal *et al.* (2010) obtuvieron 125 bandas con un 95% de polimorfismo y una media de 7.1 bandas por iniciador, pero con un PIC más alto de 0.36, sirviendo esto como información pionera en un intento de establecer la huella genética del germoplasma presente en distintas zonas ecológicas de Pakistán.

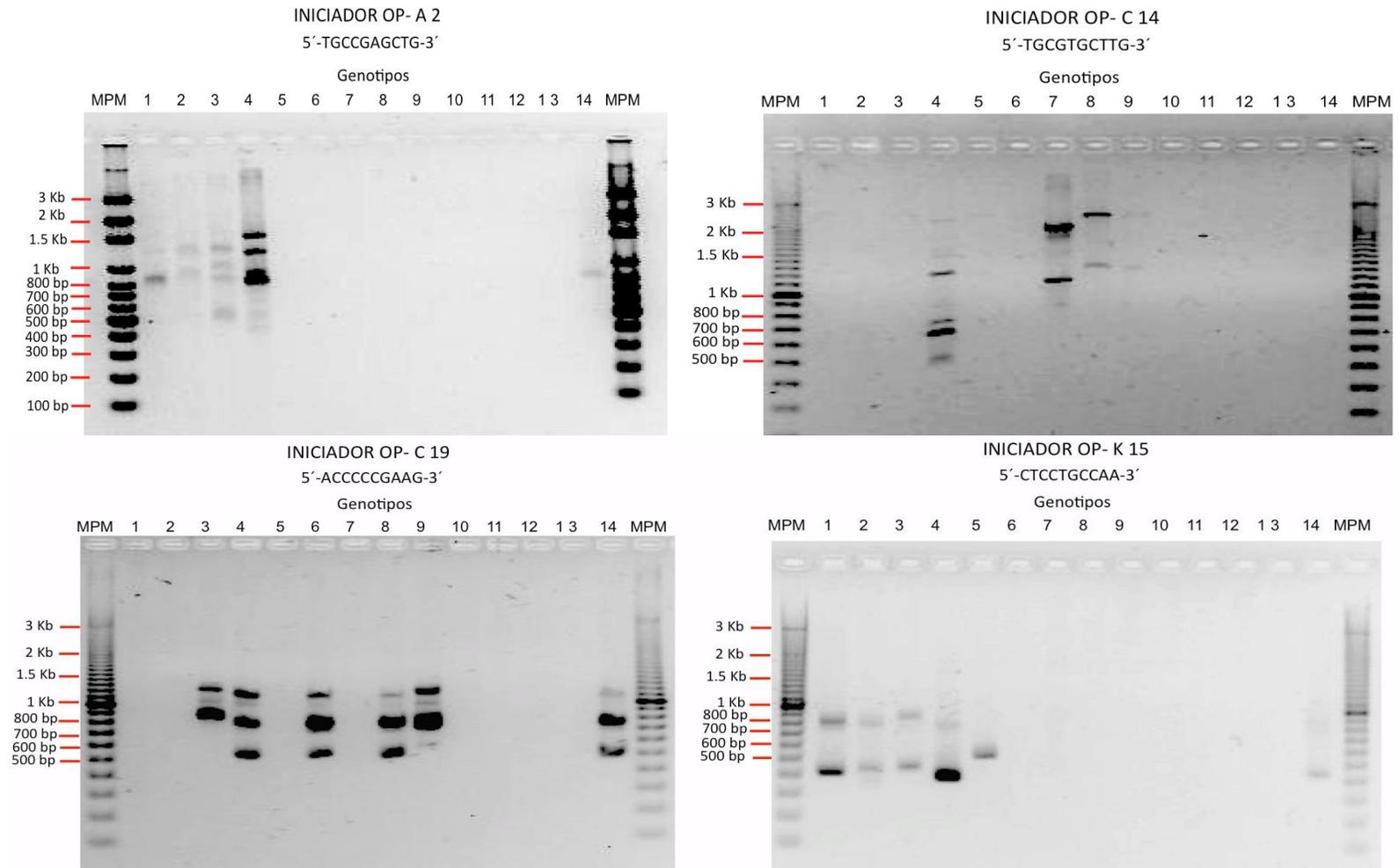


Figura 4. Electroforesis en gel de agarosa al 3% para productos de amplificación para los iniciadores OP-A 2, OP-C 14, OP-C 19 y OP-K 15. El carril MPM corresponde al marcador de peso molecular; los genotipos son: 1) FA-UANL 33A x KANSAS COLLIER; 2) A601 x FA-UANL 39; 3) FA-UANL 35A x FA-UANL 39; 4) FA-UANL 33A x FA-UANL 5; 5) FA-UANL 37A x FA-UANL 5; 6) FA-UANL 37A; 7) FA-UANL 33A; 8) FA-UANL 35A; 9) AN601; 10) FA-UANL 39; 11) FA-UANL 5; 12) KANSAS COLLIER; 13) FORTUNA y 14) KING GOLD.

4.3.1. Reproducibilidad de los resultados en marcadores RAPD

Se obtuvo un 71.3% de reproducibilidad del bandeo original respecto a las reacciones posteriores; 88.8% de los marcadores fueron reproducibles entre las diferentes PCR, este valor fue mayor al de 76% reportado por Skroch y Nienhuis (1995); además, un 92% fue reproducible entre las extracciones diferentes, lo que significa que la técnica puede aplicarse dentro de un genotipo utilizando una o varias plantas, ya que las bandas amplifican de igual manera (Pammi *et al.*, 1994). Las Figuras 5 y 6 corresponden a los productos de amplificación para los iniciadores OP-E 20 con el genotipo FA-UANL 33A x FA-UANL 5 y OP-I 14 con el genotipo FA-UANL 35A x FA-UANL 39, respectivamente.

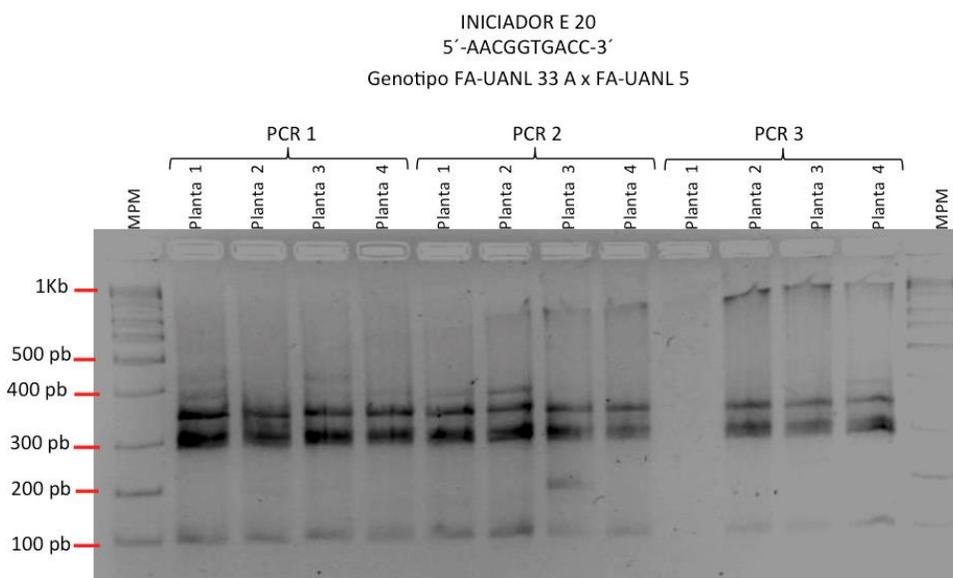


Figura 5. Electroforesis en gel de agarosa al 3% para productos de amplificación para el iniciador OP-E 20 con el genotipo FA-UANL 33A x FA-UANL 5. El carril MPM corresponde al marcador de peso molecular; cada carril corresponde a una planta diferente en diferente PCR por genotipo y por iniciador.

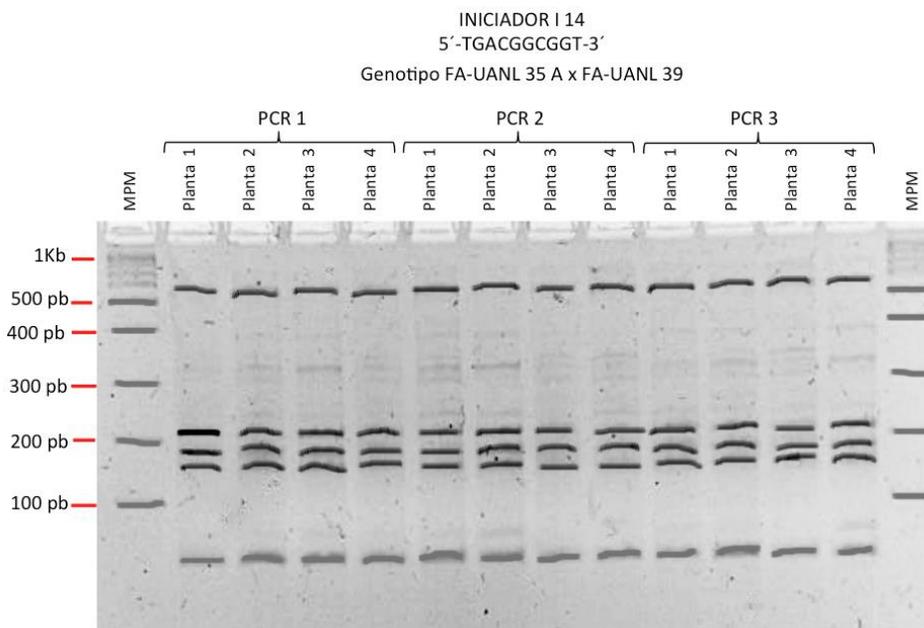


Figura 6. Electroforesis en gel de agarosa al 3% para productos de amplificación para el iniciador OP-I 14 con el genotipo FA-UANL 35A x FA-UANL 39. El carril MPM corresponde al marcador de peso molecular; cada carril corresponde a una planta diferente en diferente PCR por genotipo y por iniciador.

4.4. Selección de Marcadores RAPD Candidatos a Secuenciación

Se logró identificar 53 marcadores RAPD que tenían la característica de estar presentes en los genotipos altos en azúcares (AN601 x FA-UANL 39, FA-UANL 35A x FA-UANL39 y FA-UANL 33A x FA-UANL 5) mientras que estaban ausentes en los genotipos más bajos en azúcares (FA-UANL 37, FA-UANL 33 y FA-UANL 35). El Anexo 4 muestra la identificación de los fragmentos secuenciados.

4.5 Análisis estadístico

Se realizó la Prueba de Mantel (1967) con la cual se obtuvieron las siguientes matrices de distancias: 1) Distancia Euclidiana para Azúcares, 2) Distancia Dice (1-S) para AFLP, 3) Distancia Dice (1-S) para RAPD y 4) Distancia Dice (1-S) para AFLP y RAPD. Las correlaciones entre las matrices se muestran en el Cuadro 4.

Una correlación cercana a uno, indica que la información de las distancias obtenidas con dos tipos de datos proporciona la misma información; una correlación cercana a cero indica que las matrices proporcionan información diferente y por lo tanto complementaria (Balzarini *et al.*, 2006).

La correlación entre la matriz de Distancia Euclidiana para Azúcares y la Distancia Dice (1-S) para RAPD obtuvo un valor de 0.11 lo que estadísticamente no es significativo (p -Valores mayores a 0.05 no son significativos), puede ser que las distancias obtenidas con los marcadores RAPD respecto a los marcadores morfológicos para azúcares sean ligeramente poco informativo contrario a la correlación entre la matriz de Distancia Euclidiana para Azúcares y la Distancia Dice (1-S) para AFLP que tuvo un valor de -0.03.

Cuadro 4. Coeficientes de correlación lineal entre las matrices.

