

**UNA APROXIMACIÓN AL MAPA GENÉTICO DE LA CAÑA DE AZÚCAR
(*Saccharum* spp.) PARA LAS VARIETADES COLOMBIANAS USANDO
MARCADORES MOLECULARES MICROSATÉLITES**

NATALIA CAMPILLO PEDROZA



UNIVERSIDAD DEL QUINDÍO
FACULTAD DE CIENCIAS BÁSICAS Y TECNOLOGÍAS
PROGRAMA DE BIOLOGÍA
ARMENIA
2009

**UNA APROXIMACIÓN AL MAPA GENÉTICO DE LA CAÑA DE AZÚCAR
(*Saccharum* spp.) PARA LAS VARIEDADES COLOMBIANAS USANDO
MARCADORES MOLECULARES MICROSATÉLITES**

NATALIA CAMPILLO PEDROZA

UNIVERSIDAD DEL QUINDÍO
FACULTAD DE CIENCIAS BÁSICAS Y TECNOLOGÍAS
PROGRAMA DE BIOLOGÍA
ARMENIA
2009

**UNA APROXIMACIÓN AL MAPA GENÉTICO DE LA CAÑA DE AZÚCAR
(*Saccharum* spp.) PARA LAS VARIEDADES COLOMBIANAS USANDO
MARCADORES MOLECULARES MICROSATÉLITES**

NATALIA CAMPILLO PEDROZA

Trabajo de grado presentado como requisito parcial para optar al título de
Biólogo (a)

Director
JERSHON LÓPEZ GERENA
Biólogo, Ph.D.

UNIVERSIDAD DEL QUINDÍO
FACULTAD DE CIENCIAS BÁSICAS Y TECNOLOGÍAS
PROGRAMA DE BIOLOGÍA
ARMENIA
2009

Nota de aprobación

El trabajo de grado titulado “una aproximación al mapa genético de caña de azúcar (*Saccharum* spp.) para las variedades colombianas usando marcadores moleculares microsátélites”, presentado por la estudiante NATALIA CAMPILLO PEDROZA, para optar al título de Biólogo (a), fue revisado por el jurado y calificado como:

JERSHON LÓPEZ GERENA
Director

NHORA CECILIA RODRÍGUEZ
Jurado

EDWIN DAVID MORALES
Jurado

HEIBER CÁRDENAS
Jurado

AGRADECIMIENTOS

A Dios por ser la fuerza que siempre me impulso para continuar a pesar de las dificultades.

A la mujer que ha dado todo por mi, quién ha sido mi hombro, mi fuerza de apoyo, a quien le debo todo lo que soy, el tesoro más grande de mi vida, mi madre.

A mi hermana Diana mi amiga y fiel acompañante, por todas sus palabras y apoyo incondicional durante este duro proceso.

A Eliana Macea por haberme brindado su amistad, por todos sus aportes y ayuda en este trabajo.

A los profesores Maria Dolly García y Hernando Hurtado por su valiosa asesoría y acompañamiento al final de este trabajo.

A CENICAÑA y COLCIENCIAS quienes financiaron este proyecto.

TABLA DE CONTENIDO

	Página
1. INTRODUCCIÓN	1
2. MARCO TEÓRICO	5
2.1. Generalidades de la caña de azúcar	5
2.2. Morfología	5
2.3. Taxonómica citogenética y genómica	8
2.4. Marcadores moleculares	12
2.4.1. Tipos de marcadores moleculares	14
2.4.1.1. Marcadores moleculares basados en la secuencia del ADN	15
2.5. Mapeo genético en plantas	21
2.5.1. Poblaciones de mapeo	23
2.5.2. Población de mapeo F2	24
2.5.3. Identificación de polimorfismo	25
2.5.4. Análisis de ligamiento	26
2.5.5. Mapeo genético en caña de azúcar	28
2.5.6. Análisis de QTL	30
2.5.6.1. Análisis de un sólo marcador	33
2.5.6.2. Mapeo por intervalo simple	33
2.5.6.3. Mapeo por intervalo compuesto	34
2.5.6.4. Mapeo por intervalo múltiple	34
3. OBJETIVOS	35
3.1. General	35
3.2. Específicos	35
4. METODOLOGÍA	36
4.1. Material Experimental	36
4.1.1. Población de mapeo	36
4.1.2. Características analizadas	37
4.1.3. Marcadores moleculares y detección	38
4.1.4. Análisis de ligamiento	42

4.1.5. Análisis de características cuantitativas QTL	43
5. RESULTADOS Y DISCUSIÓN	45
6. CONCLUSIONES	68
7. RECOMENDACIONES	72
8. BIBLIOGRAFÍA	73
ANEXO A UBICACIÓN GEOGRÁFICA DE LA HACIENDA CACHIMBALITO (INGENIO INCAUCA)	78
ANEXO B. MÉTODO DE EXTRACCIÓN DE ADN DE CAÑA DE AZÚCAR (SACCHARUM SPP)	79
ANEXO C PROTOCOLO DE TINCIÓN CON PLATA PARA GELES DE POLIACRILAMIDA	81

INDICE DE TABLAS

	Página
Tabla 1. Características generales de los miembros del género <i>Saccharum</i> , se muestran los niveles de ploidía y aneuploidia.	9
Tabla 2. Mapas genéticos publicados hasta la fecha en orden cronológico.	29
Tabla 3. Se relacionan los componentes de la producción de sacarosa y biomasa que fueron analizados en el presente estudio y la edad del cultivo en la que fueron evaluadas.	38
Tabla 4. Marcadores microsatélites evaluados.	41
Tabla 5. Marcadores evaluados en este estudio y su ubicación en previos Reportes.	44
Tabla 6. Asociaciones estadísticamente significativas para cuatro Marcadores.	51
Tabla 7. Asociación entre marcadores SSR y las características analizadas.	52
Tabla 8. Marcadores asociados a diferentes características a la misma edad del cultivo.	53
Tabla 9. Se muestran los resultados del análisis de regresión logística incluyendo solo los marcadores asociados identificados con el ANOVA, para altura a los 4 meses.	58
Tabla 10. Se muestran los resultados del análisis de regresión logística incluyendo todos los alelos marcadores evaluados en éste estudio (45), para altura a los 4 meses	59
Tabla 11. Se muestran los resultados del análisis de regresión logística incluyendo solo los alelos marcadores identificados con ANOVA, para porcentaje de fibra a los 12 meses.	59
Tabla 12. Se muestran los resultados del análisis de regresión logística incluyendo todos los alelos marcadores evaluados en éste estudio (45), para porcentaje de fibra a los 12 meses.	60

Tabla 13. Se muestran los resultados del análisis de regresión logística incluyendo solo los alelos marcadores identificados con ANOVA, para clorofila a los 4 meses.	62
Tabla 14. Se muestran los resultados del análisis de regresión logística incluyendo todos los alelos marcadores evaluados en éste estudio (45), para clorofila a los 4 meses.	63
Tabla 15. Se muestra la razón de probabilidad para la asociación de los alelos marcadores en el análisis de regresión logística, donde se incluyen sólo los alelos marcadores asociados identificados con el ANOVA y el análisis incluyendo los 45 alelos marcadores.	64
Tabla 16. Se muestra la asociación individual de marcadores encontrada con el ANOVA, y las combinaciones de alelos asociados a las características con porcentajes de explicación de la variación fenotípica alta y modelos estadísticamente significativos encontrados con el análisis de regresión logística.	66

INDICE DE FIGURAS

	Página
Figura 1. Imagen de la caña de azúcar.	6
Figura 2. Organización esquemática del genoma del cultivar francés R570 de caña de azúcar.	11
Figura 3. Árbol de la evolución de los marcadores moleculares	14
Figura 4. Se muestra el polimorfismo del microsatélite en la secuencia y cómo es detectado con electroforesis	18
Figura 5. Pasos en la identificación y caracterización de Características cuantitativas (QTL) para su uso en los programas de mejoramiento genético	32
Figura 6. Esquema de la organización de los ADN en el gel de poliacrilamida	39
Figura 7. Grupo de ligamiento formado con un LOD 3 y fracción de recombinación 4.	45

1. INTRODUCCIÓN

El sector azucarero ha tenido un impacto importante en el desarrollo socio - económico del país y de la región, .está conformado por trece ingenios azucareros, más de 1.500 agricultores, más de cuarenta empresas de alimentos y bebidas, once productores de alcohol y licores, dos co - generadores de energía, un productor de papel, una industria sucro-química y una gran cantidad de proveedores de bienes y servicios que el gremio demanda (Asocaña 2001-2002).

En el año 2008 el sector generó cerca de 36.000 empleos directos y 216.000 indirectos; a nivel económico el valor de la producción de azúcar y alcohol en ese mismo año fué de 1.100 millones de dólares (calculado con el promedio del dólar de ese año), de los cuales cerca de 200 millones provinieron de las exportaciones de azúcar hacia Chile, Perú, Estados Unidos, Haití entre otros; estos ingresos permitieron hacer aportes a la economía nacional representados en 1% del PIB total, 3% del PIB Industrial y el 4% del PIB Agrícola, mientras que para la región, éstas cifras correspondieron al 6%, 12% y 47%, respectivamente (<http://www.asocana.com.co/publico/historia.aspx> consultada el 1 de Marzo de 2009).

El gremio azucarero se ha desarrollado principalmente en el valle del río Cauca, de las 406.000 hectáreas que comprende el valle, cerca de 205.664 hectáreas

están dedicadas al cultivo de la caña de azúcar, el área sembrada se distribuye en los departamentos de Cauca, Caldas, Risaralda y Valle del Cauca (Quintero., *et al.* 2008); las condiciones agro – climáticas de estas regiones han permitido cosecha y molienda de caña de azúcar durante todo el año, lo que ha posibilitado la especialización en este cultivo, obteniéndose una de las más altas productividades en el mundo. (<http://www.asocana.com.co/publico/historia.aspx> consultada el 1 de Marzo de 2009)

Para desarrollar unidades productivas empresariales que estén avanzando hacia una agricultura cada vez más moderna en busca de la máxima eficiencia del cultivo, se ha hecho necesario el uso de nuevas tecnologías en el desarrollo de variedades de caña más productivas que tengan en cuenta las condiciones propias del terreno.

(http://www.agrocadenas.gov.co/azucar/documentos/caracterizacion_azucar.pdf. consultado el 1 de Marzo de 2009).

La generación de variedades se ha realizado durante años, siguiendo un proceso cíclico que comprende tres pasos (Semagn *et al.*, 2006c):

1. Reunir los recursos de diversidad genética necesarios para el proceso de mejoramiento a partir de recursos adaptados o exóticos y realizar recombinaciones para generar nuevas combinaciones génicas.
2. Selección y evaluación para identificar recombinantes superiores, lo que incluye el tiempo de selección (generación tardía o generación temprana), selección del ambiente (favorable y estrés), el número de años y localización del estudio.
3. Liberación, distribución y comercialización de los nuevos cultivares.

Las principales limitaciones en este proceso de mejoramiento son:

1. El número de generaciones y el largo tiempo necesario para alcanzar el objetivo de introgresión.
2. La transferencia simultánea de otros genes flanqueantes del gen de interés a partir del parental donador.

Para la solución de estas limitaciones se ha desarrollado una herramienta que asiste al fitomejorador en la toma de decisiones, conocida como Selección Asistida por Marcadores (MAS, Marker Assisted Selection) (Semagn *et al.*, 2006c).

En la selección asistida por marcadores, el uso de los marcadores moleculares reduce los problemas convencionales en el mejoramiento de plantas tradicional; se usa la información obtenida del marcador molecular como guía para

seleccionar los individuos a utilizar dentro del programa de mejoramiento como parentales o variedades élite, con esto es posible disminuir el tiempo de los ciclos y aumentar la eficiencia del proceso (Semagn *et al.*, 2006c).

Implementar la selección asistida por marcadores dentro de los programas de mejoramiento implica dar un primer paso y este es el desarrollo de mapas genéticos con marcadores moleculares para realizar análisis de QTL (quantitative trait loci ó loci de características cuantitativas) con el fin de identificar asociaciones de los marcadores analizados con las características de importancia agronómica para la producción.

En el presente trabajo se realizó una aproximación al mapa genético de las variedades de caña de azúcar (*Saccharum spp*) para las variedades colombianas utilizando marcadores moleculares microsatélites incluidos dentro de un análisis de QTL.

Este estudio es la primera aproximación genómica que se realiza para las variedades de caña de azúcar colombianas, aporta información preliminar que contribuye con la aplicación a futuro de la selección asistida por marcadores en los programas de mejoramiento colombianos para el desarrollo de nuevas variedades de caña de azúcar que satisfagan las necesidades productivas del sector.

2. MARCO TEÓRICO

2.1 Generalidades de la caña de azúcar

El sector azucarero es considerado como una de las agroindustrias más importantes en el país por sus significativos aportes en materia económica y social.

El azúcar utilizada como materia prima dentro del sector, proviene casi exclusivamente de la caña de azúcar (*Saccharum spp.*).

2.2 Morfología

La apariencia de la caña de azúcar en general es muy similar a la apariencia de otras gramíneas, su hábito de crecimiento se caracteriza por un agrupamiento de tallos sólidos formando cepas (figura 1). Un tallo normal de caña esta formado por varios nudos, donde se localiza la yema, y un entrenudo. El número, diámetro, color y hábito de crecimiento del tallo dependen principalmente de las variedades.

El tamaño o longitud de los tallos depende, en gran parte, de las condiciones agro ecológicas de la zona donde crece y del manejo que se le de a la variedad, sin embargo, en promedio un individuo maduro puede medir de 1.5 a 3 metros de altura y tener un diámetro de 2.5 a 7.5 cm. (Kole, 2007; CENICAÑA, 1995).



Figura 1. Imagen de la caña de azúcar mostrando su morfología, se observa la flor y su inflorescencia, la morfología de la hoja y el tallo.

Las hojas de la caña de azúcar se originan en los nudos y se distribuyen en posiciones alternas a lo largo del tallo a medida que éste crece. Cada hoja está formada por la lámina foliar y por la vaina. La unión entre estas dos partes se denomina lígula y en cada extremo de ésta existe una aurícula con pubescencia variable. La forma y el color de la lígula, así como la forma de la aurícula, son características importantes en la diferenciación de las variedades de la caña de azúcar. (CENICAÑA, 1995).

La caña de azúcar presenta dos fases de desarrollo. La vegetativa, originada por la división celular en los puntos de crecimiento; y la reproductiva o de floración, que ocurre cuando las condiciones ambientales de foto período, temperatura,

disponibilidad de agua y nivel de nutrimentos en el suelo son favorables (CENICAÑA, 1995).

La inflorescencia de la caña de azúcar es una panícula sedosa en forma de espiga. Está constituida por un eje principal con articulaciones en las cuales se insertan las espiguillas, una frente de la otra; éstas contienen una flor hermafrodita con tres anteras y un ovario con dos estigmas. Cada flor está rodeada de pubescencias largas dando a la inflorescencia un aspecto sedoso. En cada ovario hay un óvulo el cual, una vez fertilizado, da origen al fruto o cariósido. Por lo tanto lo que comúnmente se conoce como semilla es una cariósido. El fruto es de forma ovalada de 0.5 mm. de ancho y 1.5 mm. de largo, aproximadamente (CENICAÑA, 1995).

