

TRABAJO DE GRADO

Análisis y validación de microsatélites EST-SSRs como candidatos a marcadores moleculares de la tolerancia a estrés hídrico en una colección de clones de melina (*Gmelina arborea* Roxb.)

DIANA ALEJANDRA QUISOBONY PAREDES

Presentado para obtener el título de

BIÓLOGA

PONTIFICIA UNIVERSIDAD JAVERIANA

FACULTAD DE CIENCIAS

CARRERA DE BIOLOGÍA

BOGOTÁ, D.C.

NOVIEMBRE 2011



PONTIFICIA UNIVERSIDAD JAVERIANA

FACULTAD DE CIENCIAS

BIOLOGÍA

**ANÁLISIS Y VALIDACIÓN DE MICROSATÉLITES EST-SSRs COMO
CANDIDATOS A MARCADORES MOLECULARES DE LA TOLERANCIA A
ESTRÉS HÍDRICO EN UNA COLECCIÓN DE CLONES DE MELINA (*Gmelina
arborea* Roxb.)**

DIANA ALEJANDRA QUISOBONY PAREDES

Presentado para obtener el título de

BIÓLOGA

Ingrid Schuler
Decana Académica

Andrea Forero
Directora de Carrera

PONTIFICIA UNIVERSIDAD JAVERIANA

FACULTAD DE CIENCIAS

BIOLOGÍA

**ANÁLISIS Y VALIDACIÓN DE MICROSATÉLITES EST-SSR'S COMO
CANDIDATOS A MARCADORES MOLECULARES DE LA TOLERANCIA A
ESTRÉS HÍDRICO EN UNA COLECCIÓN DE CLONES DE MELINA (*Gmelina
arborea* Roxb.)**

DIANA ALEJANDRA QUISOBONY PAREDES

Presentado para obtener el título de

BIÓLOGA

Wilson Terán

Director Trabajo de Grado

Diana Álvarez

Par académico Trabajo de Grado

NOTA DE ADVERTENCIA

Artículo 23 de la resolución No 13 de Julio de 1946

“La Universidad no se hace responsable por los conceptos emitidos por sus alumnos en sus trabajos de tesis. Sólo velará porque no se publique nada contrario al dogma y la moral católica y porque la tesis no contenga ataques personales contra persona alguna, antes bien se vea en ellas el anhelo por buscar la verdad y la justicia”

AGRADECIMIENTOS

Agradezco a Wilson Terán, Pilar Márquez y Carolina Rosero quienes condujeron este proyecto, a Jairo Moreno y Gina Solano por guiarme en los primeros pasos de estos estudios, a Manuel Paredes y Lilian Torres por ser incondicionales en su ayuda y compartir sus conocimientos, a los integrantes de la UBV de quienes recibí todo el apoyo.

A mis padres por apoyarme siempre en el desarrollo de mis sueños

Y a mis amigas por estar ahí sin condiciones

ÍNDICE DE TABLAS Y FIGURAS

TABLAS

Tabla 1: Familias de Factores de Transcripción involucrados en respuesta/tolerancia a estrés hídrico.....	8
Tabla 2: Genes de respuesta a estrés hídrico portadores de secuencias microsatelitales (SSR's) identificados en <i>Gmelina arborea</i> Roxb. (Rosero <i>et al.</i> 2011) y seleccionados en este estudio.....	9
Tabla 3: Identificación de la ubicación de las secuencias microsatelitales y condiciones de amplificación utilizadas.....	23
Tabla 4: Resultados de los productos amplificados EST-SSRs.....	27
Tabla 5: Clusters generados a partir del análisis de UPGMA con distancias mínimas de Nei (1972).....	29
Tabla 6: Índices de heterocigosis y polimorfismo.....	29