Matriz	1.- Distancia Euclidiana para Azúcares	2.- Distancia Dice (1-S) para AFLP	3.- Distancia Dice (1-S) para RAPD	4.- Distancia Dice (1-S) para AFLP y RAPD
1.- Distancia Euclidiana para Azúcares	1			
2.- Distancia Dice (1-S) para AFLP	-0.03	1		
3.- Distancia Dice (1-S) para RAPD	0.11	0.07	1	
4.- Distancia Dice (1-S) para AFLP y RAPD	0.01	0.97	0.09	1

Valores de significancia (p) para los coeficientes de correlación lineal entre matrices *

1.- Distancia Euclidiana para Azúcares	-----			
2.- Distancia Dice (1-S) para AFLP	0.493	-----		
3.- Distancia Dice (1-S) para RAPD	0.280	0.364	-----	
4.- Distancia Dice (1-S) para AFLP y RAPD	0.388	0.0001	0.279	-----

* p -Valores > 0.05 no significativo

Estos datos pueden explicarse por la proporción de marcadores obtenidos con ambas técnicas, ya que una alta significancia se alcanza con 0.01 al unir ambos marcadores moleculares y compararlos contra los datos morfológicos. Las Figuras 7 y 8 muestran ejemplos de gráficas de correlaciones entre las matrices.

Geleta *et al.* (2006) mencionaron en su estudio que los coeficientes de correlación obtenidos por la prueba de Mantel que se usó para comparar las matrices de los diferentes marcadores, mostraron que los valores estimados de la relación genética para marcadores AFLP y SSR, así como para los marcadores morfológicos y SSR se relacionaron significativamente ($p < 0.001$). Sin embargo, los datos morfológicos y AFLP mostraron correlación no significativa ($p > 0.05$). Ambos conjuntos de datos de AFLP y SSR permitieron que todas las accesiones se identifiquen de forma única; siendo que dos accesiones no podían ser distinguidas por los datos morfológicos.

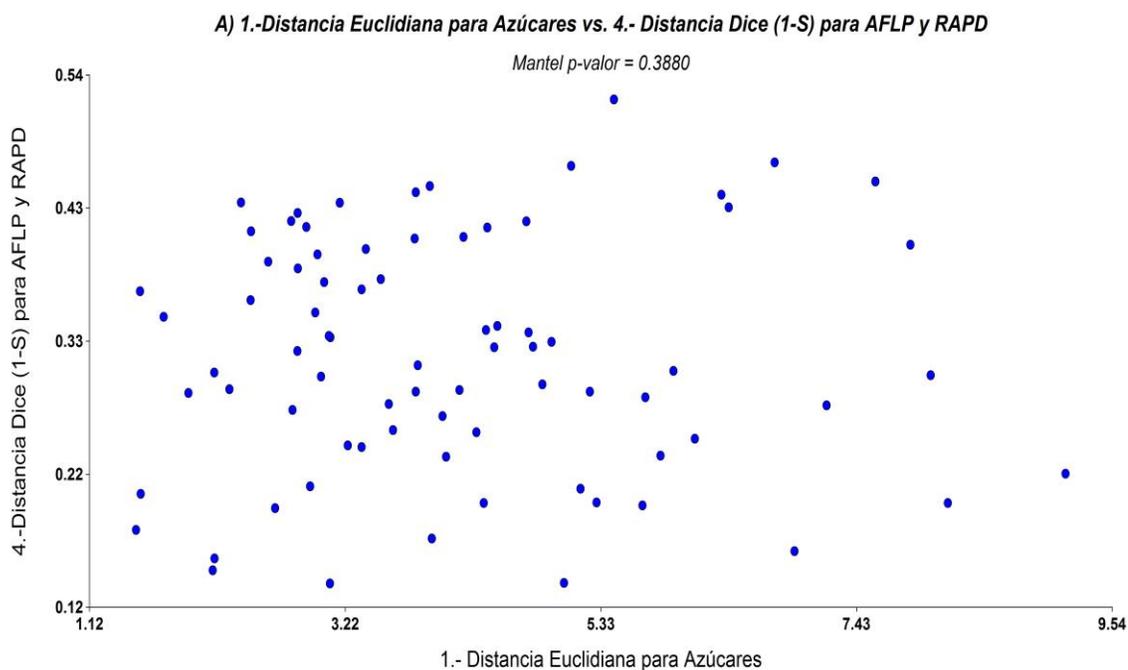


Figura 7. Correlaciones entre matrices 1.-Distancia Euclidiana para Azúcares vs. 4.-Distancia Dice (1-S) para AFLP y RAPD.

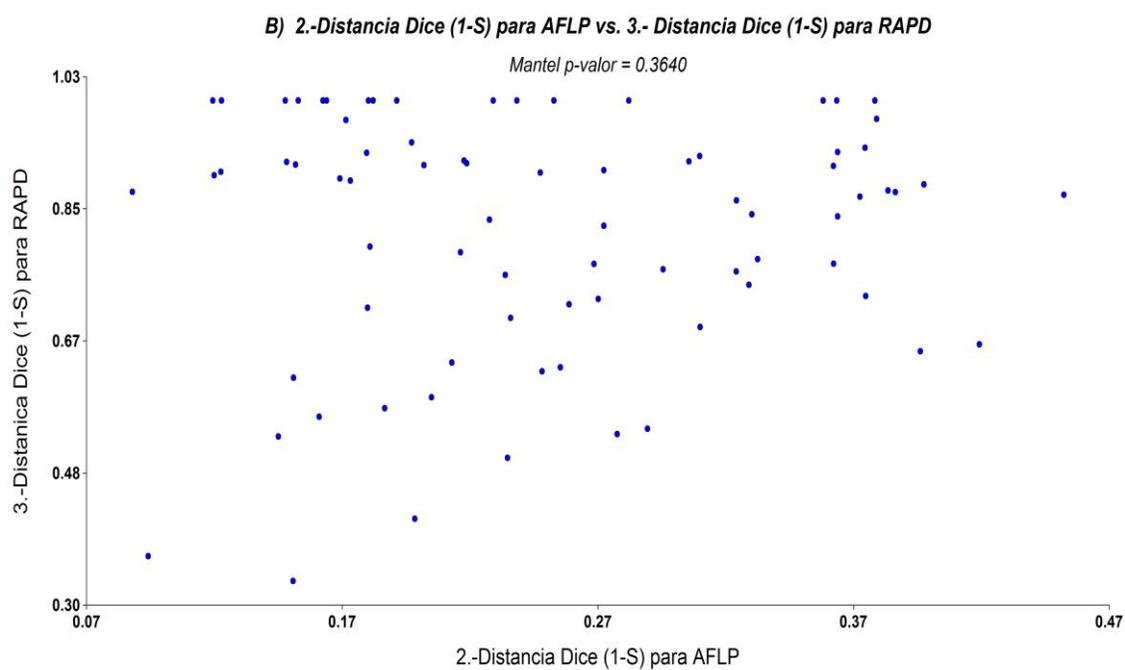


Figura 8. Correlaciones entre matrices 2.-Distancia Dice (1-S) para AFLP vs. 3.-Distancia Dice (1-S) para RAPD.

4.5.1. Análisis de Agrupamiento

El dendograma construido por el método de UPGMA para la matriz de distancia Euclidiana con los datos morfológicos se presenta en la Figura 9; éste agrupó diferencialmente a las líneas R más altas en azúcares que son los genotipos KANSAS COLLIER y FA-UANL39 con una distancia de 1.5 entre los genotipos; además, se agruparon los híbridos más altos en azúcares (FA-UANL 35 x FA-UANL 39 y AN601 x FA-UANL 39) con una distancia entre ellos de 2.27 y una distancia de 2.44 comparados con las líneas R. Los genotipos más bajos en azúcares (FA-UANL 35, FA-UANL 33, FA-UANL 33 x KANSAS COLLIER y AN601) se separaron a una distancia de 5.02.

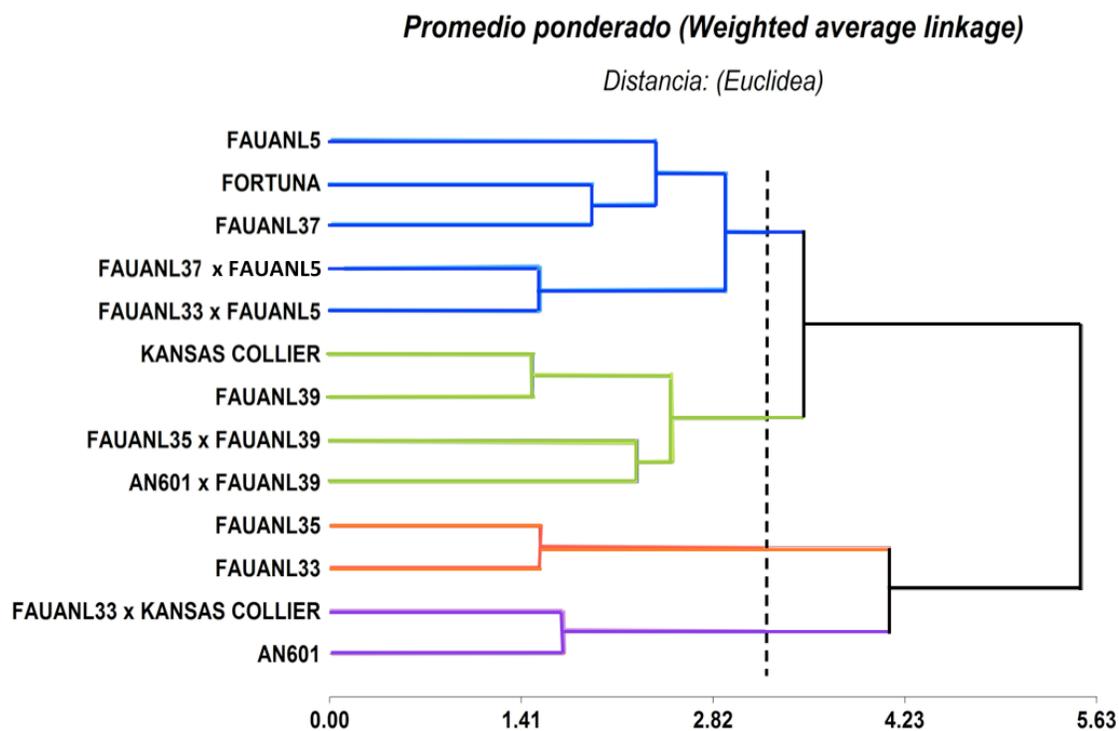


Figura 9. Dendrograma UPGMA con datos morfológicos de azúcares reductores (Glucosa y Fructosa). Correlación cofenética= 0.76.

En cuanto a las distancias Dice ($\sqrt{(1-S)}$) con los datos moleculares obtenidos de los AFLP y RAPD se construyó el dendograma con el método UPGMA indicado en la Figura 10; en éste análisis se incluyó como testigo negativo una variedad comercial de grano KING GOLD. Los híbridos experimentales se separaron ampliamente del resto de los genotipos con una distancia de 0.62; formándose un subgrupo de los híbridos más altos en azúcares (FA-UANL 37 x FA-UANL 5, FA-UANL 35 x FA-UANL 39 y AN601 x FA-UANL 39) a una distancia de 0.54. Las líneas R (FA-UANL 37, FA-UANL 35 y FA-UANL 33) formaron un subgrupo a una distancia de 0.42.