Las exigencias de humedad y variación de temperatura para obtener los máximos niveles de sacarosa han llevado a que en la mayor parte de las regiones azucareras del mundo, con excepción del valle del Cauca, Hawai y Perú la cosecha de caña se realice únicamente durante una época del año, conocida como zafra.

A diferencia de lo que sucede en los países no tropicales, en los países tropicales la caña de azúcar puede ser cultivada todo el año, y por lo tanto puede crecer continuamente hasta la cosecha (Kole, 2007).

2.3 Taxonómica citogenética y genómica.

La caña de azúcar está clasificada taxonómicamente así:

Familia: Poaceae

Sub- Familia: Panicoideae

Tribu: Andropogoneae

Sub- Tribu: Saccharastrae

Género: Saccharum

En la clasificación original de Linneo se encuentran seis especies dentro del género *Saccharum*: *S. officinarum* conocida como la caña noble, *S. spontaneum*, *S. robustum*, *S. edule*, especies silvestres, *S. barberi*, y *S. sinensi* conocidos como híbridos antiguos. (Mukherjee, 1957, Kole, 2007).

Cuatro géneros están estrechamente relacionados con las seis especies de *Saccharum*: *Erianthus*, *Miscanthus*, *Narenga*, y *Sclerostachya*, probablemente estos géneros han contribuido al origen genético de la caña de azúcar, por esta razón los cuatro géneros y las seis especies hacen parte de lo que actualmente se conoce como complejo *Saccharum*. (Mukherjee, 1957, Kole, 2007, Jannoo *et al.*, 1999).

En general, el complejo *Saccharum* se caracteriza por tener alto nivel de ploidía (individuos con más de dos dotaciones cromosómicas básicas) y aneuploidía

(pérdida o ganancia de uno o más cromosomas pero no la dotación cromosómica completa) (Grivet & Arruda. 2001).

En particular el género *Saccharum*, puede ser visto como un complejo cuya diversidad genética está relacionada con diferentes niveles de ploidía causados por hibridaciones ínter específicas o ínter genéricas (Tabla 1) (Jannoo *et al.*, 1999).

Tabla 1 Características generales de los miembros del género *Saccharum*, se muestran los niveles de ploidía y aneuploidía. Tomado de: Buzacott 1965 & Daniela *et al.*, 1987 citado por Kole. 2007.

Especies	Clasificación	Número de cromosomas
<i>S. barberi</i>	Hibrido antiguo	2n= 11-120 2n= 60-80 con
<i>S. edule</i>	especie silvestre	formas aneuploides
<i>S. officinarum</i>	Caña noble	2n= 80
<i>S. robustum</i>	especie silvestre	2n= 60 – 200
<i>S. sinensi</i>	Hibrido antiguo	2n= 80-140
<i>S. spontaneum</i>	especie silvestre	2n=40-128

Hasta finales del siglo XIX, *S. officinarum*, *S. barberi* y *S. sinense* hicieron parte de los cultivares comerciales, pero la necesidad de obtener especies resistentes a diversas enfermedades, el aumentar los niveles de producción de azúcar y el descubrimiento de la fertilidad sexual en los híbridos, llevó a los mejoradores de Java e India a realizar los primeros híbridos interespecíficos, donde se

involucraron especialmente *S. officinarum*, *S. spontaneum* y *S. barberi* (Grivet & Arruda, 2001).

A partir de repetidos retrocruzamientos de híbridos con *S. officinarum* se obtuvieron individuos que recuperaron su alta producción de azúcar y resistencia a enfermedades. Durante este proceso conocido como nobilización, *S. officinarum* usado como parental hembra transfirió su dotación somática ($2n$) a la F_1 del primer retrocruzamiento. La fusión de núcleos o endoduplicación en la meiosis II ha sido sugerida por Bath y Gill (1985), como explicación para esta atípica transmisión; por su parte *S. spontaneum* transfirió su número haploide (n). A partir del segundo retrocruzamiento la transmisión de cromosomas fue normal (Jannoo *et al.*, 1999); por lo tanto las variedades comerciales actuales son el producto de la nobilización y posteriores cruzamientos interespecíficos, seguidos de sucesivos procesos de selección.

A partir de los procesos mencionados anteriormente, en Colombia se tienen variedades como ICA 69 -11 proveniente del cruzamiento de la variedad D 200/36 (madre) con Co 775 (padre) y MZC 74-275 originada a partir del cruzamiento de las variedades POJ 2878 (madre) y CP 57-603 (padre). La variedad MZC 74-275 es alta productora de sacarosa, se encuentra actualmente en 0.9% del área total sembrada en caña de azúcar (CENICAÑA), mientras que la variedad ICA 69 -11 tiene baja producción de azúcar y actualmente no esta sembrada en el país, estas

variedades fueron seleccionadas en el presente estudio como parentales para producir la población de mapeo por su diferencia en la producción de sacarosa.

Se estima a partir de estudios genómicos y citogenéticas desarrollados en el cultivar francés R570, que las variedades actuales de caña de azúcar se caracterizan por un alto número de cromosomas ($2n=100-130$) donde el 70 - 80% son aportados por *S. officinarum*, el 10 – 20% por *S. spontaneum* y el 10% son recombinantes de las dos especies (Figura 2) (D'Hont *et al.*, 1996), es importante resaltar que si bien existe recombinación, ocurre en una baja proporción.

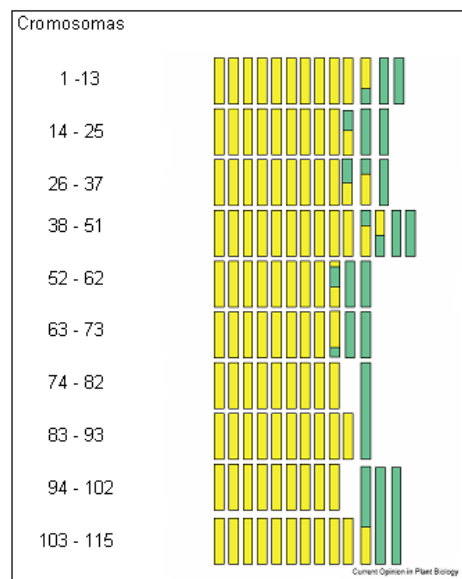


Figura 2. Organización esquemática del genoma del cultivar Francés R570 de caña de azúcar, cada barra representa un cromosoma; las regiones en amarillo representan las regiones originadas a partir de *S. officinarum* y las barras verdes las que se originaron de *S. spontaneum*. Los cromosomas alineados en la misma fila son homólogos. Tomado de: Grivet y Arruda. 2001

El tamaño genómico estimado para el cultivar Francés R570 ($2n=115$) es de 10.000 Mpb; para *S. officinarum* se estima una tamaño genómico de 930 Mpb, y para *S. spontaneum* de 750 Mpb, esto indica que el genoma de esta variedad de caña de azúcar es significativamente más grande que el de sus parentales y las especies de gramíneas relacionadas, tales como arroz (860 Mpb, $2n=24$), sorgo (1600 Mb, $2n=20$) o maíz (5500 Mb, $2n=20$), esta información da evidencia del alto nivel de ploidía de los cultivares de caña de azúcar (Grivet & Arruda, 2001 ; D`Hont & Glazmann, 2001 citados por Kole, 2007);

Por su alto nivel de ploidía, aneuploidía, origen biespecífico de los cromosomas, la existencia de diferencias estructurales entre cromosomas de las dos especies y los procesos de recombinación ínter específica ocurridos durante el mejoramiento de la caña de azúcar, esta puede catalogarse como una de las estructuras genómicas más complejas en las plantas de importancia agronómica (Grivet & Arruda, 2001; D`Hont *et al.*, 1996).

2.3 Marcadores Moleculares

Se define como marcador molecular a todo segmento específico de ADN (Ferreira y Grattapaglia, 1998; Semagn *et al.*, 2006a) que permite detectar diferencias entre individuos en el genoma, generalmente ellos no son genes blanco, pero si pueden encontrarse cerca o pueden estar ligados a genes de interés, actuando de esta

manera como “señas” o “marcas” dentro del genoma (Jones *et al.*, 1997; Collard *et al.*, 2005).

Cuando a un marcador molecular se le verifica un comportamiento de acuerdo con las leyes básicas de la herencia Mendeliana se define como marcador genético (Ferreira y Grattapaglia, 1998).

Los marcadores moleculares han “evolucionando” paulatinamente muy de la mano con las necesidades y desafíos que constantemente se presentan en la investigación, la figura 3 muestra la historia del sistema de marcadores genéticos existente y los que se esperan a futuro.

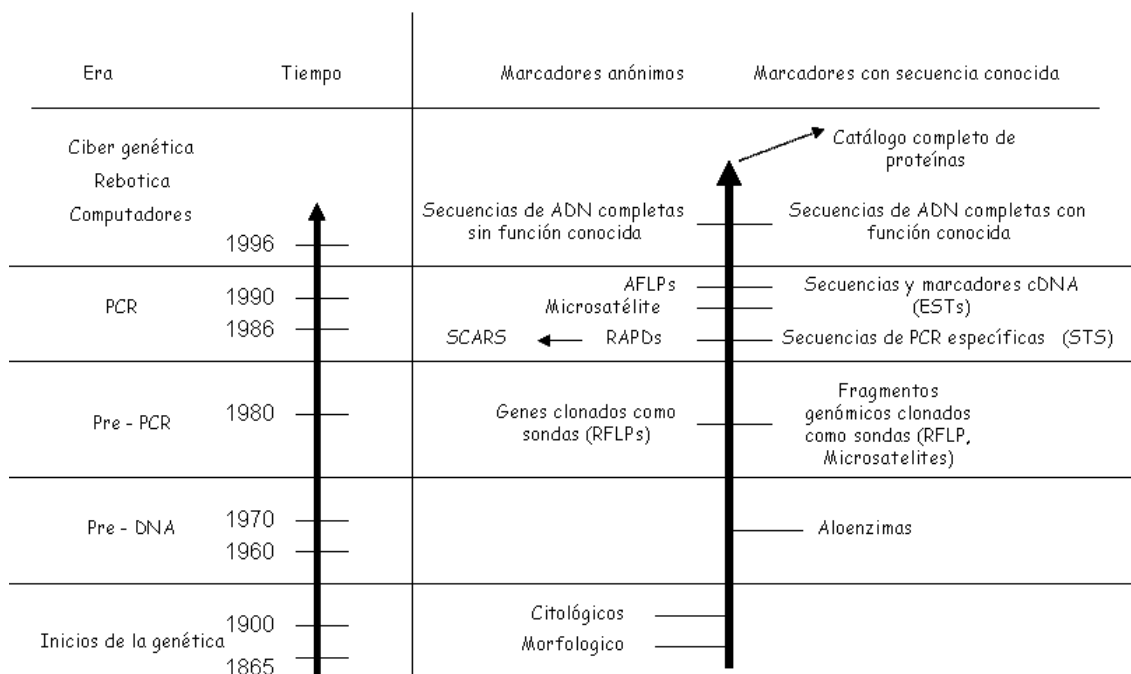


Figura 3 Árbol de la evolución de los marcadores moleculares basado en el original realizado por el Dr. David O’Melley de la universidad de Carolina del Norte. Tomado de: Hui, 1998.

2.3.1 Tipos de marcadores moleculares

En la actualidad existen tres tipos de marcadores moleculares, morfológicos, bioquímicos y de ADN (Collard *et al.*, 2005).

Inicialmente los marcadores moleculares usados para el análisis y mejoramiento genético fueron los marcadores morfológicos, son características fenotípicas de fácil identificación visual como enanismo, deficiencia clorofílica, color de pétalos o morfología foliar que son producto de la expresión de un solo gen, (Ferreira y Grattapaglia, 1998). Por lo tanto al hacer la observación de la característica fenotípica se estaba observando el gen con sus variantes alélicas determinadas y su segregación en la población de estudio.

Los marcadores morfológicos no son muy numerosos entre individuos de la misma especie, adicionalmente el mayor número de marcadores de este tipo son disponibles sólo para algunas plantas modelo, siendo una limitante para su uso en otras especies de las que no se tiene información a este nivel. Por esta razón se desarrollaron otro tipo de marcadores conocidos como marcadores bioquímicos o isoenzimáticos, basados en las diferencias encontradas en las formas moleculares de la misma enzima (isoenzimas) (Ferreira y Grattapaglia, 1998).

Las isoenzimas cumplen la misma función catalítica pero difieren en su estructura molecular, estas diferencias en su estructura son las que se emplean como marcadores entre individuos (Ferreira y Grattapaglia, 1998).

Las dificultades con estas técnicas están relacionadas con el hecho de que el análisis es a nivel del producto de la expresión del gen determinado y no directamente en el segmento de ADN usado como marcador, lo que introduce un grado de incertidumbre mayor debido a que el producto de la expresión del gen, en estos casos las características fenotípicas observadas (marcadores moleculares) y las diferentes estructuras moleculares de las isoenzimas (marcadores bioquímicos), pueden ser afectadas por variables externas, como interacciones epigenéticas ó modificaciones postranscripcionales, entre otras, pudiéndose presentar casos donde sean evidentes las diferencias, pero que estas no sean dadas directamente por diferencias a nivel genético, sino por procesos posteriores a la expresión del gen.

2.3.1.1 Marcadores moleculares basados en la secuencia del ADN

Con la ayuda de las nuevas herramientas de la biología molecular, principalmente el descubrimiento de la reacción en cadena de la polimerasa (PCR) por Kary Mullis en 1983 (Segman *et al.*, 2006a), se fueron desarrollando nuevos marcadores moleculares, basados esta vez en el análisis directo de la secuencia de ADN.

Esta nueva clase de marcadores basados en la secuencia del ADN son selectivamente neutros, es decir, a pesar de que generalmente se encuentran en regiones no codificantes, cuando se ubican en regiones que si codifican, el producto fenotípico obtenido por la expresión de secuencias que han sido modificadas por procesos como mutaciones, translocaciones, deleciones, entre otras, (modificaciones responsables de las diferencias entre individuos) no son afectadas por el ambiente ni por estados de desarrollo de la planta, esta característica permite tener información a nivel genómico más objetiva, al haber la certeza de que dicha información no es afectada por variables o procesos externos.

En caña de azúcar los marcadores que se han empleado para desarrollar mapas genéticos han sido principalmente los RFLP, RAPD, AFLP, AP-PCR, SSR y EST-SSR (Grivet *et al.*, 1996; Da Silva *et al.*, 1995; Al-Janabi *et al.*, 1993; Hoarau *et al.*, 2001; Aitken *et al.*, 2005; Oliveira *et al.*, 2007, Singh *et al.*, 2006, Garcia *et al.*, 2006).

Según Chambers and MacAvoy (2000) (citado por Semagn *et al.*, 2006a) los microsatélites también conocidos como secuencias simples repetidas (SSR, Simple Sequence Repeat) son, pequeñas secuencias repetidas (secuencias motivo) de 2 - 6 pb (pares de base) flanqueadas por secuencias altamente conservadas (figura 4).

Los microsatélites tienen un alto nivel de variación en muchas especies de animales y plantas. La forma común de estas repeticiones es de dos nucleótidos (di - nucleótidos), como $(CA)_n:(GT)_n$, $(GA)_n:(CT)_n$, $(CG)_n:(GC)_n$ y $(AT)_n:(TA)_n$ donde n es el número de repeticiones. Los microsatélites con tres y cuatro nucleótidos (tri- y tetranucleotidos, respectivamente), también son encontrados pero en menor frecuencia (Hui, 1998).