FIGURAS

Figura 1: Fenómenos de <i>slippage</i> durante la replicación de secuencias microsatelitales tomado de Ellegren (2004).....	16
Figura 2: Análisis de la variabilidad genética en la población de 33 clones de <i>Gmelina arborea</i> de la colección de Pizano S.A (Solano 2010).....	18
Figura 3: Patrones de geles de electroforesis en agarosa al 2%.....	24
Figura 4: Primer CL3C1. Rango alélico esperado, Bandas stutter y Bandas más intensas con mayor peso molecular.....	26
Figura 5: Dendrograma basado en distancias mínimas de Nei (1972).....	28
Figura 6: Frecuencias alélicas observadas.....	29

1. RESUMEN

La melina (*Gmelina arborea* Roxb) es una especie forestal originaria de países asiáticos cuya distribución se ha extendido a diferentes países tropicales, incluyendo Colombia, debido a sus características de rápido crecimiento, fácil adaptación a las condiciones adversas, además de la versatilidad de usos de su madera; atributos que le han permitido posicionarse entre las especies maderables tropicales de mayor potencial económico.

Sin embargo, a pesar de su exitosa adaptación a las condiciones edafoclimáticas de la región Caribe colombiana y su conocida capacidad de tolerancia al estrés hídrico, las temporadas de sequía que se han presentado en las últimas décadas en esta región del país, limitan las posibilidades de expansión y los rendimientos de las siembras comerciales de esta especie. Los rasgos de tolerancia a estrés hídrico y la respuesta a sequía son de carácter multigénico y asimismo su expresión obedece a mecanismos regulatorios complejos. Por ello, el poder identificar genotípicamente individuos que presenten estas características es un reto para el mejoramiento genético de esta especie frente a esta problemática. Los marcadores moleculares se presentan como una herramienta básica para la selección asistida de individuos tolerantes a este tipo de estrés abiótico.

En el presente estudio se analizaron 33 clones pertenecientes al programa de mejoramiento genético de la empresa Pizano S.A. utilizando 28 parejas de primers para microsatélites génicos tipo EST-SSRs (*Expressed Sequence Tag - Single Sequence Repeats*) obtenidas a partir de un estudio de perfiles transcriptómicos bajo estrés hídrico en *Gmelina arborea* [1]. Se estandarizaron las condiciones de amplificación por PCR para 14 parejas de primers seleccionados y se realizaron perfiles de electroforesis PAGE. El índice de polimorfismo para los respectivos marcadores microsatelitales se estimó por medio del programa GenAlEx. Los resultados obtenidos en forma preliminar demostraron un alto nivel de polimorfismo en tres de los microsatélites analizados: CL473C2, CL493C1 y CL1054C1 validando su uso potencial como marcadores moleculares de tolerancia a sequía en futuros programas de mejoramiento genético aún incipientes en esta especie.

2. INTRODUCCIÓN

En Colombia, las plantaciones de melina se encuentran ubicadas mayoritariamente en regiones de Caribe seco tropical (Sur de Bolívar, Magdalena, Córdoba); las condiciones de esta ubicación favorecen el crecimiento de la especie y permiten el desarrollo de características de la madera importantes a nivel económico; sin embargo las prolongadas estaciones de sequía de la región limitan el desarrollo de las plantaciones y su rendimiento.

Los esfuerzos en la selección de individuos tolerantes a estrés hídrico han ido encaminados a la selección por marcadores morfológicos y ecofisiológicos, arrojando resultados no extrapolables para todos los clones de la población además de requerir de tiempos prolongados en términos de selección y mejoramiento. Los estudios moleculares realizados en esta especie son muy escasos, limitando el desarrollo de programas de mejoramiento asistido por marcadores moleculares. La Unidad de Biotecnología Vegetal ha sido pionera en este sentido al iniciar un programa que incluyó por un lado un proyecto de caracterización molecular de clones de melina por medio de marcadores moleculares tipo AFLPs [2] y la generación de un primer recurso genómico para la especie basado en secuenciación masiva y anotación de transcriptoma para la búsqueda de secuencias génicas involucradas en la respuesta a estrés hídrico. Este primer recurso genómico ha permitido identificar marcadores microsatelitales génicos (EST-SSRs) asociados con respuesta a estrés hídrico y cuya validación podría permitir su uso en posteriores proyectos de selección asistida y mejoramiento genético.