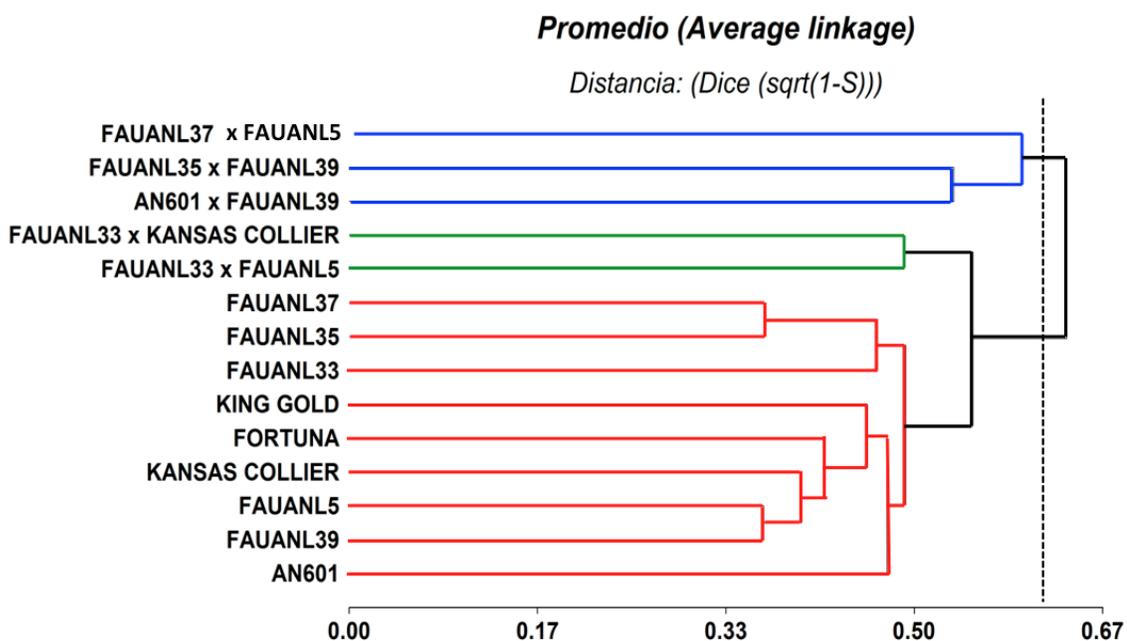


Figura 10. Dendograma UPGMA con datos moleculares AFLP y RAPD Distancia Dice ($\sqrt{(1-S)}$). Correlación cofenética =0.938.

El resultado de la construcción del dendograma mediante el método UPGMA con las matrices de distancias Gower ($\sqrt{(1-S)}$) con los datos morfológicos y datos moleculares se muestra en la Figura 11; se observó un patrón muy similar

Debido a que los distintos algoritmos de conglomeración jerárquica podrían producir agrupamientos diferentes sobre un mismo conjunto de datos, el coeficiente de correlación cofenética mayor se utiliza como criterio de selección del algoritmo. El análisis de agrupamiento también se realizó con el método de WARD; sin embargo, éste obtuvo valores menores a los de UPGMA. En el Cuadro 5 se presentan los datos de las correlaciones cofenéticas entre las matrices de distancia y el dendograma realizado.

Cuadro 5. Correlaciones cofenética del análisis de conglomerados.

Distancia correspondiente	Correlación cofenética UPGMA	Correlación cofenética WARD
Euclidiana	0.760	0.688
Dice ($\sqrt{(S-1)}$)	0.938	0.674
Gower ($\sqrt{(S-1)}$)	0.952	0.698

4.6. Generación de Marcadores SCAR

El protocolo de aislamiento y reamplificación del fragmento permitió conservar la banda originalmente obtenida por RAPD. Los marcadores secuenciados que obtuvieron una mejor calidad se muestran en el Cuadro 6.

Cuadro 6. Identificación de fragmentos RAPD secuenciados.

Identificación	Genotipo*	Iniciador	Secuencia iniciador 5´-3´	Tamaño aproximado del producto de PCR (pb)	Tamaño final de la secuencia (pb)
9	3	OP-A 2	TGCCGAGCTG	550	447
10	4	OP-A 2	TGCCGAGCTG	500	438
11	4	OP-A 2	TGCCGAGCTG	500	438
12	4	OP-A 2	TGCCGAGCTG	500	440
13	4	OP-A 2	TGCCGAGCTG	500	442
14	4	OP-A 2	TGCCGAGCTG	500	439
23	3	OP-C 1	TTCGAGCCAG	1100	1050
25	4	OP-C 1	TTCGAGCCAG	1100	992
26	4	OP-C 1	TTCGAGCCAG	800	995
29	4	OP-C 4	CCGCATCTAC	690	536
43	4	OP-C 14	TGCGTGCTTG	950	812
46	4	OP-C 14	TGCGTGCTTG	700	441
47	4	OP-C 14	TGCGTGCTTG	540	429
48	4	OP-C 19	ACCCCGAAG	1030	776
60	2	OP-F 4	GGTGATCAGG	680	359
63	3	OP-F 4	GGTGATCAGG	220	230
66	2	OP-I 3	CAGAAGCCCA	600	547
67	4	OP-I 3	CAGAAGCCCA	600	545
87	3	OP-K 15	CTCCTGCCAA	900	744
88	4	OP-K 15	CTCCTGCCAA	850	745
92	2	OP-K 20	GTGTCGCGAG	900	797
93	3	OP-K 20	GTGTCGCGAG	600	432
94	3	OP-K 20	GTGTCGCGAG	600	446
95	4	OP-K 20	GTGTCGCGAG	600	440

*Genotipo: 2) AN601 x FA-UANL 39; 3) FA-UANL 35 A x FA-UANL 39; 4) FA-UANL 33 A x FA-UANL 5

En el análisis de alineación se obtuvieron diez agrupamientos entre las secuencias del mismo iniciador RAPD al comparar entre genotipos (Cuadro 7), así como las alineaciones con algunas secuencias aún no identificados por su función genética. Con las secuencias alineadas con secuencias genómicas de sorgo, se procedió al diseño de iniciadores *forward*, los pares de iniciadores diseñados se encuentra en el Cuadro 8.

Cuadro 7. Resultados del BLAST y descripción de las alineaciones con *Sorghum bicolor* sp. y *Zea mays*, secuencias agrupadas por iniciador común. Todas las secuencias de GenBank de *Sorghum bicolor* sp. contenía al menos 100 pb de secuencias cadena arriba 5'.

Grupo	Frag- mentos	Accesión o Secuencia con mayor identidad	Error Esperado	Identida d	Definición	Referencia
1	9, 10, 11, 12, 13, 14	XM_002440182	3e-49	217/298 (75%)	<i>Sorghum bicolor</i> Proteína hipotética, mRNA	Paterson <i>et al.</i> 2009
		NM_001156026	2e-13	149/222 (67%)	<i>Zea mays</i> Purina permeasa, mRNA	Alexandrov <i>et al.</i> 2009
2	29	AC196829.2	5e-55	213/268 (79%)	Clone SB_BBc0050H06	No publicado
		XM_002459992.1	4e-50	156/183 (85%)	<i>Sorghum bicolor</i> Proteína hipotética, mRNA	Paterson <i>et al.</i> 2009
3	23 25 26	XM_002459992.1	3e-5	76/108 (70%)	<i>Sorghum bicolor</i> Proteína hipotética, mRNA	Paterson <i>et al.</i> 2009
4	92 93 94 95	XM_002458588.1	2e-28	116/144 (81%)	<i>Sorghum bicolor</i> Proteína hipotética, mRNA	Paterson <i>et al.</i> 2009

Cuadro 7. Continuación...

5	66	XM_002445441.1	3e-122	335/396 (85%)	<i>Sorghum bicolor</i> Proteína hipotética, mRNA	Paterson <i>et al.</i> 2009
	67	XM_00866475.1	3e-52	288/402 (72%)	<i>Zea mays</i> probable alpha,alpha-trealosa-fosfato sintasa, mRNA	Paterson <i>et al.</i> 2009
6	43	AC169373	4e-28	163/233 (70%)	<i>Sorghum bicolor</i> clone SB_BBc0188M08	Stanford Human Genome Center
7	46 47	XM_002450411	0.0	485/511 (94%)	<i>Sorghum bicolor</i> Proteína hipotética, mRNA	Paterson <i>et al.</i> 2009
8	48	AC196837	0.0	570/684 (83%)	<i>Sorghum bicolor</i> clone SB_BBc0073F19, secuencia completa	Stanford Human Genome Center
9	60 63	XM_002455799.1	2e-70	182/205 (89%)	<i>Sorghum bicolor</i> Proteína hipotética, mRNA	Paterson <i>et al.</i> 2009
10	87 88	Sb01g035510	1e-12	90/119 (76%)	Proteína Nuclear <i>cap-binding</i> CBP80	Paterson <i>et al.</i> 2009

Cuadro 8. Diseño de iniciadores *Forward* y *Reverse* derivados de fragmentos alineados con secuencias de *Sorghum bicolor* sp.

Secuencia (5'-3')	Longitud (pb)	Empieza	Termina	Tm °C	GC%	Auto-complementariedad
Par de iniciadores 1						
Forward GAGAGTCTTCTGTCCCGGTG	20	8	27	59.47	60	2
Reverse TGTCTGATCGGTTGGTCGAT	20	328	309	58.82	50	2
Producto esperado	321					
Par de iniciadores 2						
Forward ACCTCCTTGTCCTCAATGTCTG	20	137	156	59.68	55	3
Reverse TCCCAGCTTGTATTCGTCCC	20	377	358	59.46	55	4
Tamaño del producto esperado	241					
Par de iniciadores 3						
Forward ACCAAACCGCATTGTTCCAC	20	75	94	59.61	50	1
Reverse CAGGCAGATCAGGAGAACCC	20	317	298	59.82	60	0
Tamaño del producto esperado	243					
Par de iniciadores 4						
Forward CGCCCGTATGTGTTTCTCCT	20	195	214	60.11	55	0
Reverse GGAGTTGTAAATGGCGGTGC	20	426	407	59.83	55	2
Tamaño del producto esperado	231					
Par de iniciadores 5						
Forward GTCCCCCGATATCTCTTCGC	20	12	31	59.76	60	2
Reverse TGCGTACACGGTTTCTCACA	20	310	291	59.90	50	2
Tamaño del producto esperado	299					

Cuadro 8. Continuación...

Par de iniciadores 6							
Forward	GTCTGACTGATGGAGGCGAG	20	82	101	59.9	60	0
Reverse	CCTTGTCGTAGTGC GGTTCT	20	261	242	60.00	55	1
Tamaño del producto esperado		180					
Par de iniciadores 7							
Forward	GAGTGCTCCTACTCTTGCATCC	22	52	73	60.48	55	2
Reverse	AGCTAATGTACATCAGCGGGT	21	390	370	60.00	48	2
Tamaño del producto esperado		339					
Par de iniciadores 8							
Forward	ATCCAATCTAACTAAGCCAA	20	8	27	51.48	35	0
Reverse	CGAGTTTTGGTGT TTGTCTAT	21	698	678	54.44	38	2
Tamaño del producto esperado		691					
Par de iniciadores 9							
Forward	CGGTTGCAGGAGAGCTAACTT	21	20	40	60.34	52.38	0
Reverse	CCTCGCCTATCATCGTCTCC	20	127	108	59.48	60	0
Tamaño del producto esperado		108					
Par de iniciadores 10							
Forward	TCCTTTGTCCCTACAGTTTTTCA	23	5	27	57.88	39	1
Reverse	TCGAGTCGCCTTATCCTACCT	21	458	438	59.86	52	0
Tamaño del producto esperado		454					

Se realizaron seis repeticiones de la secuencia por cada agrupamiento, debido a la poca definición de los cromatogramas que no permitían definir su lectura como se muestra en la Figura 12.