La principal causa del polimorfismo en los microsatélites se debe a un tipo de mutación denominada “slipped-strand mispairing” este proceso de mutación ocurre durante la replicación, involucra la denaturación y el desplazamiento de las cadenas de ADN, dando como resultado un mal apareamiento de las bases complementarias, de esta manera se aumenta o disminuye el número de las secuencias repetidas, siendo esta la base del polimorfismo entre individuos (Levinson and Gutman, 1987; citado por Segman *et al.*, 2006a).

Gracias a que los motivos repetidos del microsatélite son flanqueados por secuencias conservadas, es posible diseñar iniciadores complementarios a dichas secuencias, con el fin de amplificar por medio de PCR el fragmento flanqueante y el marcador, de acuerdo a lo anterior el polimorfismo en los microsatélites es detectado al amplificar la región que contiene las secuencias repetidas, de esta manera cada microsatélite independiente del elemento repetido constituye un *locus* genético altamente variable y multialélico; cada segmento amplificado de tamaño diferente representa un alelo del mismo *locus*; estos fragmentos son

separados en geles de poliacrilamida o agarosa especial de alta resolución, permitiendo separar fragmentos que difieren en pocas pares de bases; las bandas son visualizadas directamente por coloración con nitrato de plata (AgNO_3), bromuro de etidio, a través de autorradiografía o sistemas de detección fluorescente, utilizando iniciadores marcados con radioisotopos o moléculas fluorescentes en la PCR (Ferreira y Grattapaglia, 1998).

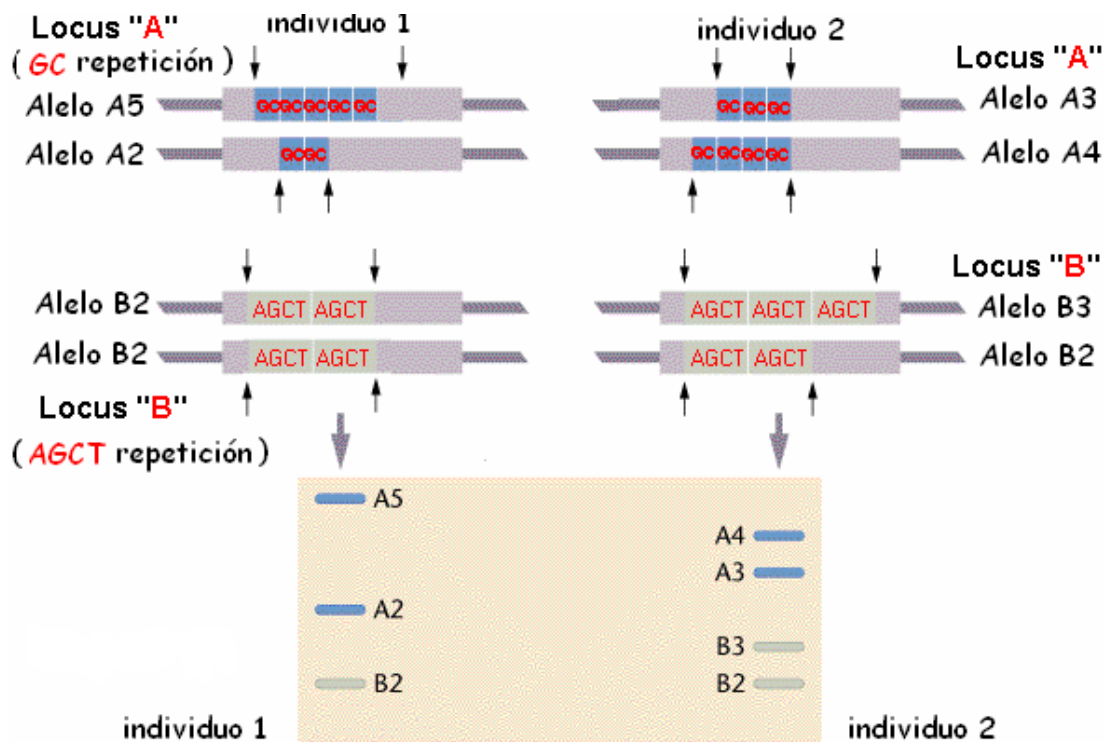


Figura 4. Se muestra el polimorfismo del microsatélite en la secuencia y cómo es detectado con electroforesis. Tomado de: Griffiths et al., 1999.

Los microsatélites han ganado especial importancia por sus diferentes atributos genéticos como hipervariabilidad, amplia distribución genómica, herencia codominante (permite visualizar los genotipos heterocigotos), reproducibilidad,

naturaleza multialélica y localización específica en el cromosoma, aunque en caña de azúcar ellos podrían estar en diferentes cromosomas debido al alto nivel de ploidía.

En caña de azúcar el ISMC (International sugarcane Microsatellite Consortium) desarrolló cerca de 200 marcadores microsatélites, Parida *et al.*, (2009), generaron 1,315 marcadores con un nivel de polimorfismo detectado entre especies, género y variedades de caña de azúcar de 88.6%.

Otro marcador que también ha cobrado gran importancia en mapeo genético de plantas son los derivados a partir de los ESTs, estos representan porciones de genes expresados en un estado de desarrollo determinado, inicialmente fueron creados para identificar genes transcritos, pero han sido utilizados en el descubrimiento de genes, en la obtención de datos en expresión de genes y regulación, determinación de secuencias, y en el desarrollo de marcadores moleculares altamente informativos como EST - RFLP, que han sido ampliamente usados en la elaboración de mapas de ligamiento genético de alta densidad, (Harushima *et al.*, 1998; Davis *et al.*, 1999, citado por semagn *et al.*, 2006a) y mapas físicos (Kurata *et al.*, 1997, citado por semagn *et al.*, 2006a).

En caña de azúcar el proyecto SUCEST desarrollado por el grupo de Brasil involucró trabajo conjunto de 24 laboratorios de secuenciación, un laboratorio bioinformático, un laboratorio coordinador, 50 grupos encargados de la búsqueda

de datos, y un grupo de relaciones internacionales, desarrollaron 291,904 secuencias a partir de 260,352 clones de 37 librerías diferentes (Telles *et al.*, 2001).

A partir de estas librerías se han desarrollado SSR que han sido utilizados en la elaboración de mapas de ligamiento genético y mapeo de QTL de importancia agronómica (Oliveira *et al.*, 2007).

A pesar de que el desarrollo de SSR a partir de secuencias EST sea más rápido, económico, eficaz, y su utilidad sea muy alta en términos de la calidad y cantidad de información que proporciona, se ha reportado que los marcadores EST- SSR tienen un bajo polimorfismo comparado con los marcadores SSR derivados de librerías genómicas (Cho *et al.*, 2000; Scott *et al.*, 2000; Eujayl *et al.*, 2002; Chabane *et al.*, 2005, citados por Semagn *et al.*, 2006a).

Teniendo en cuenta lo anterior, el uso de un marcador determinado para realizar mapeo genético depende principalmente de los recursos económicos disponibles, de la información ya existente de la especie a estudiar y la información que proporciona el marcador; de acuerdo a esto para la elaboración inicial de mapas genéticos, los marcadores ideales son los AFLP por el alto número de bandas polimórficas, que se generan; para la saturación de mapas y su aplicación en la identificación de genes involucrados en la característica de interés, marcadores

mas específicos de la secuencia como los SSR y los EST – SSR pueden ser una buena alternativa.

2.9 Mapeo genético en plantas

Un mapa genético es un modelo abstracto de la organización lineal de un grupo de genes y marcadores (Hui, 1998), los mapas genéticos indican la posición y la distancia genética relativa entre marcadores a lo largo de un cromosoma o grupo de ligamiento.

La segunda ley de Mendel plantea que en la formación de los gametos los alelos de cada gen se segregan al azar, de tal manera que cada gameto recibe uno u otro alelo con igual probabilidad. Sin embargo es bien sabido que los genes se encuentran dentro de cromosomas, y que estos y no los genes son la unidad de transmisión en la meiosis. Los genes ligados, es decir, los genes ubicados en el mismo grupo de ligamiento, no son libres para transmitirse de manera independiente, por el contrario los alelos de todos los loci de un cromosoma se transfieren como una unidad en la formación de los gametos (Klug *et al.*, 2006).

Sin embargo en la profase I de la meiosis, específicamente en el diploteno, ocurre el entrecruzamiento entre cromátidas no hermanas, dando lugar a una mezcla o recombinación de los alelos presentes en los cromosomas homólogos (Klug *et al.*, 2006).

El grado de entrecruzamiento entre dos loci es proporcional a la distancia que los separa, es decir, si dos loci se encuentran cercanos en el cromosoma, el número de gametos recombinantes con esos loci será muy bajo, pudiendo llegar a ser cero, en el caso de que los genes estén estrechamente ligados. Por el contrario si dos loci se encuentran en puntos muy distantes en el cromosoma, el número de gametos recombinantes con los loci será alto (Klug *et al.*, 2006).

Lo anterior es la base del mapeo genético, ya que permite identificar la posición relativa de cualquier tipo de marcador molecular o genes en los cromosomas.

Según Semagn *et al.*, (2006b) son cuatro los pasos necesarios para realizar un mapa genético:

1. Desarrollar una apropiada población de mapeo y determinar el tamaño de la población; Según Meksen y Kahl (2005) para hacer mapeo de QTL se requiere una población de 200 individuos.
2. Determinar el tipo de marcador molecular para genotipificar la población.
3. Detectar los polimorfismos en los parentales con el marcador seleccionado y luego genotipificar la población.
4. Realizar el análisis de ligamiento (calcular la frecuencia de recombinación por pares de marcadores, establecer los grupos de ligamiento, estimar las distancias de mapa y determinar el orden de los marcadores) usando programas estadísticos como MapMaker (Lander *et al.*, 1987)

2.9.1 Poblaciones de mapeo

Generalmente las poblaciones de mapeo se obtienen a partir de cruces controlados. Uno de los criterios más importantes en el momento de diseñar una población de mapeo es la selección de los parentales, esto es esencial para rastrear los eventos de recombinación, los parentales deben ser altamente contrastantes en las características de interés si el fin es realizar mapeo de QTL, si es mapeo convencional los parentales no necesitan ser contrastantes.

Otro criterio importante es determinar que tipo de cruzamiento se realizará, y esta decisión depende de varios aspectos, uno de ellos el modo de reproducción de la planta (autopolinización o polinización cruzada), el tiempo disponible para la construcción de la población y la resolución de mapeo requerida (Meksen y Kahl, 2005).

Cuando el modo reproductivo de la especie objeto de estudio es la autopolinización, las poblaciones de mapeo se originan a partir de parentales homocigóticos, de esta manera, la segunda generación filial (F_2), el retrocruzamiento (BC), las líneas endogámicas recombinantes (RILs), los dobles haploides (DH) y las líneas isogénicas cercanas (NILs), pueden ser usadas para generar las poblaciones de mapeo.

En las especies de polinización cruzada la situación es mucho más compleja debido a que éstas no toleran la endogamia, trayendo como consecuencia depresión endogámica. Para estas especies los pseudo cruces de prueba de dos vías, las familias que comparten un padre “half-sib” y los dos padres “full – sib” derivadas de cruces controlados, se han propuesto como poblaciones de mapeo (Burr *et al.*, 1988; He *et al.*, 2001; Doerge, 2002, citados por Semagn *et al.*, 2006b)

2.9.2 Población de mapeo F_2

Para obtener una población de mapeo F_2 Se seleccionan como parentales dos líneas puras, es decir completamente homocigóticos, diferentes en las características a estudiar; la progenie de este cruce se denomina primera generación filial (F_1); posteriormente son cruzados dos individuos de la F_1 para generar la segunda generación filial (F_2); la segregación esperada para esta población es de 1:2:1 (parental 1: heterocigoto: parental 2).

La gran desventaja con esta población es que no puede ser fácilmente preservada, ya que las plantas F_2 no son inmortales y la F_3 es el resultado del cruzamiento de dos individuos F_2 , por lo que genéticamente no son idénticas, pues ya habrían ocurrido dos eventos de recombinación.

Cuando se desarrolla la población para la caña de azúcar, por ser esta una especie con alto número de cromosomas, no es necesario llegar hasta la segunda

generación filial (F_2) dado que con la primera generación filial (F_1) se obtiene un gran número de eventos de recombinación, suficientes para obtener la información necesaria para el análisis de ligamiento.

2.9.3 Identificación de polimorfismo

Después de tener la población de mapeo, y haber seleccionado el marcador molecular, el tercer paso es identificar los marcadores polimórficos en los parentales y posteriormente evaluarlos en la población segregante.

Dependiendo de la población de mapeo, se esperan determinadas proporciones de segregación determinadas. Las desviaciones significativas de las proporciones esperadas, pueden ser analizadas por medio de una prueba Chi-cuadrado (χ^2).

La genotipificación en especies diploides es una tarea relativamente sencilla, sin embargo en las especies poliploides se presentan algunas dificultades. En primer lugar en los organismos poliploides se espera un gran número de alelos por cada marcador, siendo muy difícil determinar el genotipo de cada individuo (Wu *et al.*, 1992). Por otro lado el hecho de que la constitución genómica (auto o alopoliploide) de los poliploides sea poco clara, dificulta la determinación de los patrones de herencia. Para solucionar estos problemas Wu *et al.*, (1992), plantearon una aproximación, consiste en analizar la segregación de cada fragmento basados en la presencia y ausencia en la progenie, es decir, un

fragmento representado por una sola dosis (o un solo alelo) en plantas heterocigotas puede segregarse en una proporción 1:1 (presencia: ausencia); los fragmentos que están presentes en una sola dosis y que segregan en un radio 1:1 en los gametos de la planta se denominan fragmentos de una sola dosis (SD). Los marcadores de una sola dosis son detectados de acuerdo a dos criterios, primero, el fragmento debe estar presente en uno de los padres y ausente en el otro, y segundo, debe segregarse en una proporción 1:1, estos criterios son válidos si la especie es poliploide, aloploide o autopoliploide (Wu *et al.*, 1992).

Un marcador de una sola dosis es equivalente a un alelo simple en autopoliploides, o un alelo en un locus heterocigoto en un genoma diploide, en aloploides; la segregación de un SD es equivalente a la segregación de un alelo simple o heterocigoto la mitad de los gametos pueden tener el fragmento de ADN y la otra mitad no (Wu *et al.*, 1992).

2.9.4 Análisis de ligamiento

El paso final en la construcción de mapas de ligamiento es la codificación de los datos para cada marcador molecular en cada individuo; esta información es organizada en una matriz de datos, a partir de la cual se calculan las fracciones de recombinación para cada marcador dividiendo el número de individuos recombinantes sobre el total de individuos analizados; las fracciones de

recombinación calculadas entre dos marcadores indican la distancia de los mismos dentro del mapa de ligamiento.

Las fracciones de recombinación calculadas no son aditivas a lo largo del cromosoma, por ende las distancias en el mapa no pueden ser expresadas por fracciones de recombinación; para solucionar este problema se han desarrollado funciones de mapeo que hacen a las fracciones de recombinación aditivas, al convertirlas en centiMorgans (cM), unidad en la que finalmente se expresan las distancias genéticas. (Hui, 1998).