Los marcadores moleculares tipo microsatélites o SSRs (*Simple Sequence Repeats*) son marcadores codominantes basados en la longitud polimórfica de productos amplificados de ADN [3] Presentan alta confiabilidad en los resultados debido a su reproducibilidad, multialelismo y especificidad, sin embargo debido a la posible presencia de artefactos durante la amplificación se recomienda realizar las matrices tomando en cuenta sólo las bandas con mayor intensidad [4]. Las secuencias microsatelitales pueden ubicarse en cualquier región del genoma, tanto en regiones no codificantes o extra génicas, como en regiones expresadas del genoma denominándose microsatélites génicos. Recientemente, las herramientas de genómica funcional y secuenciación de transcriptomas han permitido obtener a bajo costo, gran cantidad de secuencias expresadas que corresponden

generalmente a secuencias génicas incompletas conocidas como “etiquetas de secuencias expresadas” (ESTs). Este tipo de recursos ha permitido obtener e identificar un gran número de microsatélites génicos llamados EST-SSRs, que presentan ventajas como su rápida obtención a partir de bases de datos de ESTs y su localización en genes de interés involucrados en determinadas funciones fisiológicas o metabólicas involucradas con rasgos particulares. El presente estudio permitió validar tres secuencias microsatelitales tipo EST-SSRs en 33 individuos de una colección de clones élite de melina, proveniente del banco de germoplasma de los viveros de Monterrey Forestal Ltda, pertenecientes a la empresa Pizano S.A.

3. PROBLEMA DE INVESTIGACIÓN

3.1. Planteamiento del Problema:

La sequía es uno de los principales factores de estrés abiótico en plantas que genera pérdidas en los cultivos de importancia económica en todo el mundo [5]. Aproximadamente el 28% de los suelos cultivados se ven afectados por este fenómeno [6]. Los cambios en los niveles de agua en los tejidos desencadenan una serie de procesos fisiológicos en la planta como lo son el cierre estomático, cambios en el ajuste osmótico y en la duración de las fases fenológicas, senescencia y abscisión de las hojas [7], perturbaciones en la transpiración y asimilación del CO₂, además puede generar cavitación y embolismo en el xilema debido a la alta presión que experimenta el tejido [8]. Estas reacciones fisiológicas generan a su vez disminuciones en la productividad y pérdidas de individuos dentro de los cultivos [9]. La selección y obtención de individuos tolerantes a sequía permite garantizar la adaptación de los cultivos frente a este tipo de estrés abiótico y por ende aumenta la productividad de los mismos bajo condiciones limitantes de agua [10].

La domesticación de los cultivos inicia a partir de la selección de individuos en poblaciones con altos niveles de variabilidad genética donde se permite escoger los mejores rasgos para la agricultura [5, 11]. En cultivos de melina, especie promisorio maderable originaria de países asiáticos que presenta cierto grado de tolerancia a sequía, se han realizado procesos de selección a partir de individuos élite; sin embargo la selección por caracteres

morfológicos presenta desventajas como la sensibilidad al ambiente, la escasez de polimorfismos o la epístasis [12], por lo tanto se hace necesario desarrollar en forma complementaria marcadores moleculares que permitan identificar las características de interés más eficiente y rápidamente, aumentando los beneficios para los cultivadores. Dentro de este contexto, el desarrollo de marcadores moleculares que se encuentren asociados con secuencias expresadas (ESTs) durante periodos de sequía pueden constituir buenos candidatos para estudios de asociación con rasgos de tolerancia a este tipo de estrés y por ende apoyar futuros programas de selección asistida por marcadores moleculares (*MAS molecular assisted selection*) [13, 14].