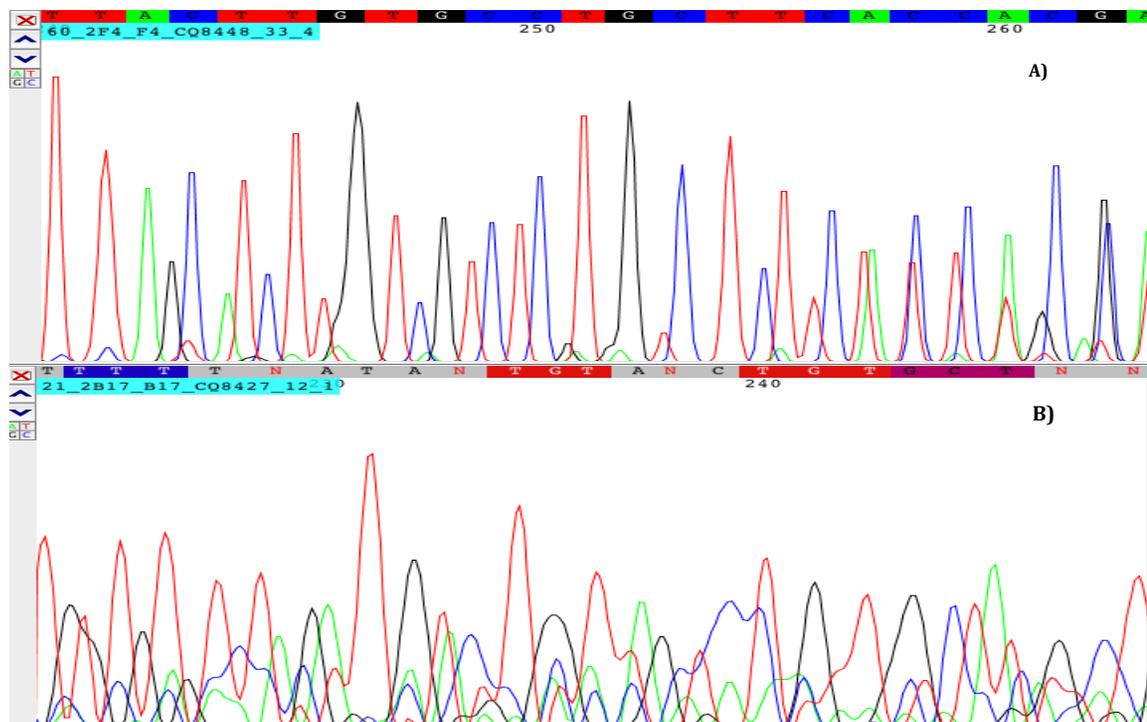


Figura 12. Ejemplo de cromatogramas obtenidos de la secuenciación de fragmentos. A) Fragmento 60, ejemplo de un cromatograma correcto: picos uniformemente espaciados, cada uno con un solo color, equidistantes. B) Fragmento 21, ejemplo de una secuenciación con poca definición.

Los iniciadores de los agrupamientos 2, 5, 6 y 9 no amplificaron ADN en la banda original secuenciada por lo que se infiere que la primera secuenciación realizada no era correcta o estuvo turbada por diferentes razones; a pesar de que el proceso de secuenciación puede estar bajo control de la computadora, hay límites de la exactitud de los datos, la repetitividad, precisión y fiabilidad. Hay una variedad de posibles fuentes de error que pueden afectar la calidad de

los datos, como puede ser el fracaso del detector para registrar colorantes fluorescentes correctamente a las inconsistencias en la coincidencia de patrones (Bergeron, 2002; Metzker, 2010).

En cuanto a los agrupamientos 1, 3, 4, 7, 8 y 10, los iniciadores *forward* y *reverse* amplificaron la banda de la longitud correspondiente al peso que se tenía estimado, como se muestra en la Figura 13.

Al lograrse la amplificación de estas bandas, se realizó la PCR con los pares de iniciadores 1, 3, 4, 7, 8 y 10 utilizando el ADN genómico total de los genotipos en el estudio, los resultados se localizan en la Figura 14. Se obtuvo la amplificación de una banda única (monomórfica) correspondiente al peso estimado para el caso de los pares 1, 3, 8 y 10; estos iniciadores se descartan para su posible uso en la identificación de individuos con alto contenido de azúcar. Al presentarse esta banda en todos los genotipos, se infiere que dichos marcadores corresponden a secuencias propias de la especie y no arrojan ningún tipo de información que nos permite diferenciar entre genotipos altos o bajos en azúcares.

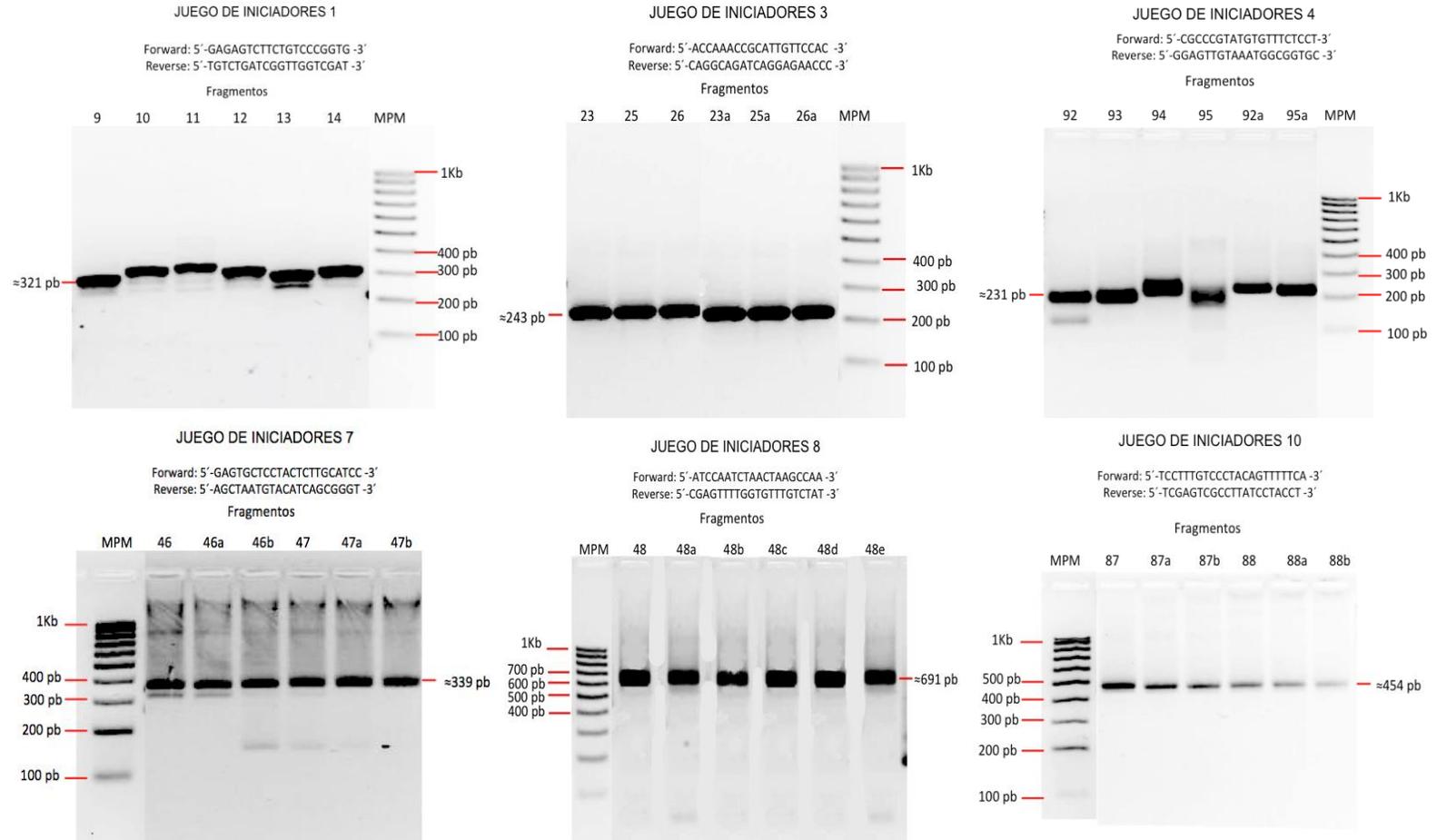


Figura 13. Electroforesis en gel de agarosa al 3%. El carril MPM corresponde al marcador de peso molecular.

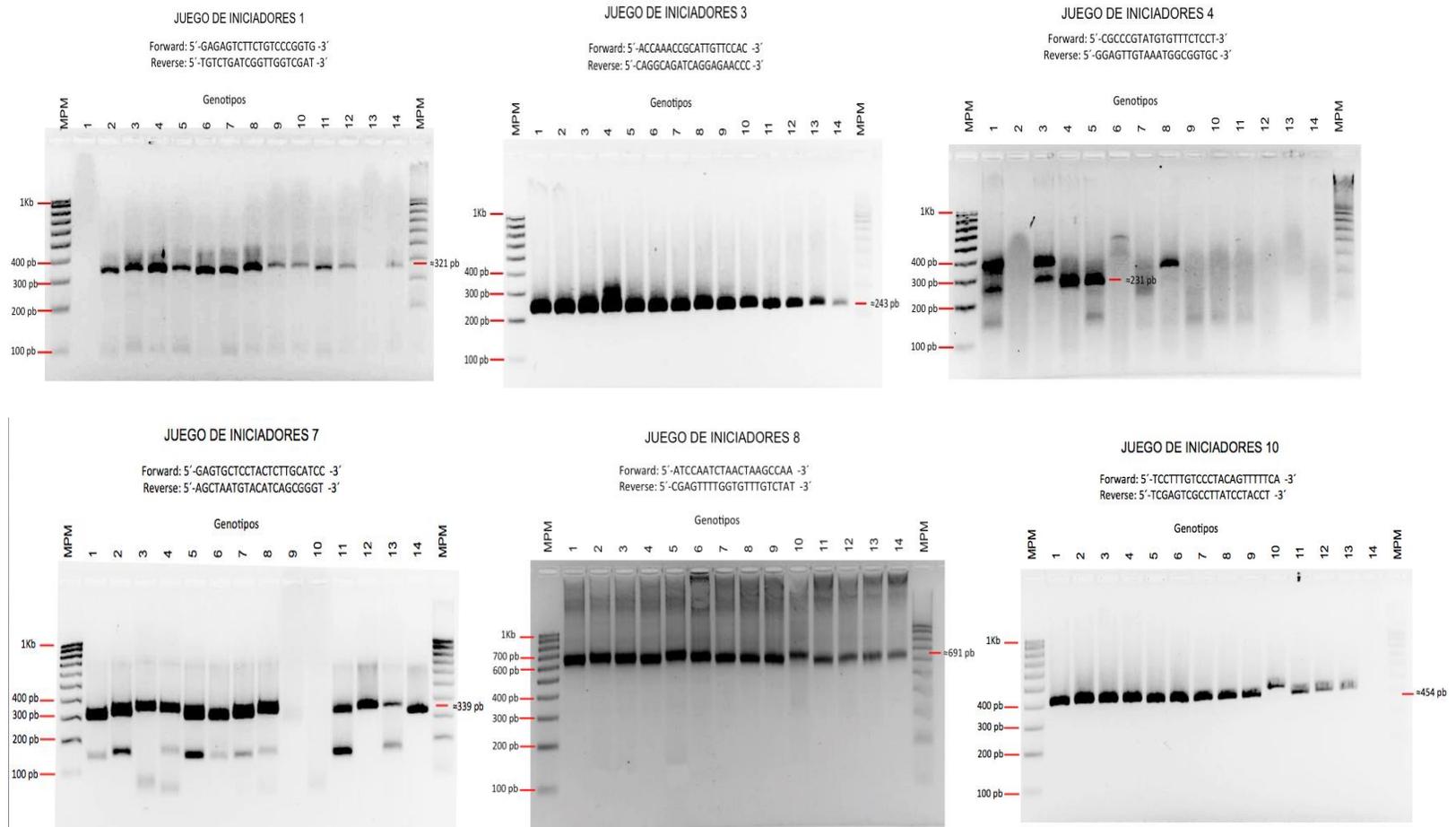


Figura 14. Electroforesis en gel de agarosa al 3%. El carril MPM corresponde al marcador de peso molecular; los genotipos son: 1) FA-UANL 33A x KANSAS COLLIER; 2) A601 x FA-UANL 39; 3) FA-UANL 35A x FA-UANL 39; 4) FA-UANL 33A x FA-UANL 5; 5) FA-UANL 37A x FA-UANL 5; 6) FA-UANL 37A; 7) FA-UANL 33A; 8) FA-UANL 35A; 9) AN601; 10) FA-UANL 39; 11) FA-UANL 5; 12) KANSAS COLLIER; 13) FORTUNA y 14) KING GOLD.

Sin embargo para los pares de iniciadores 4 y 7 se obtuvieron bandas con diferentes peso molecular, la electroforesis en agarosa al 3% no permitió la separación adecuada de las bandas, por lo que se prosiguió a realizar una electroforesis en acrilamida:bisacrilamida (40:1) al 8% para obtener una separación de las bandas con mayor definición.