La función de mapeo de Kosambi y la función de mapeo de Haldane son las funciones más utilizadas en construcción de mapas, la primera asume que la ocurrencia de dobles recombinantes es mayor que la esperada, desviando de una manera significativa los valores en las fracciones de recombinación esperadas, a este fenómeno se le conoce como interferencia, (Hui, 1998); por el contrario la función de mapeo de Haldane no asume interferencia.

Después de establecer las distancias entre pares de marcadores se hace el análisis de ligamiento, que usualmente es calculado con el logaritmo de la relación (LOD score), los valores de LOD mayores de 3 generalmente son los valores significativos para determinar un ligamiento; un valor LOD de 3 entre dos marcadores indica que el ligamiento entre ellos es 1000 veces más probable que un no ligamiento.

Los análisis explicados anteriormente son realizados en software o programas específicos para mapeo genético, los programas comúnmente usados son Mapmaker/EXP (Lander *et al.*, 1987; Lincoln *et al.*, 1993 a), Map Manager QTX (Manly *et al.*, 2001) y JoinMap (Stam, 1993).

2.9.5 Mapeo genético en caña de azúcar

La combinación de un alto nivel de ploidía, aneuploidía, y la naturaleza altamente heterocigótica de la caña de azúcar, hacen de ella, uno de los cultivos más difíciles para el mapeo genómico, mostrando desafíos que comúnmente no se presentan en especies diploides (Kole, 2007; Piperidis *et al.*, 2008). Esto hace que el proceso de mapeo genético sea mucho más lento para este tipo de especies.

Sin embargo actualmente existen para caña de azúcar dos clases de mapas genéticos (Tabla 2), los de primera y segunda generación, los mapas de primera generación comprenden todos los mapas de las especies parentales *S. officinarum* y *S. spontaneum*. A su vez, los mapas de segunda generación son los desarrollados para los cultivares comerciales como el francés R570 y el australiano Q165 (Tabla 2).

Tabla 2. Mapas genéticos publicados hasta la fecha en orden cronológico, los asteriscos encontrados al final de los paréntesis indican el número total de marcadores evaluados, los números ubicados fuera del paréntesis son los marcadores ligados a grupos de ligamiento, grupos de homología (GH) o grupos de cosegregación (GC).

	Publicación	Especie y Clon	Tipo de marcador	Número y Dosis	Grupos de Ligamiento	Cubrimiento
PRIMERA GENERACIÓN	Aljanabi <i>et al.</i> , 1993	<i>S. spontaneum</i> SES 208	AP - PCR	208 SD	41	
	daSilva <i>et al.</i> , 1993	<i>S. spontaneum</i> SES 209	RFLP	319 SD	44	
	daSilva <i>et al.</i> , 1995	Unión mapas anteriores	AP - PCR, RFLP	527 SD,DD,TD	64	93%
	Ming <i>et al.</i> , 1998	<i>S. officinarum</i> [Green German (2n=97-117)]	RFLP	270(418)*	72	2,304 cM
		<i>S. officinarum</i> [Mutak Java (2n=140)]	RFLP	206 (355)*	72	1,443 cM
		<i>S. spontaneum</i> [IND 81-146 (2n=52-56)]	RFLP	248	69	2,063 cM
	<i>S. spontaneum</i> [PIN 84-1 (2n=96)]	RFLP	182	69	1103 cM	
Grivet <i>et al.</i> , 1996	R570	RFLP	600	98	2008 cM	
Hoarau <i>et al.</i> , 2001	R570	AFLP	887 (939)*	120		
SEGUNDA GENERACIÓN	Ming <i>et al.</i> , 2002	Mapeo comparativo con <i>Sorghum bicolor</i> (L) y mapa conseso de los mapas de Ming <i>et al.</i> , 1998	RFLP	982	13 GH	Cubrimiento de 39,5 a 46% en los 4 mapas, y 70% en el mapa conseso
	Rossi <i>et al.</i> , 2003	R570 saturación mapa Hoarau <i>et al.</i> , 2001	RGA y SSR	1169	128, 7 GH	menos del 50%
	Aitken <i>et al.</i> , 2005	variedad comercial (Q165; 2n=115) X IJ76-514 (2n=80)	AFLP y SSR	910 (967)*	116	9,053, 3 cM
	García <i>et al.</i> , 2006	variedades pre-comerciales SP80-180 X SP80-4966	RFLP, SSR y AFLP	357 (1,118)*	131 GC	2,602, 4 cM
	Oliveira <i>et al.</i> , 2007	variedades pre-comerciales SP80-180 X SP80-4966	EST - SSR	664 (1669)*	192 GC, 14 GH	6,261, 1 cM

2.9.6 Análisis de QTL

Muchos de los caracteres de la caña de azúcar importantes en la producción son de naturaleza continua, como producción de sacarosa, resistencia a enfermedades, entre otras, siendo difíciles de identificar dada la ausencia de una segregación fenotípica discreta observable, además los efectos genotípicos de cada gen asociado con un carácter complejo son relativamente pequeños, por esto en los procesos de mejoramiento se hace necesario disponer de métodos de análisis de caracteres cuantitativos que revelen el componente genético.

Un método de análisis de estos caracteres cuantitativos es el análisis de QTL (loci de características cuantitativas) este análisis implica: 1. la selección y cruzamiento de líneas parentales que difieren en uno o más caracteres cuantitativos y 2. El análisis de la segregación de la descendencia para relacionar cada QTL con un marcador de ADN conocido, o intervalo de marcadores.

Una de las aplicaciones directas del análisis de QTL es en la selección asistida por marcadores (MAS: Marker assisted selection), metodología que se utiliza en los programas de mejoramiento, con el objetivo de asistir al fitomejorador dando criterios indirectos que le aportan información para la selección.

Para implementar la selección asistida por marcadores se deben realizar una serie de análisis (Figura 5), que inician con la elaboración de mapas genéticos

utilizando marcadores moleculares que a su vez serán utilizados en el análisis de asociación con las características agronómicas que se desean evaluar, posteriormente se definen y caracterizan las asociaciones, las interacciones entre marcadores y el efecto de cada marcador en la característica de interés, todo esto con miras a definir con que tipo de combinaciones de alelos marcadores, y en qué condiciones se obtienen variedades con las características agronómicas deseadas.

El análisis de características cuantitativas o QTL, se basa en la detección de asociaciones entre el fenotipo y el genotipo; en este caso, los marcadores moleculares se utilizan para dividir la población segregante con base en la presencia o la ausencia de un locus marcador particular y determinar si las diferencias entre los locus son significativas.

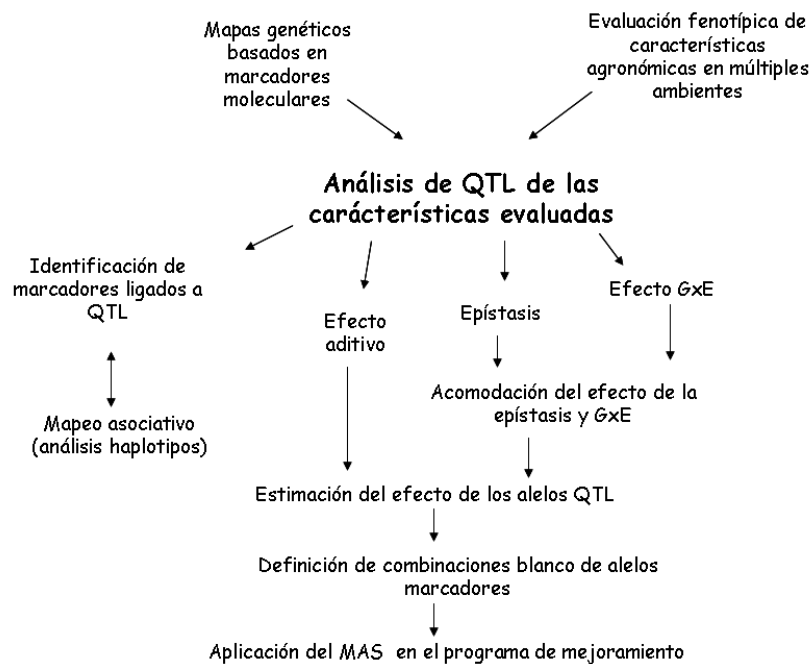


Figura 5. Pasos en la identificación y caracterización de loci de características cuantitativa (QTL) para su uso en los programas de mejoramiento. Tomado de: Semagn et al., (2006 b).

Cuando se presentan diferencias significativas entre medias se puede decir que el marcador está ligado o asociado a una característica cuantitativa (QTL).

Lo anterior se debe a las diferencias en las frecuencias de recombinación, la frecuencia hace que la media poblacional del marcador ligado varíe con respecto al no ligado, el marcador que no está ligado al QTL segrega independientemente, por lo tanto no se esperaría encontrar diferencias significativas en las medias de los grupos genotipados (Collard *et al.*, 2005).

Existen cuatro métodos principales para detectar QTL:

- Análisis de un solo marcador.
- Mapeo por intervalo simple.
- Mapeo por intervalo compuesto
- Mapeo por intervalo múltiple.

2.9.6.1 Análisis de un sólo marcador

Éste análisis se emplea para identificar QTL asociados a un solo marcador, para este análisis no es necesario un mapa genético previo y se puede realizar con métodos estadísticos como: test t, ANOVA y regresión lineal en software básico de estadística (Collard *et al.*, 2005).

2.9.6.2 Mapeo por intervalo simple

Se utilizan mapas de ligamiento e intervalos entre pares adyacentes de marcadores ligados a lo largo de un cromosoma simultáneamente. Se considera como más poderoso estadísticamente que el análisis de un solo marcador; MapMaker QTL y QGene son los programas comúnmente usados para realizar estos análisis (Collard *et al.*, 2005).

2.9.6.3 Mapeo por intervalo compuesto

Éste es el método más usado, combina intervalos de mapeo, regresión lineal y marcadores genéticos haciéndolo más preciso y efectivo comparado con los otros métodos de análisis; los programas comúnmente utilizados son QTL cartographer, MapManager QTX y PLAB QTL.

2.9.6.4 Mapeo por intervalo múltiple

El intervalo de mapeo múltiple detecta todos los QTL que pueden estar asociados a una característica, bajo la restricción de que son múltiples los QTL implicados y que estos son independientes, además asume que no hay ligamiento ni interacción genética entre ellos, como se observa este método es útil cuando la característica de interés que esté controlada por varios genes (Hui. 1998).

3. OBJETIVOS

3.1 General

Realizar una aproximación de mapa genético para las variedades colombianas de caña de azúcar, usando marcadores microsatélites en el cruzamiento MZC 74-275 x ICA 69-11.

3.2 Específicos

- Identificar grupos de ligamiento en variedades colombianas de caña de azúcar usando marcadores microsatélites
- Identificar posibles asociaciones o QTL ligados a componentes de producción de biomasa y sacarosa.

5. METODOLOGÍA

El presente estudio fué desarrollado en la estación experimental del Centro de Investigación de la Caña de Azúcar de Colombia (CENICAÑA) ubicado en Florida, Valle del Cauca, dentro del grupo de investigación Macroproyecto alta sacarosa estable, reconocido por COLCIENCIAS como categoría A.

4.1 Material Experimental

4.1.1 Población de mapeo

Los primeros pasos para el desarrollo de este trabajo fueron la selección de la población de mapeo y su siembra en campo. La población de mapeo se originó del cruzamiento entre MZC 74 - 275 e ICA 69 - 11, estas variedades son contrastantes en sacarosa la primera generación filial constituida por 300 individuos, fue sembrada en campo con dos repeticiones de cada individuo en Cachimbalito, Hacienda ubicada en el Departamento del Cauca (Anexo A), presenta una precipitación media anual de 1520 mm, el suelo se encuentra clasificado como AZ5H5 dentro de la zonificación agroecológica realizada por el centro de investigación de la caña de azúcar de Colombia (CENICAÑA), es un suelo con textura franca gruesa, superficial a muy superficial

4.1.2 Características analizadas

A cada uno de los individuos de la progenie y a los padres se les evaluó, la producción de clorofila, el diámetro y la altura a los 4, 12 y 13 meses de edad, estas características se evaluaron en estas edades del cultivo por que a los 4 meses se puede predecir que tanta biomasa y azúcar puede tener el cultivo, en esta edad la caña está afirmada en el suelo y comienza su verdadero desarrollo, acumulando biomasa y azúcar; a los 12 y 13 meses de edad del cultivo se presenta la máxima producción y acumulación de sacarosa, esta es la edad de corte con la que trabajan los ingenios azucareros.

Los datos obtenidos se registraron con tres repeticiones, estos valores fueron promediados; la cantidad de clorofila fue determinada con un clorofilómetro (Minolta SPAD - 502), el diámetro con un calibrador o pie de rey y la altura con una regla metálica de dos metros; después de la cosecha de la población a los 13 meses se tomaron al azar tres tallos de cada parcela, estos fueron cortados hasta obtener 12 trozos, para un total de 36 trozos por parcela, los tallos fueron llevadas al laboratorio de química de CENICAÑA donde, para cada uno de los individuos con su repetición se determinó, los porcentajes de sacarosa, pureza, fibra, azúcares reductores totales, no sacarosas, humedad, materia seca, brix, ARE (Azúcares Reductores Estimados) y ART (Azúcares Reductores Totales) todas

estas características son determinantes en la producción y calidad del azúcar generado por la planta (Tabla 3).

Tabla 3. Se relacionan los componentes de la producción de sacarosa y biomasa que fueron analizados en el presente estudio y la edad de cultivo en la que fueron evaluadas.

características analizadas	Componente producción		Mes toma de datos		
	Sacarosa	Biomasa	4	12	13
Sacarosa	x			x	x
Pureza	x			x	x
Fibra	x			x	x
Azúcares reductores	x			x	x
Azúcares reductores totales (ART)	x			x	x
Azúcares reductores estimados (ARE)	x			x	x
No sacarosas	x			x	x
Humedad	x			x	x
Materia seca	x			x	x
Brix	x			x	x
Altura		x	x	x	x
Diámetro		x	x	x	x
Clorofila		x	x	x	x

4.1.3 Marcadores moleculares y detección

Para realizar el análisis molecular de la población de mapeo, se extrajo el ADN de los 300 individuos empleando el protocolo de extracción de Gilbertson *et al.*, (1991) modificado (Anexo B). Posteriormente, se evaluó la calidad del ADN y se cuantificó en un gel de agarosa al 0.8% para establecer su concentración, después el ADN de cada individuo fue diluido con agua ultra pura y desionizada a una concentración de 4ng/ul para ser utilizado en la amplificación de los SSR.

Para facilitar la lectura y el registro de los datos estos fueron organizados de forma tal que los padres (MZC 74-275, ICA 69-11) quedaran ubicados en cada

uno de los extremos del gel, dando espacio para el ADN de 60 individuos más (se usaron peines de 68 pozos), así se formaron cinco grupos (llamados peines), cada uno con el ADN de 60 individuos más los parentales (figura 6), finalmente el ADN de los 300 individuos de la población fue almacenado en una placa para PCR de 384 pozos a -20°C .