En la electroforesis para los productos de PCR con el juego de iniciadores 4, *Forward:* 5'-CGCCCGTATGTGTTTCTCCT-3' y *Reverse:* 5'-GGAGTTGTAAATGGCGGTGC-3' (Figura 15) se discernió una banda de 231 pb presente en los genotipos altos en azúcares que son los híbridos F1: A601 x FA-UANL 39, FA-UANL 35A x FA-UANL 39, FA-UANL 33A x FA-UANL 5 y FA-UANL 37A x FA-UANL 5, dicha banda no amplificó en ningún otro genotipo, pero tampoco se encontró en ningún progenitor, por lo que resulta complicado confirmar su identidad como marcador SCAR. Es necesario realizar un análisis de segregación en individuos de líneas avanzadas con estos iniciadores. La información de la secuencia correspondiente se localiza en la Figura 16.

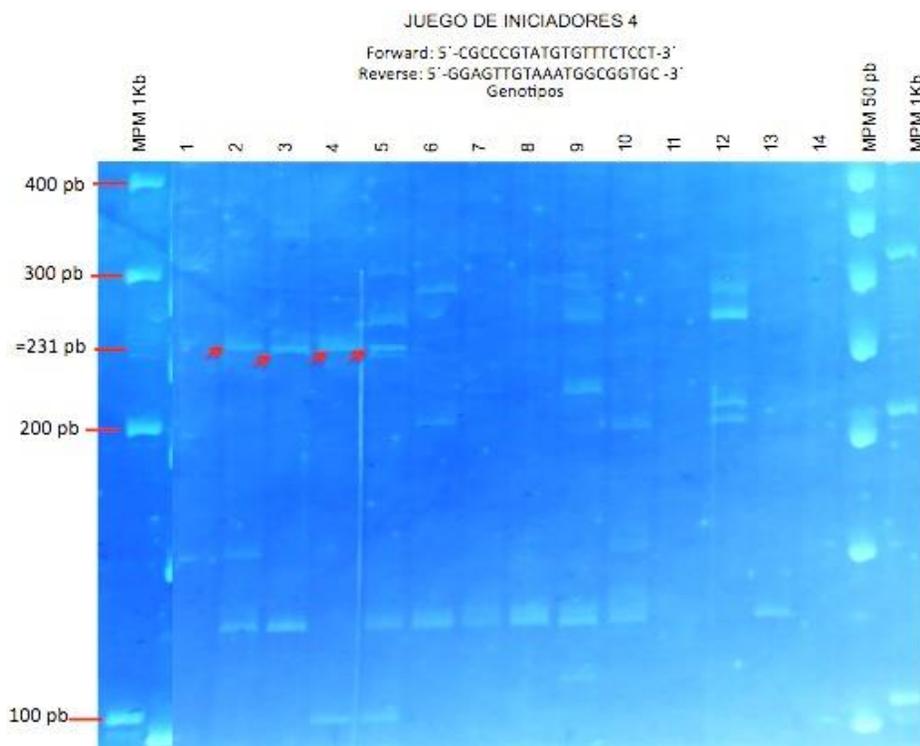


Figura 15. Electroforesis en gel de acrilamida: bis-acrilamida (40:1) al 8%. El carril MPM corresponde al marcador de peso molecular; los genotipos son: 1) FA-UANL 33A x KANSAS COLLIER; 2) A601 x FA-UANL 39; 3) FA-UANL 35A x FA-UANL 39; 4) FA-UANL 33A x FA-UANL 5; 5) FA-UANL 37A x FA-UANL 5; 6) FA-UANL 37A; 7) FA-UANL 33A; 8) FA-UANL 35A; 9) AN601; 10) FA-UANL 39; 11) FA-UANL 5; 12) KANSAS COLLIER; 13) FORTUNA y 14) KING GOLD.

```

1  GTGTCGCGAG CAAGAGGGGA AGCAGTGACA GAGCGAAGTA AATGTAAATG GCCTGGTCCG 60
   OP-K20
61  GGGCGGCTGC GGGCGGAGGA GGGGATGACG CGCGGTACAA AGAGTCCGCG GGGGTGAGGG 120
121 CGTCCACGA GGCCCTCCGC AAAGTAGGCC TCTGCTCCGA CTGGACGTAC TCAGTATTCT 180
180 GCAGCATCCG TCCTCGCCCG TATGTGTTTC TCCTCCCTGT CGTCCTTTCC CCGATCCCTC 240
   FORWARD 4
241 TAGCTACTAG AGGAAAACT AGAGACCTCT CTTGCTTTGC TTCAGTCCG TCCGGACAGC 300
301 CGACACCAAC AACGAACTCG TCTTCTAGT AACTTAGCTA GCTAGCTACT GTCTTCATAT 360
361 TCGATCATGT GTGTGCATGC CCCTCCTGCG CGCGGCGGCA ACGGCTGCAC CGCCATTTAC 420
   REVERSE 4
421 AACTCCCTCC ACTCTGTGAC TGCTTTCCTT TTGTTTCCTC GACACACAT 469

```

Figura 16. Secuencia de 469 pb de la región polimórfica de *Sorghum bicolor* sp., mostrando los sitios de unión del iniciador RAPD (OP-K 20) y los marcadores SCAR (*Forward 4* y *Reverse 4*).

En lo que concierne a el juego de iniciadores 7, *Forward*: 5'-GAGTGCTCCTACTCTTGCATCC-3' y *Reverse*: 5'-AGCTAATGTACATCAGCGGGT-3', la electroforesis para los productos de PCR se hallan en la Figura 17, dónde se distinguió una banda de 341 pb presente en los genotipos altos en azúcares que son los híbridos F1: A601 x FA-UANL 39 y FA-UANL 33A x FA-UANL 5, dicha banda no amplificó en ningún otro genotipo, pero sí se encontró en los progenitores B de éstas mismas líneas que son: FA-UANL 39 y FA-UANL 5, además de estar presente en el progenitor KANSAS COLLIER. Sin embargo, una vez más es necesario realizar un análisis de segregación en individuos de líneas avanzadas con estos iniciadores. La información de la secuencia correspondiente se ubica en la Figura 18.

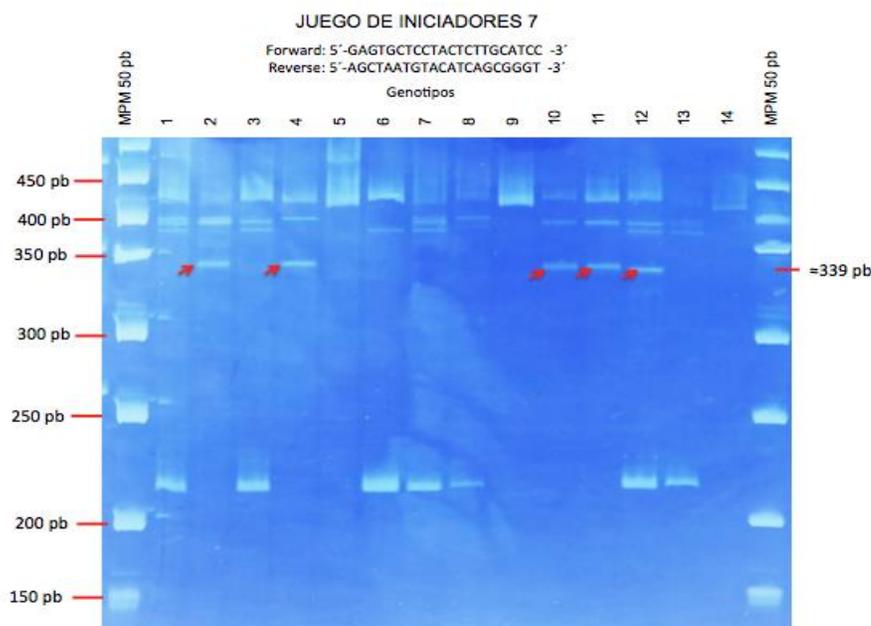


Figura 17. Electroforesis en gel de acrilamida: bis-acrilamida (40:1) al 8%. El carril MPM corresponde al marcador de peso molecular; los genotipos son: 1) FA-UANL 33A x KANSAS COLLIER; 2) A601 x FA-UANL 39; 3) FA-UANL 35A x FA-UANL 39; 4) FA-UANL 33A x FA-UANL 5; 5) FA-UANL 37A x FA-UANL 5; 6) FA-UANL 37A; 7) FA-UANL 33A; 8) FA-UANL 35A; 9) AN601; 10) FA-UANL 39; 11) FA-UANL 5; 12) KANSAS COLLIER; 13) FORTUNA y 14) KING GOLD.

```

1  GTGCGTGCTT GTGACTCCAA AGTCCGGCGA CAACCTCCAC AACCTGACAG TGAGTGCTCC 60
   OP-C14

61  TACTCTTGCA TCCTCAAAC T CAGTTCATCT GGT CAGTCTA CATCTAGATG TGAAAGTTTC 120
   FORWARD 7

121 TTTTATAAAA ATTTTAGCCC AGATCAACT ATTTATGTGT ATATTTGCAT AAAAAGAGAG 180

181 AATACTAGTT CATGTTCTGA TTACAGCTAG CACACAATTT GCCTTGCCA GAACATATCA 240

241 CCTAGTATTT GTTAGCTATG CTCTACTAGT ACAGTCAAAA TATCTATCTT CATGGTAGGT 300

301 GTATTGTGCT CTTAGTGGAG GAGGAAAAAG TTTCTATAGA TGTGCACGTG AAAATAAACC 360

361 TCACCTTAAA CCCGCTGATG TACATTAGCT TCACTCTCCT ACATTACCTA CAAC      414
   REVERSE 7

```

Figura 18. Secuencia de 414 pb de la región polimórfica de *Sorghum bicolor* sp., mostrando los sitios de unión del iniciador RAPD (OP-C 14) y los marcadores SCAR (*Forward 7* y *Reverse 7*).

5. CONCLUSIONES

En este trabajo de investigación se obtuvieron las siguientes conclusiones:

- Para el objetivo 1: se estableció la diversidad genética del subconjunto de accesiones de la colección de sorgo dulce de la FA-UANL mediante la identificación de 549 marcadores AFLP y 256 marcadores RAPD con un índice de contenido polimórfico de 0.26 y 0.21, respectivamente. Ambos marcadores permitieron la estimación de la variación genética en sorgo al presentar un coeficiente de distancia genética de 0.42 (Dice ($\sqrt{S-1}$)) en el agrupamiento. La técnica de AFLP resulta ser ventajosa para evaluar la variación entre genotipos, al tener un mayor rendimiento con 18.3 bandas por juego de iniciadores; sin embargo, por su tamaño no son adecuados para el diseño de iniciadores SCAR.

- Para el objetivo 2: la técnica de marcadores RAPD permitió la identificación de 53 marcadores polimórficos presentes en los genotipos altos en azúcares que asimismo estaban ausentes en los genotipos más bajos en azúcares con rango de tamaño de 220 a 2800 pb.

- Para el objetivo 3: se diseñaron dos juegos de iniciadores SCAR provenientes de fragmentos que alinearon con secuencias genómicas de sorgo reportadas en el NCBI. El juego de iniciadores 4 amplificó una banda de 231 pb presente en los genotipos altos en azúcares; y el juego de iniciadores 7 amplificó una banda de 341 pb presente en los genotipos altos en azúcares, dicha banda se encontró en los progenitores B de éstas mismas líneas y en el progenitor KANSAS COLLIER.

Conforme a lo anterior se recomienda realizar un análisis de segregación en individuos de líneas avanzadas con estos iniciadores para comprobar la utilidad de los SCAR obtenidos.

6. LITERATURA CITADA

- Alexandrov, N.N., V.V. Brover, S. Freidin, M.E. Troukhan, T.V. Tatarinova, H. Zhang, T.J. Swaller, Y.P. Lu, J. Bouck y R.B. Flavell. 2009. Insights into corn genes derived from large-scale cDNA sequencing. *Plant Mol Biol* 10:179-194.
- Almodares, A. y M.R. Hadi. 2009. Production of bioethanol from sweet sorghum : A review. *African Journal of Agricultural Research* 4(9):772-780.
- Alvarado, J.I, B. Hernández, A. Morales, E. Ávila, S. Payán y M. Camarillo. 2011. Genotipos de sorgo dulce potenciales para producción de bioetanol en el Valle de Mexicali. *Investigación y Ciencia* 52:15-20.
- Ayana, A., T. Bryngelsson, y E. Bekele. 2000. Genetic variation of Ethiopian and Eritrean sorghum (*Sorghum bicolor* (L.) Moench) germplasm assessed by random amplified polymorphic DNA (RAPD). *Genetic Resources and Crop Evolution* 47:471-482.
- Balzarini, M., A. Arroyo, C. Bruno, y J. Di Renzo. 2006. Análisis de Datos de Marcadores con InfoGen. En: XXXV Congreso Argentino de Genética. San Luis, Argentina. 96 p.
- Balzarini, M., L. Gonzalez, E. Tablada, F. Casanoves, J. Di Rienzo y C. Robledo. 2008. Manual de usuario. Editorial Brujas. Córdoba, Argentina. 336 p.
- Bergeron, B. 2002. *Bioinformatics Computing*. Prentice Hall PTR. New York, USA. 438 p.
- Botstein, D., R. White, M. Skolnick, M. y R. Davis. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet* 32:314-331.
- Boora, K.S., R. Frederiksen y C. Magill. 1999. A molecular marker that segregates with sorghum leaf blight resistance in one cross is maternally inherited in another. *Molecular & General Genetics* 261(2):317-22.

- Calviño, M., y J. Messing. 2012. Sweet sorghum as a model system for bioenergy crops. *Current Opinion in Biotechnology* 23(3):323-9.
- Calviño, M., M. Miclaus, R. Bruggmann y J. Messing. 2009. Molecular markers for sweet sorghum based on microarray expression data. *Rice* 2(2-3):129-142.
- Chang, J., J. Cui, W. Xue y Q. Zhang. 2012. Identification of molecular markers for a aphid resistance gene in sorghum and selective efficiency using these markers. *Journal of Integrative Agriculture* 11(7):1086-1092.
- Chen, Q., S. Wei, Z. Deng, L. Yin, B He, y X. Kong. 2009. Optimization of DNA extraction from seeds of *Sorghum sudanense* (Piper) Stapf. *Notulae Botanicae Horti Agrobotanici Cluj-Napoca* 37(1):256-260.
- Choi, Y.E., C.H. Ahn, B.B Kim y E. Yoon. 2008. Development of species specific AFLP-derived SCAR marker for authentication of *Panax japonicus* C.A. MEYER. *Biol Pharm Bull* 31:135-138
- Clayton, W.D. 1961. Proposal to conserve the generic name *Sorghum* Moench. (Gramineae) versus *Sorghum* Adans (Gramineae). *Taxonomy* 10:242-243.
- Cuadras, C.M. 1981. Métodos de análisis multivariante. Ed. EUNIBAR. Barcelona, España. 642 p.
- Dahlberg, J.A., X. Zhang, G.E. Hart, y J.E. Mullet. 2002. Comparative assessment of variation among sorghum germplasm accessions using seed morphology and RAPD measurements. *Crop Science* 42(1):291.
- Devaiah, K.M y P. Venkatasubramanian. 2008. Development of SCAR marker for authentication of *Pueraria tuberosa* (Roxb. ex. Willd.) DC. *Current Science* 94: 1306-1309.
- Dhanya, K., S. Syamkumar y B. Sasikumar. 2009. Development and application of SCAR marker for the detection of papaya seed adulteration in traded black pepper powder. *Food Biotechnol* 23:97-106.
- Dice, L.R. 1945. Measures of amount of ecologic association between species. *Ecology* 26(3):297-302.
- Elangovan, M., P. Kiran, N. Seetharama y J.V. Patil. 2013. Genetic diversity and heritability characters associated in Sweet Sorghum [*Sorghum bicolor* (L.) Moench]. *Sugar Tech* 16(2):200-210.

- Elbehri, A., A. Segerstedt y P. Liu. 2013. Biofuels and the sustainability challenge: A global assessment of sustainability issues, trends and policies for biofuels and related feedstocks. FAO, Rome, Italy. 73 p.
- Franco, T.L. y R. Hidalgo. 2003. Análisis estadístico de datos de caracterización morfológica de recursos fenotípicos. Boletín Técnico No. 8. Instituto Internacional de Recursos Fitogenéticos (IPGR) Cali, Colombia. 89 p.
- Geleta, N., M.T. Labuschagne y C.D. Viljoen. 2006. Genetic diversity analysis in sorghum germplasm as estimated by AFLP, SSR and morpho-agronomical markers. *Biodiversity and Conservation* 15(10): 3251-3265.
- Gower, J. 1971. A general coefficient of similarity and some of its properties. *Biometrics* 27(4):857-871.
- Horejsi, T., J.M. Box y J.E. Staub. 1999. Efficiency of randomly amplified polymorphic DNA to sequence characterized amplified region marker conversion and their comparative polymerase chain reaction sensitivity in cucumber. *J Amer Soc Hort Sci* 124(2):128-135.
- Iqbal, A., B. Sadia, A.I. Khan, F.S. Awan, R.A. Kainth y H.A. Sadaqat. 2010. Biodiversity in the sorghum (*Sorghum bicolor* (L.) Moench) germplasm of Pakistan. *Genetics and Molecular Research* 9:756-764.
- Jaiswal, P., A. Sane, S. Ranade, P. Nath y P.V. Sane. 1998. Mitochondrial and total DNA RAPD patterns can distinguish restorers of CMS lines in sorghum. *Theoretical and Applied Genetics* 96(6-7):791-796.
- Jiang, G. 2010. Molecular markers and marker-assisted breeding in plants. In: S. Anderson (ed.). *Plant Breeding from Laboratories to Fields*. Croatia: InTech. pp. 45-83.
- Kalendar, R., D. Lee y A.H. Schulman. 2011. Java web tools for PCR, in silico PCR, and oligonucleotide assembly and analysis. *Genomics* 98(2):137-44.
- Kasai, K., Y. Morikawa, V.A. Sorri, J.P. Valkonen, C. Gebhardt y K.N. Watanabe. 2000. Development of SCAR markers to the PVY resistance gene Ryadg based on a common feature of plant disease resistance genes. *Genome* 43:1-8.
- Kesawat, M.S. y B.K. Das. 2009. Molecular markers: It's application in crop improvement. *Journal Crop Science Biotechnology* 12(4):169-181.
- Kiran, U., S. Khan, K.J. Mirza, M. Ram y M.Z. Abdin. 2010. SCAR markers: a potential tool for authentication of herbal drugs. *Fitoterapia* 81(8):969-76.

- Kosaric, N., F. Vardar-Sukan y T. Senn. 2001. The Biotechnology of ethanol classical and future applications. M. Roehr, Ed. WILEY-VCH. Weinheim, Germany. 232 p.
- Kumar, P., V.K. Gupta, A.K. Misra, D.R. Modi, y B.K. Pandey. 2009. Potential of molecular markers in plant biotechnology. *Plant Omics Journal* 2(4):141-162.
- Ladhalakshmi, D., A. Vijayasamundeeswari, V. Paranidharan y R. Samiyappan. 2009. Molecular identification of isolates of *Peronosclerospora sorghi* from maize using PCR-based SCAR marker. *World J Microb Biot* 25:2129-2135.
- Lazarín, K.J., F. Zavala, J. Aranda, E. García, K. Arévalo y M. Aguilar. 2013. Dinámica de acumulación de azúcares durante el desarrollo de sorgo dulce. *En: XXXVIII Ciclo de Seminarios verano 2013. Área de Estudios de Posgrado. Campus de Ciencias Agropecuarias. General Escobedo, N.L., México. 6 y 7 de junio 2013. pp 143-148.*
- Lekgari, A., y I. Dweikat. 2014. Assessment of genetic variability of 142 sweet sorghum germplasm of diverse origin with molecular and morphological markers. *Open Journal of Ecology* 4:371-393.
- Leps, J. y P. Smilauer. 2003. Multivariate analysis of ecological data using CANOCO. Ed University Press. Crambridge. United Kingdom. 269 p.
- Levitus, G., V. Echenique, C. Rubinstein, E. Hopp y L. Mroginski. 2010. Biotecnología y mejoramiento vegetal II. INTA. Argentina. 648 p.
- LI-COR Inc. 2014. CodonCode Aligner. V 5.0.1.
- Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Research* 27:209-220.
- Menkir, A., P. Goldsbrough y G. Ejeta. 1997. RAPD based assessment of genetic diversity in cultivated races of sorghum. *Crop Science* 37(2):564.
- Menz, M.A., R.R. Klein, N.C. Unruh, W.L. Rooney, P.E. Klein y J.E. Mullet. 2004. Genetic diversity of public inbreds of sorghum determined by mapped AFLP and SSR markers. *Crop Science* 44(4):1236-1244.
- Metzker, M.L. 2010. Sequencing technologies-the next generation. *Nature Reviews. Genetics* 11(1):31-46.
- Michelmore, R.W., I. Paran y R.K. Kesseli. 1991. Identification of markers linked to disease-resistance genes by bulked segregant analysis: A rapid method

to detect markers in specific genomic regions by using segregating populations. *Proceedings of National Academy of Sciences of the United States of America* 88:9828-9832.

- Möller, E.M., G. Bahnweg, H. Sandermann y H.H. Geiger. 1992. A simple and efficient protocol for isolation of high molecular weight DNA from filamentous fungi, fruit bodies, and infected plant tissues. *Nucleic Acids Research* 20(22):6115-6116.
- Moore, D., D. Dowhan, J. Chory y R.K. Ribaud. 2002. Isolation and purification of large DNA restriction fragments from agarose gels. *In*: F. Ausubel, R. Bent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith y K. Struhl (ed.). *Current protocols in molecular biology*. Wiley. New York, USA. pp. 2.6.-2.6.11.
- Morris, G.P., P. Ramu, S.P. Deshpande, C.T. Hash, T. Shah, H.D. Upadhyaya, O. Riera-Lizarazu, P.J. Brown, C.B. Acharya, S.E. Mitchel, J. Harriman, J.C. Glaubitz, E.S. Blucker y S. Kresovich. 2013. Population genomic and genome-wide association studies of agroclimatic traits in sorghum. *Proceedings of the National Academy of Sciences of the United States of America* 110(2):453-458.
- Murray, S.C., W.L. Rooney, M.T. Hamblin, S.E. Mitchell, y S. Kresovich. 2009. Sweet sorghum genetic diversity and association mapping for brix and height. *The Plant Genome Journal* 2(1):48-62.
- Norusis, M.J. 2008. SPSS® Statics 17.0. Procedures companion. Ed. Prentice Hall. Pearson Education. USA. 648 p.
- Pammi, S., K. Schertz, G. Xu, G. Hart y J.E. Mullet. 1994. Random-amplified-polymorphic DNA markers in sorghum. *Theoretical and Applied Genetics* 89:80-88.
- Paran, I. y R.W. Michelmore. 1993. Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce. *Theoretical and Applied Genetics* 85(8):985-993.
- Paterson, A.H., J.E. Bowers, R. Bruggmann, I. Dubchak, J. Grimwood, H. Gundlach, G. Haberer, U. Hellsten, T. Mitros, A. Poliakov, J. Schmutz, M. Spannagl, H. Tang, X. Wang, T. Wicker, A.K. Barthi, J. Chapman, F.A. Feltus, U. Gowik, I.V. Grigoriev, E. Lyons, C. Maher, M. Martis, A. Narechania, P. Otiillar, B. Penning, A. Salomov, Y. Wang, L. Zhang, N. Carpita, M. Freeling, A. Hash, B. Keller, P. Klein, S. Kresovich, M. McCann, R. Ming, D. Paterson, R. Mehboob-ur, D. Ware, K. Mayer, J.