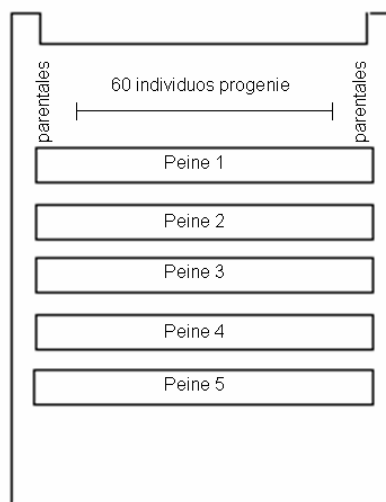


Figura 6. Esquema de la organización de los ADN en el gel de poliacrilamida, con cinco siembras. Se muestran 5 peines que en conjunto forman la población de mapeo (300 individuos), en los extremos se ubica cada uno de los parentales y en el medio de ellos se encuentran 60 individuos de la progenie, esta organización se tiene para cada uno de los peines que se visualiza en el esquema.

Posterior al proceso de extracción, cuantificación y evaluación de la calidad del ADN, se inició la estandarización de las condiciones de amplificación de los microsatélites (SSR), que básicamente consistió en establecer, por medio de ensayos de amplificación, las condiciones precisas para que el SSR pudiera ser

amplificado y visualizado con una buena resolución, nitidez y reproducibilidad de las bandas. Para lograr esto, se amplificaron todos los SSR con las siguientes condiciones de amplificación:

Reactivos	Concentración inicial	concentración final	Volumen de cada reactivo para una reacción (ul)
Buffer	10X	1X	1,5
MgCl₂	25mM	2,0mM	1,2
dNTP's	8mM	0,192mM	0,36
Iniciadores	10mM	0,67mM	1
Taq polimerasa	5U	0,01U	0,03
ADN	4ng/ul	1,3ng/ul	5
Agua	-	-	5,91
volumen de la reacción	-	-	15

y el programa de amplificación siguiente:

94°C por 3 minutos

94°C por 30 segundos	} 35 Ciclos
55°C por 30 segundos	
72°C por 45 segundos	

72°C por 5 minutos

10°C ∞

A los SSR que después de amplificados con estas condiciones no estuvieran estandarizados, se les ajustaron sus condiciones cambiando la concentración del cloruro de magnesio y de temperatura de anillamiento hasta que estuviese

completamente ajustado. Las amplificaciones se realizaron en un termociclador BioRad C1000 Thermal Cycler™ y un JM research PT100.

Con las condiciones de PCR establecidas, se evaluaron los SSR en los parentales, con el fin de encontrar polimorfismo entre ellos; un total de 23 marcadores microsatélite fueron seleccionados (Tabla 4), posteriormente estos se evaluaron en la población de mapeo.

Tabla 4. Marcadores microsatélites evaluados, se muestra la temperatura de anillamiento y la concentración de MgCl₂ con la que se amplificaron.

Microsatélites	Secuencia iniciador 5'-3'	Secuencia iniciador 3'-5'	Temperatura (°C)	MgCl ₂
CG319	CCT TTC ATC CAC AGA GGA CAG	GGT TCA CCG AAG CAA GAG AAC	55	2,0
CIR11	CCA CCA TCT TTT CGC ACC AG	GCA GCA CCA ACC ATA ATC AT	55	2,0
CIR48	CGG TCG TGA GAA GGA AAG	CGC ACC AGA ACA CGT ATA	60	2,0
CIR36	CAA CAA TAA CTT AAC TGG TA	CTG TCC TTT TTA TTC TCT TT	55	3,0
CL1557	ACC AGC GAT GGG TGC TAA	GCC TCT TCC AGT TCC AGC T	55	2,0
CIR31	ATT TGG GTA AGG ATG GAT	CCT AAT GAT ACG CTT TGA	55	2,0
CV37	GGATGGACGACGTGTCCTGG	ATAAAGTGGCCGCTTGGATTGA	62	2,0
CS597	GCA CAC CAC TCG AAT AAC GGA T	AGT ATA TCG TCC CTG GCA TTC A	60	2,0
CS668	ACG CTT GCG TGC TCC ATT	CCA ATC GTG CCA CTG TAG TAA G	62	2,0
DUQ08	TCA TGG ATG CTT TTG CGT	ACA GAA CAA CGC CGC TGT	53	2,5
CS67	TGC GGA TGA TGA GAT CAG CA	TGG ACA CGA TCA AAC AGC AGT G	55	2,0
CG236	ACC GAC CTC AAA GAA GAC GGA T	CAG TAT GAG TGA AGC ACG GGA AC	55	2,0
CIR52	ACA AGG GAA GAC GAC AAA TCA G	ACC AAA CCA CCA AGC AAA	55	2,5
BS334	CAA TTC TGA CCG TGC AAA GAT	CGA TGC GCT TGA TTG CGA ATG	55	2,0
CIR10	ACA CCA CTC ACA TCC ACT TG	TGA TAC ACC ATT GTT GAT GC	55	2,5
CIR19	GGT TCC AAA ATA CAC AAA	CAA TCT TAT CTA CGC ACT T	55	2,3
CIR18	GGG TGT TCT GTT GAG CA	GAG GTA GGA GGG AGT GTT	55	1,5
CIR40	CCG TAT TCT TCT TTT CTT CA	TTC AGG CGA GTC ATC ATT	55	2,0
CIR24	AGA TGA ACC CAA AAA CTT A	TTA CTC CGC CTC TTT ACT	55	2,0
CIR32	CAC AAA CAT CCC TCA ACT	ATT CTT GCA TTA AAC CTC TC	55	1,8
CIR47	GCA ATG GAG GTA GGA ATG	TAG AAT CAC CCA AAA ATA AA	53	2,0
CIR15	CTT GGA CCC GTT CTT GGA TG	AGC ACT GAG GCG ACT TAC CC	54	1,7

Los productos de PCR fueron separados en geles de poliacrilamida denaturantes al 5%, con una proporción acrilamida/bisacrilamida de 19:1, úrea 7M. Para visualizar los fragmentos separados los geles se tiñeron con nitrato de plata (AgNO₃) (Anexo C); en los geles de progenie se determinó la presencia o

ausencia de las bandas polimórficas, registrándose 1 o 0 en una matriz de datos respectivamente.

4.1.4 Análisis de ligamiento

En el análisis de los datos sólo se incluyeron los marcadores que segregaron en una proporción 1:1 y 3:1, para determinar que marcadores cumplían con esta característica se realizó una prueba chi - cuadrado (χ^2).

Con la información obtenida a partir de los marcadores seleccionados se construyó una matriz binaria con los marcadores ubicados en las columnas y los individuos en las filas, como lo requiere el programa MAPMAKER v2.0 para Macintosh (Lander *et al.*,1987), en el que se realizó el análisis de ligamiento colocando como parámetros de identificación un LOD 3 y un porcentaje de fracción de recombinación (θ) de 4, estos parámetros son los parámetros estadísticos, el valor de LOD 3 es la probabilidad umbral para establecer un ligamiento, este parámetro indica que el ligamiento es 1.000 veces más probable que el no ligamiento, por otro lado, el valor de θ permite estimar el orden de los marcadores que se encontraron ligados en términos de la fracción de recombinación, es decir sólo los marcadores que presenten un 4% del máximo de la fracción de recombinación que se espera (50%) serán los que se incluirán

dentro del mapa, ya que este valor indica que los marcadores se encuentran cercanos dado que sólo el 4% de los gametos fue recombinante.

4.1.5 Análisis de características cuantitativas QTL

Posterior al análisis de ligamiento se realizó el análisis de un solo marcador, con un análisis de varianza (ANOVA) donde cada característica es explicada por el factor de alelos marcadores, con el fin de encontrar asociaciones entre los marcadores incluidos en el análisis de ligamiento y las 13 características evaluadas de componentes de producción de sacarosa y biomasa. Este procesamiento se realizó con el programa SAS versión 9.1.3.

Dado que son 45 alelos incluidos dentro del análisis de QTL, la probabilidad de que cada uno de éstos marcadores se encuentre asociado a una característica no puede ser la misma que se tiene para la asociación de un solo alelo a una característica $p \leq 0.05$; con el objetivo de ser más exigentes en términos del nivel de significancia estadístico de la asociación, se utilizó la corrección de Bonferroni para el valor de significancia, que consiste en dividir el valor de significancia para la asociación de un sólo alelo sobre el número de alelos evaluados ($0.05/45 = 0.001$) de esta manera sólo las asociaciones que tengan un $p \leq 0.001$ serán considerados como significativas estadísticamente.

Para analizar el comportamiento de todos los alelos marcadores en conjunto y definir las combinaciones que explican variaciones fenotípicas en la población, se

realizó un análisis de regresión logística, en el cual la variable respuesta es cada una de las características y las variables explicativas son los alelos marcadores. Ésta técnica también permitió evaluar el aporte de cada marcador a la característica a través de la razón de probabilidad (odds ratio).

La variable respuesta fue discretizada con base en los valores del tercer cuartil para cada una de las características; los datos cuyos valores sean superiores al tercer cuartil fueron codificados como 1, que significa un valor alto de la variable y los inferiores como 0, que significa un valor bajo de la variable. Este análisis fue realizado con el programas Statgraphics Centurión.

5. RESULTADOS Y DISCUSIÓN

Un total de veintidós marcadores (45 alelos) fueron evaluados en la población de mapeo, a los cuales se les realizó el análisis de ligamiento usando MAPMAKER 2.0 para Macintosh (Lander *et al.*, 1987), con un LOD de 3 y θ 4; en el análisis se encontró un grupo de ligamiento con longitud total de 176.791 cM y 7 alelos ligados de los 45 analizados (figura 7).

En los mapas genéticos construidos en Australia con SSR y AFLP por Aitken *et al.*, (2005) y Piperidis *et al.*, (2008) y el mapa construido en Brasil con SSR, AFLP y RFLP por García *et al.*, (2006), se reporta la ubicación de 11 de los 22 marcadores evaluados en este estudio (Tabla 5).

Tabla 5 Marcadores utilizados en este estudio y su ubicación en previos reportes. GH= Grupo de homología, GL= Grupo de ligamiento, GC=Grupo de cosegregación.

Publicación	Marcador	GH	GL	GC
Aitken <i>et al.</i> , 2005	CIR10	3	23,29	
	CIR15	6	82,71	
	CIR19	4	78,12b	
	CIR24	3	67	
	CIR32	3	93,28,63	
	CIR36	1	17,42,9,19,114,27	
	CIR52	3	83,21	
	CL1608	3	65,24,108	
García <i>et al.</i> , 2006	CIR36	1	I-3;I-5;I-4;I-1	
	CIR32	11	XI-1;XI-2	
	CIR18			13

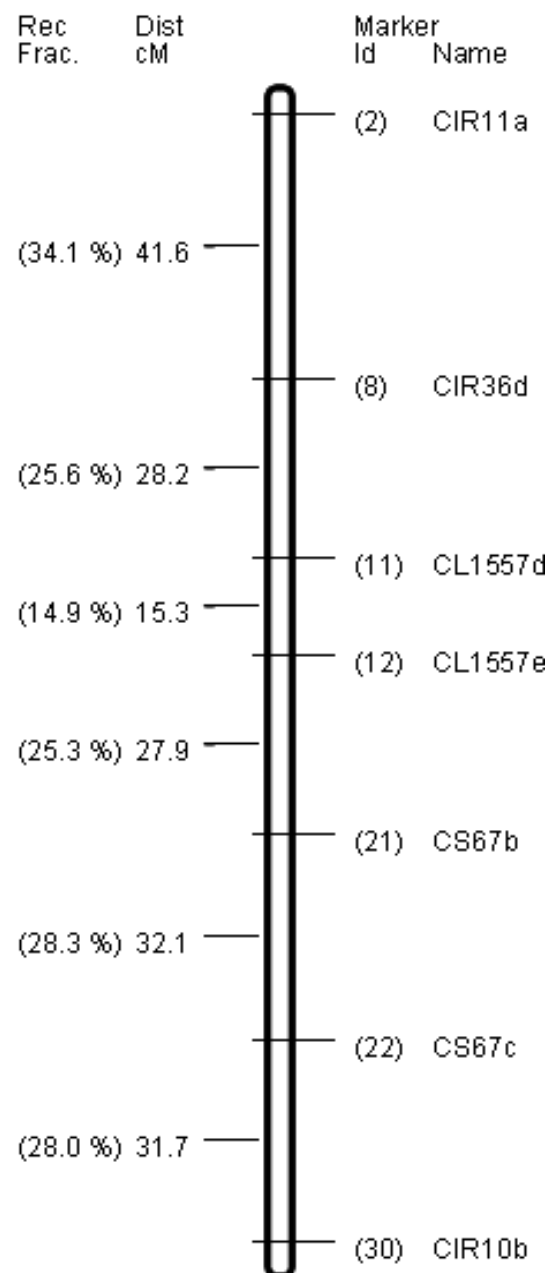


Figura 1. Grupo de ligamiento formado con un LOD 3 y fracción de recombinación 4.

Una de las causas del bajo número de marcadores usados en este estudio (22), es la complejidad y especificidad requerida para la estandarización de los microsatélites, lo cual dificulta en gran medida la reacción de amplificación (PCR). Dado que el tejido de la caña de azúcar contiene altos niveles de polisacáridos y compuestos polifenólicos, estos en su forma oxidada se enlazan covalentemente al ADN dándole una coloración café, creando así un problema de contaminación que limita su uso en gran medida, cuando dicho ADN es utilizado para PCR (Aljanabi *et al.*, 1999).

La caña de azúcar por su tamaño genómico y sus altos niveles de ploidía y aneuploidía, es catalogada como la planta cultivada con estructura genética más compleja y difícil para el mapeo genómico entre los cultivos de importancia agronómica, especialmente dentro de los cereales (D'Hont *et al.*, 1996; Piperidis *et al.*, 2008).

Es así como se estima que en los mapas genéticos de Francia y Australia catalogados como los más grandes hasta el momento, que fueron construidos con marcadores SSR y AFLP, se cubrió menos del 50% del genoma a pesar del gran número de marcadores mapeados en las dos variedades (R570 y Q165) (Piperidis *et al.*, 2008).

El mapa genético de caña construido en Francia por Hoarau *et al.*, (2001) en la variedad R570, contiene 1169 marcadores distribuidos en 128 grupos de

ligamiento que fueron asignados a 7 de los 8 grupos de homología esperados; el mapa del cultivar Australiano construido en la variedad Q165 por Aitken *et al.*, (2005), contiene 968 marcadores SSR y AFLP distribuidos en 136 grupos de ligamiento que fueron asignados a 8 grupos de homología.

Dada la alta complejidad genómica de la caña de azúcar, el entendimiento, así como el conocimiento del mosaico genómico y genético, la biología molecular y citogenética de las variedades de caña de azúcar, es necesario para facilitar y complementar el desarrollo y futura aplicación de los mapas de ligamiento en la selección asistida por marcadores dentro de los programas de mejoramiento.

Este tipo de información no se conoce para las variedades de caña de azúcar colombianas pero si para la variedad francesa R570, sin embargo por el hecho de que la caña de azúcar es el producto de un proceso de nobilización, subsecuente selección y posteriores cruzamientos interespecíficos independientes en las diferentes partes donde se desarrollaron, hace que las variedades comerciales alrededor del mundo sean diferentes entre si en términos genómicos, por lo tanto la información genómica y citogenética ya obtenida para variedades comerciales, no debe ser extrapolada completamente a las variedades colombianas.

Una de las razones que sustenta lo anterior, es que la especie *S. spontaneum*, tiene un amplio rango en el número de cromosomas con cinco citotipos

principales $2n=64$, $2n= 80$, $2n= 96$, $2n=112$ y $2n=128$, que fueron incluidos en los cruzamientos interespecíficos con *S. officinarum* en el proceso de nobilización.