- Messing y D.S. Rokhsar. 2009. The *Sorghum bicolor* genome and the diversification of grasses. *Nature* 457(7229):551-556.
- Pecina, V., J.L. Anaya, A. Zamarripa, N. Montes, C. Nuñez, J.L. Solis, M.R. Aguilar y L. Prom. 2012. Genetic diversity of sweet sorghum germplasm in Mexico using AFLP and SSR markers. *Pesquisa Agropecuaria Brasil* 47(8):1095-1102.
- Prakash, S.P., K.R. Biji, S.M. Gomez, K.G. Murthy y R.C. Babu. 2006. Genetic diversity analysis of sorghum (*Sorghum bicolor* L. Moench) accessions using RAPD markers. *Indian Journal Crop Science* 1(1-2):109-112.
- Qazi, H.A., S. Paraniye y S. Bhargava. 2012. Stem sugar accumulation in sweet sorghum-activity and expression of sucrose metabolizing enzymes and sucrose transporters. *Journal of Plant Physiology* 169:605-613.
- Raghuwanshi, A. y R.G. Birch. 2010. Genetic transformation of sweet sorghum. *Plant Cell Reports* 29(9):997-1005.
- Ray, T. y S.C. Roy. 2009. Genetic diversity of *Amaranthus* species from the indo-gangetic plains revealed by RAPD analysis leading to the development of ecotype-specific SCAR marker. *Journal Heredity* 100:338-347.
- Rincón, F., H. Johnson, J. Crossa, y S. Taba. 1996. Cluster analysis, and approach to sampling variability in maize accessions. *Maydica*. 41:307-316.
- Ritter, K.B., C.L. McIntyre, I.D. Godwin, D.R. Jordan y S.C. Chapman. 2007. An assessment of the genetic relationship between sweet and grain sorghums, within *Sorghum bicolor* ssp. *bicolor* (L.) Moench, using AFLP markers. *Euphytica* 157(1-2):161-176.
- Sambrook, J., E.F. Fritsch, T. Maniatis. 1989. *Molecular cloning: a laboratory manual*. 2ª Ed., 3 vol. Cold Spring Harbor Laboratory Press. New York, USA. 253 p.
- Schut, J.W., X. Qi y P. Stam. 1997. Association between relationship measures based on AFLP markers, pedigree data and morphological traits in barley. *Theoretical and Applied Genetics* 95(7):1161-1168.
- Shegro, A., M.T. Labuschagne, A. van Biljon y N.G. Shargie. 2013. Assessment of genetic diversity in sorghum accessions using amplified fragment length polymorphism (AFLP) analysis. *African Journal of Biotechnology* 12(11):1178-1188.

- Shen, F., L. Kumar, J. Hu y J.N. Saddler. 2011. Evaluation of hemicellulose removal by xylanase and delignification on SHF and SSF for bioethanol production with steam-pretreated substrates. *Bioresource Technology* 102(19):8945-8951.
- Shoemaker, C.E. y D.I. Bransby. 2010. Chapter 9: The role of sorghum as a bioenergy feedstock. *In: Sustainable alternative fuel feedstock opportunities, challenges and roadmaps for six U.S. regions*. Ankeny, Iowa, USA: Soil and Water Conservation Society. pp. 149-159.
- Simpson, W. y K. Gil. 2004. Análisis de AFLP de ADN genómico. I Congreso Internacional de Agrobiotecnología. Facultad de Agronomía, Universidad Autónoma de Nuevo León. 29 de noviembre-1 de diciembre. Marín, N.L., México. 18 p.
- Singh, H.P. y H.C. Lohithaswa. 2006. Sorghum. *In: C. Kole (ed.). Genome mapping and molecular breeding in plants: Volume 1*. Pennsylvania, USA. pp. 257-302.
- Singh, M., K. Chaudhary y K.S. Boora. 2006. RAPD-based SCAR marker SCA 12 linked to recessive gene conferring resistance to anthracnose in sorghum [*Sorghum bicolor* (L.) Moench]. *Theoretical and Applied Genetics* 114:187-192.
- Skroch, P. y J. Nienhuis. 1995. Impact of scoring error and reproducibility RAPD data on RAPD based estimates of genetic distance. *Theoretical and Applied Genetics*. 91(6-7):1086-1091.
- Sperling, D. 1989. Transportation energy: when muddling through is not enough. *Transportation Research A* 23(3):179-181.
- Srinivasa, R.P., S. Rao, N. Seetharama, A. Umakath, R.P. Sanjana, B. Reddy y C. Gowda. 2009. Sweet sorghum for biofuel and strategies for its improvement. *Information Bulletin No. 77*. International Crops Research Institute for the Semi-Arid Tropics. Andhra Pradesh, India. 80 p.
- Srivastava, M.K., C. Li y Y. Li. 2012. Development of sequence characterized amplified region (SCAR) marker for identifying drought tolerant sugarcane genotypes. *Australian Journal of Crop Science* 6(4):763-767.
- Vinutha, K.S., L. Rayaprolu, K. Yadagiri, A. Umakanth, J.V. Patil y R.P. Srinivasa. 2014. Sweet sorghum research and development in India: Status and prospects. *Sugar Tech* 16(2):133-143.

- Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. Van De Lee, A. Frijters, J. Pot, J. Poleman, M. Kuiper y M. Zabeau. 1995. AFLP : a new technique for DNA fingerprinting. *Nucleic Acids Research* 23(21):4407-4414.
- Wang, M.L., C. Zhu, N. Barkley, Z. Chen, J.E. Erpelding, S.C. Murray, M.R. Tuinstra, T. Tesso, G.A. Paderson y J. Yu. 2009. Genetic diversity and population structure analysis of accessions in the US historic sweet sorghum collection. *Theoretical and Applied Genetics* 120(1):13-23.
- Weir, B.J., R.G. Pierre y R.N. Chibbar. 1996. Isolation of DNA for RAPD analysis from leaves of the saskatoon (*Amelanchier alnifolia* Nutt.) and other horticultural crops. *Canadian Journal of Plant Science* 76:819-824.
- Weising K., H. Nybom, K. Wolff y G. Kahl. 2005. DNA fingerprinting in plants: Principles, methods and applications (2nd Edition). CRC Press, New York, USA. 444 p.
- Wetton, J.H., R.E. Carter, D.T. Parkin y D. Walters. 1987. Demographic study of a wild house sparrow population by DNA fingerprinting. *Nature* 327:147-149.
- Williams, J.G., A.R. Kubelik, K.J. Livak, J.A. Rafalski y S.V. Tingey. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research* 18(22):6531-6535.
- Yan, W. 2003. GGE Biplot Analysis: A graphical tool for breeders, geneticists, and agronomist. CRC Press LLC. USA. 267 p.
- Zheng, L., X. Guo, B. He, L. Sun, Y. Peng, S. Dong y T. Liu. 2011. Genome-wide patterns of genetic variation in sweet and grain sorghum (*Sorghum bicolor*) Genome-wide patterns of genetic variation in sweet and grain sorghum (*Sorghum bicolor*). *Genome Biology* 12: R114.
- Zou, G., S. Yan, G. Zhai, Z. Zhang, J. Zou y Y. Tao. 2011. Genetic variability and correlation of stalk yield-related traits and sugar concentration of stalk juice in a sweet sorghum (*Sorghum bicolor* L. Moench) population. *Australian Journal of Crop Science* 5(10):1232-1238.

7. ANEXOS

ANEXO 1. Protocolo de Extracción de ADN vegetal con el método de Möller *et al.*, (1992) modificado (Chen *et al.*, 2009).

1.- Seleccionar de 60 a 100 mg de tejido foliar limpio, moler con nitrógeno líquido con la ayuda de un mortero y poner el tejido en un microtubo de 2 ml conteniendo 500 μ l de buffer TES (Tris-HCl pH 8.0, 100 mM, EDTA 10mM, SDS 2%); 50 μ g de Proteasa K; mezclar vigorosamente e incubar 30 min a 65 °C. Durante la incubación la muestra se agita ocasionalmente.

2.- Agregar 500 μ l de Buffer de extracción CTAB (NaCl 1.4 M, CTAB 1%), incubar por 10 minutos a 65 °C.

3.- Agregar un volumen igual de Cloroformo:isoamílico 24:1 v/v, mezclar cuidadosamente e incubar 30 min a 0 °C, centrifugar a 4 °C, 13000 g² por 10 min.

4.-Transferir el sobrenadante a un microtubo nuevo de 1.5 ml, añadir 225 μ l de NH₄Ac 5M, mezclar cuidadosamente e incubar en hielo 30 min, centrifugar a 4 °C, 13000 g² por 10 min

5.- Sobrenadante a un microtubo nuevo de 1.5 ml, agregar 30 μ g de RNasa e incubar 30 minutos a 37 °C, centrifugar a 4 °C, 13000 g² por 10 min.

6.- Transferir el sobrenadante a un microtubo nuevo de 1.5 ml, agregar 0.55 volumen de isopropanol frío, dejar precipitar de 15 a 30 min en hielo. Centrifugar a 4 °C, 13000 g² por 10 min.

7.- Remover el sobrenadante y lavar la pastilla con etanol frío al 70% dos veces, secar a temperatura ambiente y disolver en 50 µl de buffer TE (Tris-HCl pH 8.0, 100 mM; EDTA 1mM).

ANEXO 2. Secuencia nucleotídica de los iniciadores utilizados en la amplificación selectiva para la generación de AFLP en genotipos de sorgo de la FA-UANL.

Combinación	Iniciador	Secuencia de nucleótidos 5´-3´
*M-CAA x E-AAG	<i>MseI</i>	5'-GAT-GAG-TCC-TGA-GTA-ACA-A-3'
	<i>EcoRI</i>	5'-GAC-TGC-GTA-CCA-ATT-CAA-G-3'
*M-CAA x E-AGG	<i>MseI</i>	5'-GAT-GAG-TCC-TGA-GTA-ACA-A-3'
	<i>EcoRI</i>	5'-GAC-TGC-GTA-CCA-ATT-CAG-G-3'
*M-CAC x E-AAC	<i>MseI</i>	5'-GAT-GAG-TCC-TGA-GTA-ACA-C-3'
	<i>EcoRI</i>	5'-GAC-TGC-GTA-CCA-ATT-CAA-C-3'
*M-CAC x E-ACT	<i>MseI</i>	5'-GAT-GAG-TCC-TGA-GTA-ACA-C-3'
	<i>EcoRI</i>	5'-GAC-TGC-GTA-CCA-ATT-CAC-T-3'
*M-CAG x E-ACA	<i>MseI</i>	5'-GAT-GAG-TCC-TGA-GTA-ACA-G-3'
	<i>EcoRI</i>	5'-GAC-TGC-GTA-CCA-ATT-CAC-A-3'
*M-CAG x E-ACG	<i>MseI</i>	5'-GAT-GAG-TCC-TGA-GTA-ACA-G-3'
	<i>EcoRI</i>	5'-GAC-TGC-GTA-CCA-ATT-CAC-G-3'
*M-CAT x E-ACC	<i>MseI</i>	5'-GAT-GAG-TCC-TGA-GTA-ACA-T-3'
	<i>EcoRI</i>	5'-GAC-TGC-GTA-CCA-ATT-CAC-C-3'
*M-CAT x E-AGC	<i>MseI</i>	5'-GAT-GAG-TCC-TGA-GTA-ACA-T-3'
	<i>EcoRI</i>	5'-GAC-TGC-GTA-CCA-ATT-CAG-C-3'
*M-CTA x E-ACA	<i>MseI</i>	5'-GAT-GAG-TCC-TGA-GTA-ACT-A-3'
	<i>EcoRI</i>	5'-GAC-TGC-GTA-CCA-ATT-CAC-A-3'
*M-CTA x E-ACG	<i>MseI</i>	5'-GAT-GAG-TCC-TGA-GTA-ACT-A-3'
	<i>EcoRI</i>	5'-GAC-TGC-GTA-CCA-ATT-CAC-G-3'
*M-CTC x E-AAC	<i>MseI</i>	5'-GAT-GAG-TCC-TGA-GTA-ACT-C-3'
	<i>EcoRI</i>	5'-GAC-TGC-GTA-CCA-ATT-CAA-C-3'
*M-CTC x E-ACT	<i>MseI</i>	5'-GAT-GAG-TCC-TGA-GTA-ACT-C-3'
	<i>EcoRI</i>	5'-GAC-TGC-GTA-CCA-ATT-CAC-T-3'
*M-CTG x E-AAG	<i>MseI</i>	5'-GAT-GAG-TCC-TGA-GTA-ACT-G-3'
	<i>EcoRI</i>	5'-GAC-TGC-GTA-CCA-ATT-CAA-G-3'
*M-CTG x E-AGC	<i>MseI</i>	5'-GAT-GAG-TCC-TGA-GTA-ACT-G-3'
	<i>EcoRI</i>	5'-GAC-TGC-GTA-CCA-ATT-CAG-C-3'
*M-CTT x E-ACC	<i>MseI</i>	5'-GAT-GAG-TCC-TGA-GTA-ACT-T-3'
	<i>EcoRI</i>	5'-GAC-TGC-GTA-CCA-ATT-CAC-C-3'
*M-CTT x E-AGG	<i>MseI</i>	5'-GAT-GAG-TCC-TGA-GTA-ACT-T-3'
	<i>EcoRI</i>	5'-GAC-TGC-GTA-CCA-ATT-CAG-G-3'
*M-CTA x E-AAG	<i>MseI</i>	5'-GAT-GAG-TCC-TGA-GTA-ACT-A-3'
	<i>EcoRI</i>	5'-GAC-TGC-GTA-CCA-ATT-CAA-G-3'
*M-CTA x E-AGG	<i>MseI</i>	5'-GAT-GAG-TCC-TGA-GTA-ACT-A-3'
	<i>EcoRI</i>	5'-GAC-TGC-GTA-CCA-ATT-CAG-G-3'
*M-CAG x E-AAG	<i>MseI</i>	5'-GAT-GAG-TCC-TGA-GTA-ACA-G-3'
	<i>EcoRI</i>	5'-GAC-TGC-GTA-CCA-ATT-CAA-G-3'
*M-CAG x E-AGG	<i>MseI</i>	5'-GAT-GAG-TCC-TGA-GTA-ACA-G-3'
	<i>EcoRI</i>	5'-GAC-TGC-GTA-CCA-ATT-CAG-G-3'

Anexo 2. Continuación...

*M-CAT x E-AAC	<i>MseI</i>	5'-GAT-GAG-TCC-TGA-GTA-ACA-T-3'
	<i>EcoRI</i>	5'-GAC-TGC-GTA-CCA-ATT-CAA-C-3'
*M-CAT x E-AGG	<i>MseI</i>	5'-GAT-GAG-TCC-TGA-GTA-ACA-T-3'
	<i>EcoRI</i>	5'-GAC-TGC-GTA-CCA-ATT-CAG-G-3'
*M-CAC x E-ACA	<i>MseI</i>	5'-GAT-GAG-TCC-TGA-GTA-ACA-C-3'
	<i>EcoRI</i>	5'-GAC-TGC-GTA-CCA-ATT-CAC-A-3'
*M-CAC x E-AGC	<i>MseI</i>	5'-GAT-GAG-TCC-TGA-GTA-ACA-C-3'
	<i>EcoRI</i>	5'-GAC-TGC-GTA-CCA-ATT-CAG-C-3'
*M-CAA x E-ACA	<i>MseI</i>	5'-GAT-GAG-TCC-TGA-GTA-ACA-A-3'
	<i>EcoRI</i>	5'-GAC-TGC-GTA-CCA-ATT-CAC-A-3'
*M-CAA x E-ACG	<i>MseI</i>	5'-GAT-GAG-TCC-TGA-GTA-ACA-A-3'
	<i>EcoRI</i>	5'-GAC-TGC-GTA-CCA-ATT-CAC-G-3'
*M-CTC x E-ACC	<i>MseI</i>	5'-GAT-GAG-TCC-TGA-GTA-ACT-C-3'
	<i>EcoRI</i>	5'-GAC-TGC-GTA-CCA-ATT-CAC-C-3'
*M-CTC x E-AGC	<i>MseI</i>	5'-GAT-GAG-TCC-TGA-GTA-ACT-C-3'
	<i>EcoRI</i>	5'-GAC-TGC-GTA-CCA-ATT-CAG-C-3'
*M-CTG x E-AAC	<i>MseI</i>	5'-GAT-GAG-TCC-TGA-GTA-ACT-G-3'
	<i>EcoRI</i>	5'-GAC-TGC-GTA-CCA-ATT-CAA-C-3'
*M-CTG x E-ACG	<i>MseI</i>	5'-GAT-GAG-TCC-TGA-GTA-ACT-G-3'
	<i>EcoRI</i>	5'-GAC-TGC-GTA-CCA-ATT-CAC-G-3'

ANEXO 3. Descriptiva de datos marcadores RAPD.

Iniciador	BP	BM	BT	PIC	AMP	PDICMA
A1	5	0	5	0.23	21.54	3.70E-19
A 2	7	0	7	0.20	15.38	6.60E-21
A 3	11	0	11	0.26	25.17	2.10E-18
A 10	5	0	5	0.30	29.23	6.70E-26
A 15	1	0	1	0.23	15.38	7.80E-30
B 15	3	0	3	0.30	25.64	1.40E-27
B 17	9	0	9	0.19	13.68	1.20E-16
B 18	4	0	4	0.18	13.46	8.70E-17
C 1	2	0	2	0.26	19.23	4.90E-30
C 2	2	0	2	0.21	15.38	1.70E-22
C 4	2	0	2	0.13	7.69	0.00E+00
C 6	3	0	3	0.16	10.26	3.90E-20
C 10	3	0	3	0.20	15.38	4.40E-19
C 11	7	0	7	0.25	23.08	2.00E-18
C 12	3	1	4	0.13	5.77	0.00E+00
C 13	2	0	2	0.13	7.69	0.00E+00
C 14	9	0	9	0.13	7.69	0.00E+00
C 15	1	0	1	0.13	7.69	0.00E+00
C 18	3	0	3	0.36	41.03	4.10E-32
C 19	7	0	7	0.19	12.09	1.80E-21
C 20	4	0	4	0.13	7.69	0.00E+00
D 20	2	0	2	0.37	53.85	2.70E-28
E 12	2	0	2	0.13	7.69	0.00E+00
E 14	4	0	4	0.21	17.31	8.40E-16
E 15	6	0	6	0.27	23.08	3.60E-25
E 19	0	1	1	sd	0	1
E 20	7	0	7	0.29	39.56	1.90E-16
F 1	6	0	6	0.21	19.23	2.70E-16
F 2	4	0	4	0.29	25	6.40E-25
F 3	9	0	9	0.24	19.66	6.00E-22
F 4	9	0	9	0.19	13.68	5.90E-17
F 5	7	0	7	0.19	14.29	8.60E-19
F 6	4	0	4	0.13	7.69	0.00E+00
F 7	0	1	1	sd	0	1
F 8	2	0	2	0.25	23.08	2.60E-21
F 9	6	0	6	0.29	34.62	7.00E-24
F 10	3	0	3	0.25	17.95	2.80E-29
I 3	9	0	9	0.26	21.37	2.50E-19
I 6	11	1	12	0.24	18.59	5.40E-19
I 14	9	0	9	0.26	21.37	1.40E-21
K 1	5	0	5	0.25	20.00	6.40E-24
K 3	8	0	8	0.16	10.58	2.00E-19
K 4	3	3	6	0.23	8.97	2.60E-10

Anexo 3. Continuación...

K 6	2	0	2	0.13	7.69	0.00E+00
K 7	7	0	7	0.23	19.78	1.30E-20
K 8	3	1	4	0.13	5.77	0.00E+00
K 9	0	1	1	sd	0	1
K 11	4	0	4	0.16	9.62	6.40E-19
K 14	1	0	1	0.13	7.69	0.00E+00
K 15	5	0	5	0.20	15.38	1.40E-20
K 16	1	0	1	0.23	15.38	7.80E-30
K 19	6	5	11	0.19	6.99	2.70E-10
K 20	4	0	4	0.20	13.46	4.20E-20
Total	242	14	256	0.22	17.34	5.66E-02

Nota: BM=Bandas monomórficas; BP= Bandas polimórficas; BT=Bandas totales; PIC= Índice de contenido polimórfico; AMP=Porcentaje de amplificación; PDICMA= Probabilidad de que dos individuos compartan el mismo alelo; E=Exponencial.

ANEXO 4. Identificación de los productos totales de PCR de fragmentos RAPD recuperados y secuenciados.

Identificación de los fragmentos seleccionados	Nombre del Iniciador OP-	Secuencia de nucleótidos 5'-3'	Tamaño aproximado del Producto total de PCR en pares de bases (pb)
5	A 2	5'-TGCCGA'GCTG-3'	550
6	A 2	5'-TGCCGAGCTG-3'	1285
7	A 2	5'-TGCCGAGCTG-3'	600
9	A 2	5'-TGCCGAGCTG-3'	550
10	A 2	5'-TGCCGAGCTG-3'	500
11	A 2	5'-TGCCGAGCTG-3'	500
12	A 2	5'-TGCCGAGCTG-3'	500
13	A 2	5'-TGCCGAGCTG-3'	500
14	A 2	5'-TGCCGAGCTG-3'	500
16	A 10	5'-GTGATCGCAG-3'	550
17	A 15	5'-TTCCGAACCC-3'	1400
21	B 17	5'-AGGGAACGAG-3'	800
22	B 17	5'-AGGGAACGAG-3'	740
23	C 1	5'-TTCGAGCCAG-3'	1100
25	C 1	5'-TTCGAGCCAG-3'	1100
26	C 1	5'-TTCGAGCCAG-3'	800
27	C 2	5'-GTGAGGCGTC-3'	1000
29	C 4	5'-CCGCATCTAC-3'	690
30	C 6	5'-GAACGGACTC-3'	950
31	C 6	5'-GAACGGACTC-3'	770
32	C 6	5'-GAACGGACTC-3'	680
43	C 14	5'-TGCGTGCTTG-3'	950
44	C 14	5'-TGCGTGCTTG-3'	1250
45	C 14	5'-TGCGTGCTTG-3'	780
46	C 14	5'-TGCGTGCTTG-3'	700
47	C 14	5'-TGCGTGCTTG-3'	540
48	C 19	5'-ACCCCCGAAG-3'	1350
49	C 19	5'-ACCCCCGAAG-3'	1030
50	E 15	5'-ACGCACAACC-3'	500
52	E 20	5'-AACGGTGACC-3'	500
53	F 1	5'-ACGGATCCTG-3'	900
54	F 1	5'-ACGGATCCTG-3'	400
60	F 4	5'-GGTGATCAGG-3'	400
62	F 4	5'-GGTGATCAGG-3'	680
63	F 4	5'-GGTGATCAGG-3'	220
66	I 3	5'-CAGAAGCCCA-3'	600
67	I 3	5'-CAGAAGCCCA-3'	600
68	I 6	5'-AAGGCGGCAG-3'	250
70	I 6	5'-AAGGCGGCAG-3'	460

Anexo 4. Continuación...

71	I 14	5'-TGACGGCGGT-3'	700
72	I 14	5'-TGACGGCGGT-3'	1400
73	I 14	5'-TGACGGCGGT-3'	450
74	K 1	5'-CATTCGAGCC-3'	720
75	K 1	5'-CATTCGAGCC-3'	720
87	K 15	5'-CTCCTGCCAA-3'	900
88	K 15	5'-CTCCTGCCAA-3'	850
90	K 19	5'-CACAGGCGGA-3'	400
91	K 19	5'-CACAGGCGGA-3'	500
92	K 20	5'-GTGTCGCGAG-3'	900
93	K 20	5'-GTGTCGCGAG-3'	600
94	K 20	5'-GTGTCGCGAG-3'	600
95	K 20	5'-GTGTCGCGAG-3'	600