Hower Roach (1969, citado por D'Hont *et al.*, 1996) en cruzamientos de *S.officinarum* con diferentes citotipos de *S. spontaneum* ($2n=80$, $2n= 96$ y $2n=112$) reportó que la transmisión de cromosomas durante la segregación meiótica es $n + n$, es decir cada uno de los parentales dona su número haploide de cromosomas. También observó que en los cruzamientos con el citotipo de *S. spontaneum* con $2n= 64$ la transmisión es exclusivamente $2n+n$, sin embargo D'Hont *et al.*, (1996) reportaron que la transmisión $n+n$ también puede ocurrir en los cruzamientos con este último citotipo.

Si la transmisión es $n+n$ se entiende que cada uno de los parentales aportó una copia de su genoma a la progenie, y si la transmisión es $2n+n$, se infiere que uno de los padres aportó su genoma completo, mientras que el otro solo aportó la mitad, esto se verá reflejado en la composición genómica de las variedades comerciales; si la segregación, como se mostró anteriormente, cambia con el citotipo de *S. spontaneum* utilizado en el cruzamiento, es posible que las proporciones de *S. officinarum* (70 - 80%) y *S. spontaneum* (10 - 20%) reportadas por D'Hont *et al.*, (1996) para la variedad Francesa R570 no sean las mismas para las variedades colombianas, ya que para estas variedades, como ya se mencionó, aún se desconoce, su número exacto de cromosomas, las

proporciones de cada parental dentro del genoma, la forma de transmisión y de segregación de cromosomas.

El entender la forma de transmisión es informativo en la elaboración de mapas genéticos, cuya base está en el proceso meiótico de segregación, el conocer de que forma se da la segregación en la variedad, da herramientas para estimar qué se espera en términos de fracciones de recombinación y grupos de ligamiento, por otro lado si una variedad tiene una proporción mayor de un parental que de otro, es posible inferir que esta tenga en gran medida los atributos de ese parental, situación que puede dar luces en el proceso de mejoramiento y del mapeo mismo.

El número de cromosomas según Hoarau *et al.*, (2001) permite estimar el tamaño genómico de la especie, al multiplicar el número de cromosomas total de la variedad por la longitud media de los grupos de co – segregación (marcadores ligados en un grupo de ligamiento en fase de acoplamiento) más largos, quienes probablemente se encuentran cerca de la saturación. Por ejemplo, para la variedad R570, el tamaño promedio de los grupos de co – segregación más grandes que se encontraron distribuidos en tres grupos de ligamiento tuvieron un tamaño aproximado de 150 cM, este valor fué multiplicado por el número de cromosomas de la variedad ($122 \times 150 = 17.000 \text{ cM}$ aproximadamente), así se tiene que el tamaño del genoma de esa variedad es aproximadamente 17.000 cM, esta

información permite hacer estimativos a cerca de que tanto se está aportando en términos de cubrimiento del genoma, con los mapas que se desarrollen.

Si bien el entendimiento y conocimiento del genoma, la genética y la citogenética de las variedades de caña de azúcar colombianas son factores importantes dentro de los procesos de mejoramiento, no son una limitante que impida continuar con el desarrollo y construcción del mapa genético de caña de azúcar en Colombia.

El análisis de QTL se realizó con el método de un solo marcador. Para este análisis no es necesario contar con un mapa genético previo, se hace análisis de asociación de cada marcador con cada característica fenotípica evaluada en campo; los marcadores que explican la variación fenotípica observada, serán aquellos que estén asociados de una manera estadísticamente significativa a la característica evaluada, estos podrán ser reconocidos como posibles QTL.

Con el análisis de un sólo marcador se encontraron cuatro marcadores con asociaciones estadísticamente significativas a altura a los 4 meses (CIR18a; $p=0.0016$), porcentaje de Fibra a los 12 meses (CIR15a, $p=0.0016$) y Clorofila a los 4 meses (CIR32b; $p=0.0018$ y CIR1557d; $p= 0.0001$) (Tabla 7).

Tabla 7. Asociaciones estadísticamente significativas para cuatro marcadores, se muestra el marcador, la característica a la que se asoció la edad del cultivo y el valor de significancia estadística (p).

Marcador	Característica	Edad (meses)	<i>p</i>
CIR 18a	Altura	4	0,0016
CIR 15a	% Fibra	12	0,0016
CIR 32	Clorofila	4	0,0001
CL1557			

En general los marcadores se asociaron a más de una característica en diferentes edades del cultivo (Tabla 6).

La asociación de un marcador molecular a varias características está de acuerdo con lo reportado por Aitken *et al.*, (2008) donde encontraron que la mayoría de los alelos de QTL asociados afectaron a más de una característica; Ming *et al.*, (2002) también reportaron QTL controlando la producción de azúcar, el peso del tallo y Pol; igualmente Aitken *et al.*, (2006) encontraron que cuatro de los marcadores analizados estuvieron asociados a las tres características evaluadas en su estudio (Brix, estimado del total de sólidos solubles en el jugo de caña, Pol, índice de polarización del primer jugo, es una medida del contenido de sacarosa y CCS, es un estimado de la sacarosa comercialmente extraíble y es determinada a partir de una función estándar de Brix, Pol y fibra).

Dentro del análisis de QTL sobresalieron marcadores que se asociaron a diferentes características a la misma edad del cultivo. Por ejemplo CIR 18a que además de asociarse significativamente a la variable altura a los 4 meses ($P \leq 0.001$) también se asoció con siete características a los 13 meses de edad del cultivo (Tabla 8), % sacarosa ($p= 0.0372$), % brix caña ($p= 0.0459$); % humedad ($p= 0.0210$); % materia seca ($p= 0.0210$), % ARE ($p= 0.0432$); % ART($p= 0.0319$) y % Pureza ($p= 0.0417$). Así mismo el marcador CIR15a se asoció a clorofila a los 4 meses ($p=0.0323$), % humedad ($p=0.0198$), %materia seca ($p=0.0198$) y el marcador CIR32d se asoció a los 12 meses a % humedad ($p=0.0037$), % materia seca ($p=0.0037$) y % fibra ($p=0.0045$), (Tabla 8).

Tabla 8. Marcadores asociados a diferentes características a las mismas edades del cultivo. Se relaciona el marcador asociado, las características, la edad y el valor de significancia estadística (p).

Marcador	Característica	Edad (meses)	p
CIR 18a	% Sacarosa	13	0,0372
	% Brix		0,0459
	% Humedad		0,021
	% Materia Seca		0,021
	% ARE		0,0432
	% ART		0,0319
	% Pureza		0,0417
CIR 15	Clorofila	4	0,0323
	% Humedad		0,0198
CIR 32c	%Humedad	12	0,0037
	% Materia Seca		0,0037
	% Fibra		0,0045

Estos marcadores llaman la atención porque, a pesar de que las asociaciones encontradas no son estadísticamente significativas, se está observando un

posible patrón, es decir, un mismo QTL controlando los diferentes componentes de producción de sacarosa en un tiempo determinado de crecimiento del cultivo.

El encontrar QTL que controlan los diferentes componentes de producción de sacarosa en un tiempo determinado de crecimiento, permite entender la relación genotipo – tiempo que existe en las variedades que se encuentran dentro del proceso de mejoramiento. Esto permitiría a futuro desarrollar cultivares específicos para tiempos particulares de cosecha (Aitken *et al.*, 2006). de esta manera sería posible mejorar variedades que en sus primeros meses de cultivo tengan una producción de sacarosa igual o mayor a la que se encontraría a los 14 meses (fecha tradicional de cosecha), reduciéndose significativamente el tiempo de cosecha.

Piperidis *et al.*, (2008) reportaron tres marcadores asociados con brix, CIR36 ($p \leq 0.005$) a los 9 meses, CIR 10 ($p \leq 0.005$) a los 12 meses y CIR18 ($p \leq 0.01$) a los 12 meses de edad del cultivo. En el presente estudio se confirma la asociación a brix de los marcadores CIR 18 ($p \leq 0.05$) a los 13 meses y CIR 36 ($p \leq 0.01$) a los 12 meses, esto indica que las asociaciones encontradas en los dos trabajos para estos marcadores son estables en diferentes variedades, pero es posible que la expresión del gen o genes que controlan la característica sea específica del sitio geográfico en el que se encuentra la variedad, dado la diferencia en el nivel de significancia de las asociaciones reportadas por los dos estudios (CIR36

$p \leq 0.005$ y $p \leq 0.01$; CIR18 $p \leq 0.01$ y $p \leq 0.05$ para el trabajo de Piperidis *et al.*, 2008 y el presente estudio; respectivamente).

Una posible explicación para las asociaciones poco significativas es, que debido a que las características cuantitativas pueden ser el producto de la acción aditiva de la expresión de muchos genes, es posible que la expresión del gen al que está asociado el marcador tenga un pequeño efecto en la característica cuantitativa evaluada, por ende no explicaría completamente la variación fenotípica de la característica, lo que se vería reflejado en una asociación no significativa estadísticamente. Para comprobar esto se realizó un análisis de regresión logística, que permite detectar el número y el efecto de los marcadores asociados a una característica particular.

Todos los marcadores identificados con el análisis univariado fueron incluidos dentro de un análisis de regresión logística para cada una de las características con las que se encontraron asociados, éste análisis se realizó con el objetivo de comprobar dicha asociación, establecer cómo es el comportamiento de los marcadores en conjunto, determinar cuál es su aporte y efecto (positivo o negativo) a la variación de la característica en la población, y si su explicación de la variación de la característica cambia de acuerdo a la interacción con otros marcadores.

Inicialmente se realizó el análisis de regresión logística para las características a las cuales se les identificó marcadores asociados significativamente con el ANOVA, estas características fueron, altura a los 4 meses, fibra a los 12 meses y clorofila a los 4 meses.

En el análisis se incluyeron los marcadores asociados estadísticamente significativos y los demás que se encontraron asociados a la misma característica pero no estadísticamente significativos, posterior a este análisis con los marcadores asociados estadísticamente significativos identificados por el ANOVA se realizó la regresión logística para las mismas características incluyendo todos los marcadores evaluados en este estudio, obteniéndose los siguientes resultados:

Altura a los 4 meses

El análisis de regresión logística, para los marcadores identificados como asociados a la altura a los 4 meses con el ANOVA, muestra que en conjunto los marcadores explican el 4,32% de variación en la característica, sin embargo, el modelo tiene un valor de p de 0,128, lo que cuestiona su significancia estadística (Tabla 9).

Tabla 9. Se muestran los resultados del análisis de regresión logística incluyendo solo los marcadores asociados identificados con el ANOVA, para altura a los 4 meses.

Modelo estimado de regresión (Máxima probabilidad)

		<i>Error</i>	<i>Estimado</i>
<i>Parámetro</i>	<i>Estimado</i>	<i>Estandar</i>	<i>Razón de la probabilidad</i>
CIR36c=0	-0,550372	0,397973	0,576735
CIR47a=0	-0,912422	0,320147	0,40155
BS334b=0	-0,743485	0,322755	0,475454
CIR18a=0	0,161828	0,365881	1,17566

Análisis de variación

<i>Source</i>	<i>Variación</i>	<i>Df</i>	<i>p-Valor</i>
Modelo	7,13449	4	0,1289

Porcentaje de varianza explicada por el modelo= 4,32485

Cuando se incluyeron todos los alelos evaluados en este estudio (45) en el análisis de regresión logística, se encontraron marcadores diferentes a los identificados con el ANOVA, que en conjunto explican 22,4% de la variación de la altura a los 4 meses con un valor de p del modelo de 0,0087, valor que soporta su significancia estadística (Tabla 10).

Los marcadores asociados a altura a los 4 meses, fueron CIR10a, CIR10b con efecto positivo, CIR31c y CIR 31f, con efecto negativo; entre estos se destaca el CIR10b, que según la regresión logística, cuando está presente en el individuo aumenta la probabilidad en 16.161 veces de que la altura incremente (Tabla 10).

Tabla 10. Se muestran los resultados del análisis de regresión logística incluyendo todos los alelos marcadores evaluados en éste estudio (45), para altura a los 4 meses

Modelo estimado de regresión (Máxima probabilidad)

Parametro	Error		Estimado
	Estimado	Estandar	Razón de probabilidad
CIR10a=0	1,37337	0,710619	3,94862
CIR10b=0	2,7826	1,43082	16,161
CIR31c=0	-1,55516	0,774971	0,211155
CIR31f=0	-1,39365	0,664988	0,248169

Análisis de variación

Source	Variación	Df	p-Valor
Modelo	13,6053	4	0,0087

Porcentaje de variación explicado por el modelo = 22,4617

Porcentaje de Fibra a los 12 meses

Con el análisis de regresión logística incluyendo los marcadores identificados con el ANOVA, se observa que los marcadores incluidos no explican la variación de la fibra a los 12 meses; además el modelo tiene un valor de p de 1,0 lo que indica que no es estadísticamente significativo (Tabla 11).

Tabla 11. Se muestran los resultados del análisis de regresión logística incluyendo solo los alelos marcadores identificados con ANOVA, para porcentaje de fibra a los 12 meses.

Modelo estimado de regresión (Máxima probabilidad)

Parametro	Error		Estimado
	Estimado	Estandar	Razón de probabilidad
CG236f=0	-0,944208	0,295495	0,388987
CIR15a=0	-0,114544	0,304602	0,891773
CIR32d=0	0,486658	0,374675	1,62687
CIR47b=0	-0,110294	0,365686	0,895571

Análisis de Variación

Source	Variación	Df	p-Valor
Modelo	-15,1384	4	1,0000

Porcentaje de variación explicado por el modelo = -8,54161

Cuando se incluyen en el análisis los 45 alelos evaluados en este estudio, se encuentran cinco asociados al valor de la fibra a los 12 meses: CIR11b, CIR36e, CIR52a, CIR40a, CIR32a explicando en conjunto 34,20% de la variación de la característica en la población (Tabla 12).

Los marcadores CIR32a y CIR40a tienen un efecto positivo en la característica, cuando el CIR32a está presente en el individuo aumenta en 27,842 veces la probabilidad de que se incremente el porcentaje de fibra a los 12 meses; cuando el marcador CIR40a se presenta, aumenta en 15,53 veces la probabilidad de que se incremente el porcentaje de fibra a los 12 meses (Tabla 12).

Tabla 12. Se muestran los resultados del análisis de regresión logística incluyendo todos los alelos marcadores evaluados en éste estudio (45), para porcentaje de fibra a los 12 meses.

Modelo estimado de regresión (Máxima probabilidad)

		<i>Error</i>	<i>Estimado</i>
<i>Parametro</i>	<i>Estimado</i>	<i>Estandar</i>	<i>Razón de probabilidad</i>
CIR11b=0	-3,38125	1,33234	0,0340049
CIR36e=0	-3,06115	1,25849	0,0468339
CIR52a=0	-2,89476	1,19544	0,0553126
CIR40a=0	2,74311	1,18686	15,5352
CIR32a=0	3,32654	1,33237	27,842

Análisis de variación

<i>Source</i>	<i>Variación</i>	<i>Df</i>	<i>p-Valor</i>
Model	21,9432	5	0,0005

Porcentaje de varianza explicado por el modelo = 34,2092

En las características altura a los 4 meses y fibra a los 12 meses se observó que los alelos marcadores asociados, identificados con el análisis univariado no

explican un porcentaje considerable de la variación de la característica para el caso de altura a los 4 meses, y en el caso de fibra a los 12 meses no se explica variación, sumado a esto los modelos no son estadísticamente significativos (altura a los 4 meses p -valor=0.12; fibra a los 12 meses p -valor = -8,54), lo que indica que es necesario incluir más marcadores dentro del análisis de regresión ya que estos, según el análisis, son insuficientes para explicar la variación de la característica en la población (Tablas 9 y 11).

Lo anterior se confirma al realizar el análisis de regresión logística con todos los alelos marcadores evaluados en el estudio, para las mismas características, altura a los 4 meses y fibra a los 12 meses; con el análisis se observa que al incluir todos los marcadores la explicación de la varianza aumenta en 22,4% y 34,2% respectivamente, con modelos estadísticamente significativos (p -valor =0,0087 y p -valor =0,0005) (Tablas 12 y 10).

Estos resultados siguen la lógica de comportamiento esperada para características cuantitativas, donde la acción conjunta de los diferentes alelos implicados en la característica explica una mayor variación que la dada por un solo alelo, además se observa cómo el comportamiento de cada uno de los alelos está determinado por la interacción con otros, esto se evidencia cuando los marcadores identificados por el ANOVA (altura a los 4 meses: BS334b, CIR18a, CIR36c, CIR47a. Fibra a los 12 meses: CG236f, CIR15a, CIR32d, CIR47b) y los encontrados con el análisis de regresión múltiple (altura a los 4 meses: CIR10a,

CIR10b, CIR31c, CIR 31f. Fibra a los 12 meses: CIR11b, CIR36e, CIR52a, CIR40a, CIR32a) son diferentes.

Clorofila a los 4 meses

El resultado del análisis de regresión logística sólo con los marcadores identificados con el análisis univariado, arrojó un resultado diferente al encontrado para altura a los 4 meses y fibra a los 12 meses. En éste los marcadores encontrados inicialmente con el ANOVA explican 10,41% de la variación de la característica con un modelo estadísticamente significativo ($p=0.003$), donde se destacan los marcadores CIR19b y CL1608b por tener un efecto positivo en la característica (Tabla 13).

Tabla 13. Se muestran los resultados del análisis de regresión logística incluyendo solo los alelos marcadores identificados con ANOVA, para clorofila a los 4 meses.

Modelo estimado de regresión (Máxima probabilidad)

Parametro	Estimado	Error	
		Estandar	Razón de probabilidad
CIR15a=0	-0,643032	0,412809	0,525696
CIR19b=0	0,475976	0,40731	1,60958
CIR31c=0	-0,62286	0,395929	0,536408
CIR47b=0	-0,0105227	0,448475	0,989533
CL1557d=0	-0,395542	0,78053	0,673315
CL1608b=0	0,650928	0,368512	1,91732
CS668b=0	-0,977059	0,408709	0,376416

Análisis de variación

Source	Variación	Df	p-Valor
Model	15,2431	7	0,0330

Porcentaje de variación explicado por el modelo = 10,41

Sin embargo, cuando se incluyeron todos los marcadores evaluados en este estudio en el análisis de regresión logística, se observa que sólo tres marcadores CL1557d, CL1557e y CIR31c explican 16,18% de la variación fenotípica en la población, dos de ellos ya habían sido identificados con el ANOVA (CL1557d, CIR31c) (Tabla 14).

Tabla 14. Se muestran los resultados del análisis de regresión logística incluyendo todos los alelos marcadores evaluados en éste estudio (45), para clorofila a los 4 meses.

Modelo estimado de regresión (Máxima probabilidad)

		<i>Error</i>	<i>Estimado</i>
<i>Parametro</i>	<i>Estimado</i>	<i>Estandar</i>	<i>Razón de probabilidad</i>
CL1557d=0	-16,2604	505,776	8,67355E-8
CL1557e=0	15,5712	505,777	5,78729E6
CIR31c=0	-0,92178	0,33971	0,39781

Análisis de variación

<i>Source</i>	<i>Variación</i>	<i>Df</i>	<i>p-Valor</i>
Modelo	24,9947	3	0,0000

Porcentaje de variación explicado por el modelo = 16,186

A pesar de que en el análisis de regresión logística donde sólo se incluyeron los marcadores encontrados con el ANOVA y en el que se incluyeron todos los marcadores evaluados en este estudio, muestra que los marcadores CL1557d y CIR31c tiene un efecto negativo, difieren en la proporción y en la probabilidad de que disminuya la característica cuando el marcador está presente (Tabla 15).

Tabla 15. Se muestra la razón de probabilidad para la asociación de los alelos marcadores en el análisis de regresión logística, donde se incluyen sólo los alelos marcadores asociados identificados con el ANOVA y el análisis incluyendo los 45 alelos marcadores.

ALELO MARCADOR	REGRESIÓN LOGÍSTICA CON MARCADORES IDENTIFICADOS CON ANOVA			REGRESIÓN LOGÍSTICA CON LOS 45 ALELOS MARCADORES
	Razón de la probabilidad	p - Valor del modelo	Razón de la probabilidad	p - Valor del modelo
CL 1557d	0,673315	0,0330	8,67355E-8	0,0000
CIR31c	0,536408		5,78729E6	

El hecho de que las probabilidades y proporciones de disminución o aumento de la característica de un mismo marcador cuando está presente varíen de acuerdo al número de marcadores introducidos en el análisis, como sucedió para los marcadores CL1557d y CIR31c, asociados a clorofila a los 4 meses, y el hecho de que se haya identificado un marcador asociado que no había sido detectado por el ANOVA como sucedió para altura a los 4 meses y fibra a los 12 meses, soportan lo mencionado anteriormente, con relación a que el aporte de cada alelo en la variación de la característica está determinada por la interacción con otros marcadores.

El análisis de regresión logística también se realizó para todas las características evaluadas, incluyendo los 45 alelos marcadores evaluados en el estudio, éste análisis permitió identificar combinaciones de alelos que en conjunto explican una proporción determinada de la variación de las características, las combinaciones de alelos reportadas en la tabla 16 son las que mostraron modelos estadísticamente significativos; de estas combinaciones cabe resaltar ocho, que

presentan porcentajes de explicación de la variación de las características mayores de 18%, así se tienen:

Carcaterística	Varianza explicada por el modelo	p- Valor del modelo
Altura 12 meses	24,76	0,0199
Diámetro 12 meses	19,17	0,0018
Humedad 13 meses	33,9	0,0001
Sacarosa 12 meses	26,91	0,0002
Sacarosa 13 meses	34,24	0,0000
Brix 13 meses	19,55	0,0004
Fibra 12 meses	34,2	0,0005
No sacarosa 13 meses	43,54	0,0000
Azúcares reductores 12 meses	24,27	0,008

Con el análisis de regresión logística se encontraron asociaciones que no habían sido identificadas con el análisis univariado. Así se encontraron asociados ocho (8) marcadores CIR10b, CIR15b, CG319b, CIR32a, CIR36b, CIR52a, CIR11b, y CIR10a (Tabla 16); también se detectaron marcadores asociados con diámetro a los 12 meses, en conjunto estos marcadores explicaron 19,17% de la variación de la característica en la población.

El hecho de haber encontrado nuevos marcadores asociados que no habían sido detectados con el análisis de varianza, puede explicarse posiblemente porque los marcadores moleculares, como ya ha sido mencionado, varían su comportamiento de acuerdo a su interacción con los demás marcadores; el análisis univariado no permite detectar esta interacción entre marcadores, contrario a la regresión logística, éste análisis permite correlacionar todos los marcadores que aportan positiva o negativamente a la característica y establece el porcentaje de variación explicado por el comportamiento conjunto de los marcadores.

6. CONCLUSIONES

Con el análisis univariado se detectaron asociaciones estadísticamente significativas de cuatro marcadores a tres características, altura a los 4 meses (CIR18a; $p=0.0016$), porcentaje de fibra a los 12 meses (CIR15a, $p=0.0016$) y clorofila a los 4 meses (CIR32b; $p=0.0018$ y CIR1557d; $p= 0.0001$).

Con éste mismo análisis se confirma la asociación de los marcadores CIR 18 y CIR36 a brix reportados por Piperidis *et al.*, (2008), también se encontraron marcadores asociados a diferentes características en diferentes edades del cultivo con bajos niveles de significancia estadísticas.

Sin embargo el análisis de regresión logística el cual permite determinar el comportamiento de los marcadores en conjunto, mostró asociaciones muy diferentes a las encontradas con el análisis de varianza realizado inicialmente.

Con el análisis de regresión logística se observó que el efecto de cada uno de los alelos marcadores asociados a una característica es determinado por la interacción entre los diferentes marcadores que se incluyen en el análisis, por ende la identificación de las combinaciones de alelos que potencien las

características importantes agronómicamente, es sensible al número de marcadores que se incluyan en el mismo.

En este estudio se reportan las combinaciones de alelos marcadores para altura a los 12 meses, diámetro a los 12 meses, humedad a los 13 meses, sacarosa a los 12 y 13 meses, brix a los 13 meses, no sacarosas a los 13 meses y azúcares reductores a los 12 meses, que explican un porcentaje de variación de la característica en la población mayor de 18%.

Estas combinaciones de alelos que explican porcentajes de variación de la característica mayores de 18% son específicas para las condiciones ambientales que se evaluaron en este trabajo, así mismo, como el comportamiento de cada alelo marcador se encontró determinado por los alelos con los que interacciona, las combinaciones de alelos marcadores reportadas aquí también son específicas de los marcadores evaluados en este estudio.

Posiblemente al aumentar el número de marcadores y evaluarlos en conjunto para las mismas características se encontrarán nuevas asociaciones y combinaciones de alelos marcadores que expliquen un mayor porcentaje de variación de las características evaluadas.

Por ende es necesario aumentar el número de marcadores analizados para ampliar la cobertura del genoma y evaluar nuevas combinaciones, incluyendo las

reportadas en este estudio. También es necesario realizar análisis que además de mostrar el comportamiento de cada marcador en conjunto permita detectar posibles interacciones génicas como epistásis, información que será de gran utilidad en la selección de combinaciones de alelos para aplicar en la selección asistida por marcadores.

Con el fin de tener un mayor cubrimiento del genoma de la caña de azúcar y teniendo en cuenta el gran desafío que representa la saturación del mapa genético por su tamaño y complejidad genómica, una opción que puede ser contemplada es la elaboración de mapas por mapeo comparativo con especies estrechamente relacionadas con el género y con alta homología con el genoma de caña de azúcar, como arroz, sorgo y maíz,

Finalmente es necesario elaborar investigaciones que conduzcan al conocimiento citogenético básico de las variedades de caña de azúcar colombianas, que aporten información como proporción del genoma constituido por *S. officinarum* y *S. spontaneum* y número de cromosomas total ($2n$), información que aportará al conocimiento y entendimiento de la compleja estructura genómica de las variedades de caña de azúcar, que será de gran utilidad en la elaboración del mapa de las variedades colombianas y su aplicación en los programas de mejoramiento con la implementación de la selección asistida por marcadores.

Todos los resultados obtenidos en este estudio son análisis preliminares, o exploratorios que aportan información inicial sobre el genoma de la caña de azúcar, este trabajo es un inicio de un largo camino por recorrer en el entendimiento del genoma de la caña de azúcar de las variedades Colombianas.

7. RECOMENDACIONES

Se recomienda para futuros estudios aumentar el número de marcadores moleculares microsatélites e incluir marcadores con buen nivel de polimorfismo como los AFLP's, incluso marcadores como los EST – SSR que serán de mucha utilidad en la identificación de genes de importancia agronómica y combinaciones de alelos marcadores que expliquen altos porcentajes de la variación de los componentes de producción de biomasa y sacarosa.

Otra metodología que se recomienda aplicar en el proceso de cobertura del genoma es mapeo comparativo, donde se incluyan especies que tengan alta homología con el genoma de caña de azúcar, como arroz, sorgo y maíz.

Se recomienda realizar estudios citogenéticos que contribuyan con el conocimiento y entendimiento del complejo genoma de las variedades de caña de azúcar colombianas

8. BIBLIOGRAFÍA

LITERATURA CITADA

AITKEN KS, HERMANN S, KARNO K, BONNETT GD, MCINTYRE LC, JACKSON PA. 2008. Genetic Control of yield related stalk traits in sugarcane. *Theor Appl Genet* 117:1191-1203.

AITKEN KS, JACKSON PA, MCINTYRE CL. 2005. A combination of AFLP and SSR markers provides extensive map coverage and identification of homo(eo)logous linkage groups in a sugarcane cultivar. *Theor Appl Genet* 110: 789-801.

AITKEN KS, JACKSON PA, MCINTYRE CL. 2006. Quantitative trait loci identified for sugar related traits in a sugarcane (*Saccharum* spp.) cultivar x *Saccharum officinarum* population. *Theor Appl Genet* 112: 1306-1317.

AL – JANABI SM, HONEYCUTT RJ, MCCLELLAND M, SOBRAL BWS. 1993. A genetic linkage map of *Saccharum spontaneum* L. 'SES 208'. *The Genetci Society of America* 134: 1249-1260.

AL- JANABI AM, FORGET L, DOOKUN A. 1999. An Improved and rapid protocol for the isolation of polysaccharide- and polyphenol – free sugarcane DNA. *Plant Molecular and Biology Reporter* 17:1-8.

ASOCAÑA. Informe anual. 2001 – 2002.

BATH SR, GILL SS.1985. The implications of 2n egg gametes in nubilization and breeding of sugarcane. *Euphytica* 34 : 377 – 384.

CABRERA AF, SALAZAR EI. 2002. Mejoramiento genético de plantas. Universidad Nacional de Colombia. Cali, Palmira.402p.

CENICAÑA (Centro de Investigaciones de la Caña de Azúcar de Colombia).1995. El cultivo de la caña de azúcar en la zona azucarera de Colombia. C.(eds.) Cali, Colombia. 412p.

COLLARD BCY; JAHUFER MZZ; BROUWER JB; PANG ECK. 2005. An introduction to markers, quantitative trait loci (QTL) mapping and marker-

assisted selection for crop improvement: The basic concepts. *Euphytica* 142: 169 – 196.

D'HONT A, GRIVET L, FELDMANN P, RAO S, BERDING N, GLASZMANN JC. 1996. Characterisation of the double genome structure of modern sugarcane cultivars (*Saccharum* spp.) by molecular cytogenetics. *Mol Gen Genet* 250:405 – 413.

DA SILVA J, HONEYCUTT RJ, BURNQUIST W, AL-JANABI SM, ZORREÍIS ME, TANKSLEY SD, SOBRAL BWS. 1995. *Saccharum spontaneum* L. 'SES 208' genetic linkage map combining RFLP and PCR-base markers. *Molecular breeding* 1: 165- 179.

DA SILVA J, SORRELLS ME, BURNQUIST WL, TANKSLEY SD. 1993. RFLP linkage map and genome analysis of *S. spontaneum*. *Genome* 36: 782 – 791.

FERREIRA ME; GRATTAPAGLIA D. 1998. Introducción al uso de marcadores moleculares en el análisis genético. 1ed. Brasilia: EMBRAPA – CENARGEN. pp. 220.

GARCÍA AAF, KIDO EA, MEZA AN, SOUZA HMB, PINTO LR, PASTINA MM, LEITE CS, DA SILVA JAG, ULIAN EC, FIGUEIRA A, SOUZA AP. 2006. Development of an integrated genetic map of a sugarcane (*Saccharum* spp.) commercial cross, based on a maximum-likelihood approach for estimation of linkage and linkage phases. *Theor Appl Genet* 112: 298-314.

GILBERTSON RL, ROJAS MR, RUSSEL LD, MAXELL DP. 1991. The use of the asymmetric polymerase chain reaction and DNA sequencing to determine genetic variability among isolates of bean golden mosaic geminivirus in the Dominican Republic. *J. Gen. Virol.* 72: 2343 – 2848.

GRIFFITHS JF, MILLER JH, SUZUKI DT, LEWONTIN RC, WILLIM M. 1999. Introduction to genetics analysis. New York. W. H. Freeman & Co.
GRIVET L, ARRUDA P. 2001 . Sugarcane genomics: depicting the complex genome of an important tropical crop. *Current Opinion in Plant Biology* 5: 122 -127.

GIVET L, ARRUDA P. 2001. Sugarcane genomics: depicting the complex genome of an important tropical crop. *Current Opinion in Plant Biology* 5:

122 – 127.

GRIVET L, D'HONT A, ROQUES D, FELDMANN P, LANAUD C, GLAZMANN JC. 1996. RFLP mapping in cultivated sugarcane (*Saccharum* spp.): Genome Organization in a Highly Polyploid and Aneuploid Interspecific hybrid. *The Genetics Society of America* 142: 987-1000.

HOARAU JY, OFFMANN B, D'HONT A, RISTERUCCI AM, ROQUES D, GLAZMANN JC, GRIVET L. 2001. Genetic dissection of a modern sugarcane cultivars (*Saccharum* spp.) I. Genome mapping with AFLP markers. *Theor Appl Genet* 103: 84-97.

HUI LB. 1998. *Statistical genomics: linkage, mapping, and QTL analysis*. CRC press.

JACKSON AP. 2005. Breeding for improved sugar content in sugarcane. *Field Crops Research* 92: 277-290.

JANNNOO N, GRIVET L, SEGUIN M, PAULET F, DOMAINGUE R, RAO PS, DOOKUN A, D'HONT A, GLAZMANN JC. 1999. Molecular investigation of the genetic base of sugarcane cultivars. *Theor Appl Genet* 171-184.

JONES N; OUGHAM H; THOMAS H. 1997. Markers and mapping: we are all geneticists now. *New phytol* 137: 165 – 177.

KLUG WS, CUMMINGS MR, SPENCER CA. 2006. *Conceptos de genética*. 8^a ed. Madrid, Pearson Prentice Hall. 884p.

KOLE, CH. 2007. *Genome Mapping and molecular Breeding in Plants, Volumen 3 Pulses, Sugar and Tuber Crops*. Springer-Verlag Berlin Heidelberg.

LANDER ES, GREEN P, ABRAHAMSON J, BARLOW A, DALY MJ, LINCOLN SE, NEWBURG L. 1987. MAPMAKER: an interactive computer package for constructing primary linkage maps of experimental and natural populations. *Genomics* 1: 174 -181.

LINCOLN SM, DALY M, LANDER E. 1993. *Constructing genetic linkage maps with MAPMAKER/EXP. Version 3.0*. Whitehead Institute for Biomedical Research Technical Report, 3rd Edn.

MANLY KF, CUDMORE R, MEER JM. 2001. Map Manager QTX, cross-platform software for genetic mapping. *Mamm Genome* 12: 930-932.

MEKSEN K; KAHL G. 2005. The handbook of plant genome mapping. WILEY-VCH Verlag GmbH & Co. KGaA. pp.380.

MING R, LIU SC, LIN YR, DA SILVA J, WILSON W, BRAGA D, DEYNZE A, WENSLAFF TF, WU KK, MOORE PH. 1998. Detailed alignment of *Saccharum* and *Sorghum* chromosomes: comparative organization of closely related. *Genetics* 150: 1663 – 1682.

MING R, WANG YW, DRAYE X, MOORE PH, IRVINE JE, PATERSON AH. 2002. Molecular dissection of complex traits in autopolyploids: mapping QTL affecting sugar yield and related traits in sugarcane. *Theor Appl Genet* 105:332-345.

MUKHERJEE SK.1957.Origin and distribution of saccharum. *Botanical Gazette*119(1):55-61.

OLIVERA KM, PINTO LR, MARCONI TG, MARGARIDO GRA, PASTINA MM, TEIXEIRA MHL, FIGUEIRA AV, ULIAN EC, GARCÇIA AAF, SOUZA AP. 2007. Functional integrated genetic linkage map based on EST-markers for a sugarcane (*Saccharum* spp.) commercial cross. *Mol Breeding* 20: 189-208.

PARIDA SK, KALIA SK, KAUL S, DALAL V, HEMAPRABHA G, SELVI A, PANDIT A. SINGH A, GIKWAD K, SHARMA TR, SRIVASTAVA PS, SINGH KN, MOHAPATRA T. 2009. Informative genomic microsatellite markers for efficient genotyping applications in sugarcane. *Theor Appl Genet* 118(2):327-338.

PIPERIDIS N, JACKSON PA, D'HONT A, BESSE P, HOARAU JY, COURTIS B, AITKEN KA, MCINTYRE CL. 2008. Comparative genetics in sugarcane enables structured map anchorage and validation of marker-trait associations. *Mol Breeding* 21:223-247.

QUINTERO R, SÁNCHEZ A, LOMBANA A, ARBOLEDA F, TORRES JS, GONZÁLEZ JA, MURILLO CA. 2008 Grupos homogéneos de suelos del área dedicada al cultivo de la caña de azúcar en el valle del río Cauca (segunda aproximación). *Cenicaña*. Cali. 106p. (Serie Técnica no. 37)

ROSSI M, ARAUJO P, GARSMEUR O, DIAS VM, CHEN H, VANSLEYS MA, D'HONT A. 2003. Genomic distribution and characterization of EST - derived resistance gene analogs (RGAs) in sugarcane. *Mol Genet Genomics* 269: 406 – 419.

SEMAGN K, BJORNSTAD A, NDJIONDJOP MN. 2006. Progress and prospects of marker assisted backcrossing as a tool in crop breeding programs. *African Journal of Biotechnology* 5(25):2588-2603 c.

SEMAGN K, BJORNSTAD A, NDJIONJOP MN. 2006a .An overview of molecular marker methods for plants. *African journal of Biotechnology* 5(25):2540-2568 a.

SEMAGN K, BJORNSTAD A, NDJIONJOP MN. 2006b. Principles, requirements and prospects of genetic mapping in plants. *African journal of Biotechnology* 5(25):2569-2587 b.

SINGH RK, SINGH P, SINGH SP, MOHAPATRA T, SINGH SB. 2006. Mapping QTLs for sugar content and segregation analysis in sugarcane. *Sugar Cane International* 24(1): 7 – 13

STAM P. 1993. Construction of integrated genetic linkage maps by means of a new computer package: JoinMap. *Plant J* 3: 739 – 744

TELLLES GP, BRAGA MDV, ZANONI D, LIN TL, QUITZAU JAA, DA SILVA FR, MEIDANIS J. 2001. Bioinformatics of the sugarcane EST project. *Genetics and Molecular Biology* 24(1-4): 9-15.

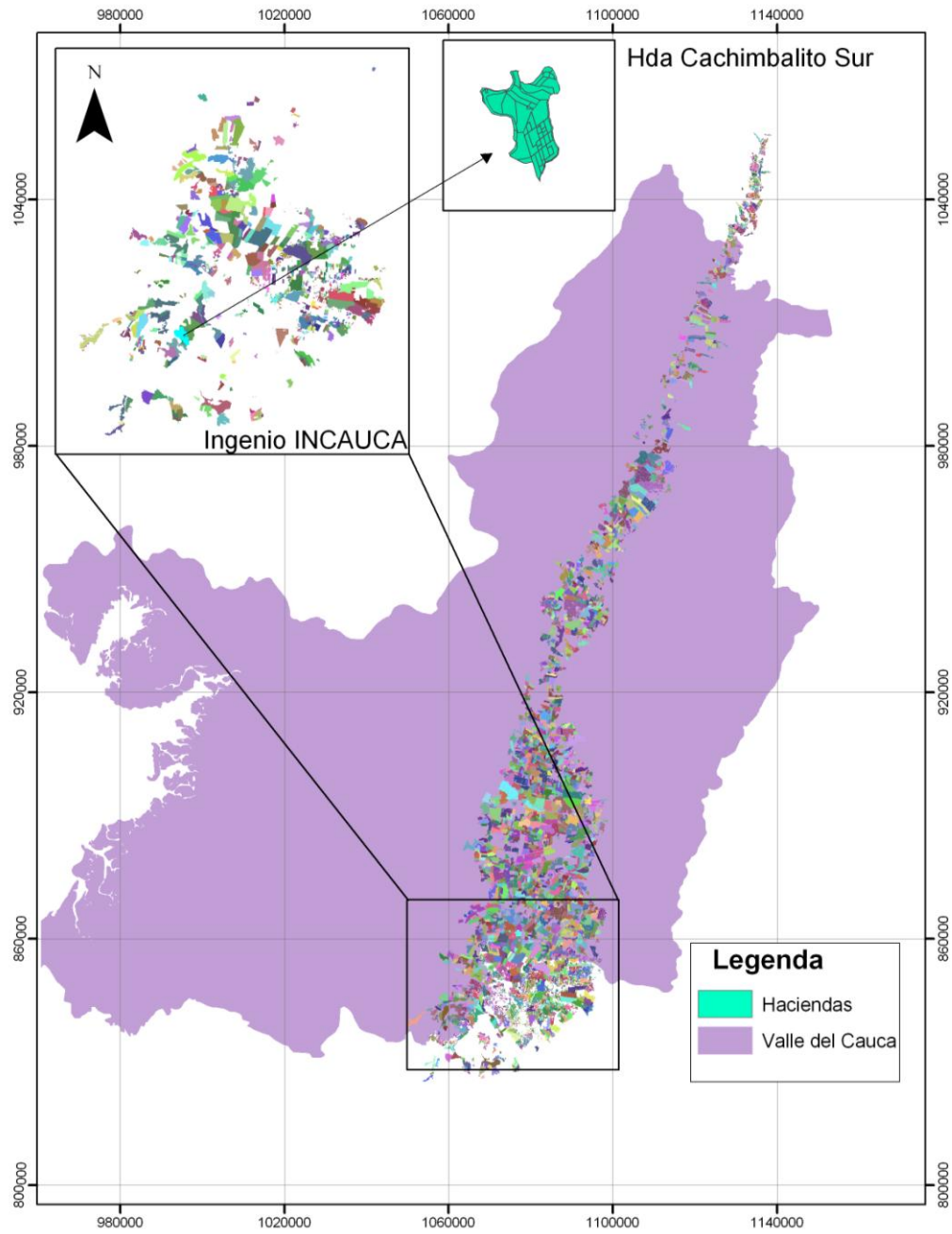
WU KK, BURNQUIST W, SORRELLS ME, TEW TL, MOORE PH, TANKSLEY SD. 1992. The detection and estimation of linkage in polyploids using single – dose restriction fragments. *Theor Appl Genet* 83:294-300.

LITERATURA CONSULTADA

BREMER G. 1961. Problems in breeding and cytology of sugar cane. *Euphytica* 10: 59-78.

CORDEIRO G, AMOUYAL O, ELIOT F, HENRY R. 2007. Pulses, Sugar and Tuber Crops. En: Chittaranjan Kole (ed) *Genome Mapping and Molecular Breeding in Plants III*. Springer - Verlag Berlin Heidelberg, pp175-203.

ANEXO A Ubicación geográfica de la hacienda Cachimbalito (ingenio Incauca)



ANEXO B. Método de extracción de ADN de Caña de Azúcar (*Saccharum spp*)

1. Congelar con nitrógeno líquido 3g de hoja joven de caña de azúcar y macerar.
2. Adicionar el macerado a un tubo con 12 ml de buffer de extracción previamente calentado a 65°C.
3. Adicionar 800ul de SDS 20% e invertir suavemente para homogenizar la mezcla.
4. Incubar en baño maría por 10 minutos a 65°C.
5. Adicionar 4000ul de acetato de potasio 5M.
6. Centrifugar 20 minutos a 10000 rpm/4°C.
7. Filtrar el sobrenadante usando gasa, ésta debe doblarse las veces que sea necesario hasta que se pueda garantizar la completa separación de la fase orgánica de la fase acuosa.
8. Adicionar isopropanol frio (-20°C).en un volumen igual al que se tiene de la fase acuosa, después de la filtración.
9. Invertir tubos hasta observar el ADN.
10. Centrifugar 20 minutos a 10000 rpm/4°C.
11. Eliminar el sobrenadante.
12. Lavar el pellet con 500ul de etanol al 70%.
13. Dejar secar el pellet, aproximadamente media hora.
14. Resuspender en 500ul de buffer TE.

15. Adicionar a cada tubo 2ul de RNAsa y colocar a 37°C por 20 minutos.

16. Cuantificar y evaluar calidad del ADN en gel de agarosa al 0.8%.

17. Almacenar a -20°C.

Nota: Antes de precalentar el buffer de extracción adicionar 34.8ul de β -mercaptoetanol y 1g de PVP (Polivinilpirrolidona) para 100 ml.

ANEXO C Protocolo de tinción con plata para geles de poliacrilamida

1. Sumergir el gel por 10 minutos en solución fijadora.
2. Hacer dos lavados, cada uno de 2 minutos.
3. Sumergir el gel en ácido nítrico durante 3 minutos.
4. Lavar el gel con agua destilada durante 1 minuto.
5. Sumergir el gel en la solución de nitrato de plata (AgNO_3) durante 20 minutos.
6. Lavar el gel durante 10 segundos.
7. Sumergir el gel en solución reveladora fría (12°C) hasta visualizar las bandas.
8. Sumergir el gel en solución de parada durante 3 minutos.
9. Lavar el gel aproximadamente 5 minutos.

Nota: El proceso de tinción debe hacerse con agitación constante; todos los lavados se hacen con agua destilada.

Preparación soluciones de tinción:

Fijadora: en 1L de agua adicionar 200ml de etanol absoluto y 20 ml de ácido acético 100% y aforar a 2L.

Ácido Nítrico: En 1L de agua diluir 30 ml de ácido nítrico y aforar a 2L.

Plata: A 2 L de agua adicionar 4g de plata (AgNO_3) y agitar hasta que esté bien diluida.

Reveladora: Adicionar a 1,5L de agua 60g de Carbonato de Sodio, agitar hasta disolver y aforar a 2L. Esta solución debe usarse fría (12-14°C).

Nota. Antes de usar esta solución adicionar 1ml de Formaldehído ($\text{H}_2\text{C}=\text{O}$) (37%) y 400ul de Tiosulfato de sodio. ($\text{Na}_2\text{S}_2\text{O}_3$). Para preparar el tiosulfato de sodio adicionar a 25 ml de agua 0.5 g de tiosulfato de sodio y aforar a 50 ml, homogenizar la solución.

Parada: Adicionar a 1L de agua 100ml de ácido acético y aforar a 2L.

Nota: Todas las soluciones se preparan con agua desionizada.