Actualmente se emplean diversos marcadores moleculares en la selección asistida entre los que se incluye los microsatélites basados en amplificación por PCR y caracterizados por su alto polimorfismo [15, 16], además la preparación de las muestras es sencilla, y no se requiere grandes concentraciones de ADN [16]. Los microsatélites pueden ubicarse en regiones génicas (exones, 5'UTR y 3'UTR) y de esta forma pueden regular la expresión de las mismas; así los polimorfismos génicos pueden asociarse con variaciones fenotípicas de interés para el mejoramiento del cultivo [17]. El estudio de secuencias SSRs asociadas a secuencias expresadas EST (EST-SSRs) durante periodos de estrés hídrico permite escoger secuencias polimórficas candidatas a marcadores moleculares con posibilidad de ser utilizados en programas de MAS dentro de las colecciones clonales de *Gmelina arborea* Roxb.

3.2. Justificación

La productividad de los cultivos de importancia económica como *G. arborea* [18], decrece debido al estrés hídrico que enfrentan por cambios en las condiciones climáticas, entre ellas la sequía. [10]. Empresas como Pizano S.A y en asociación con la cooperativa Camcore buscan garantizar la sostenibilidad de las plantaciones y recursos forestales a través del desarrollo de programas de mejoramiento y selección de individuos [2]. Por ello se hace necesaria la implementación de técnicas que permitan seleccionar mejores características de forma rápida y con mayor probabilidad de éxito. [19].

A pesar de que los marcadores morfológicos se presentan como una forma de selección de parentales [11], desventajas como la escasez de polimorfismos, la dominancia de los

fenotipos y la sensibilidad al ambiente [12], implican la búsqueda de alternativas de selección [20]. Por ello, la selección asistida por marcadores moleculares (MAS), se ha popularizado debido a su efectividad para la selección de individuos y obtención de variedades con características de interés agronómico [21]. El uso de marcadores moleculares asociados con rasgos de tolerancia a sequía permitirá la identificación y selección de individuos con esta característica, lo cual contribuirá a un aumento de la productividad de las plantaciones forestales de melina, acrecentando las utilidades de los productores [2]

3.3. Pregunta de investigación

¿Qué marcadores microsatelitales tipo EST-SSRs son polimórficos en una colección de 33 clones élite de *G. arborea* y pueden ser candidatos para estudios de asociación con tolerancia a sequía?

4. MARCO TEÓRICO

4.1. MARCO CONCEPTUAL

4.1.1. Melina (*Gmelina arborea* Roxb.)

Gmelina arborea es una especie leñosa, de la familia Lamiaceae (aunque comúnmente asociada a la familia Verbenaceae), de mediano a gran crecimiento, que crece naturalmente en regiones tropicales y subtropicales de Asia [22]. Pueden encontrarse individuos desde los 50 hasta 1300 metros de altitud [22], los cuales llegan a alcanzar hasta 30 m de altura, con un diámetro de tronco de hasta 1 m de ancho. Esta especie presenta características relevantes para el sector forestal: se le considera especie pionera debido a su capacidad de colonización de nichos ecológicos de forma eficiente, propiedad que es muy apreciada en proyectos de reforestación debido a que presenta una de las mayores tasas de crecimiento en especies maderables [22, 23]. Sus excelentes propiedades como la adaptabilidad a suelos secos, la resistencia al fuego, y la versatilidad de usos de su madera la hacen igualmente muy atractiva para el sector forestal industrial en países tropicales [1, 24]. En Colombia, los

cultivos de esta especie para uso maderable iniciaron en 1960 en la región Caribe, actualmente se presentan alrededor de 11 mil hectáreas industriales distribuidas en los departamentos de Bolívar, Magdalena, Córdoba, Llanos Orientales, Tolima y Cundinamarca [2].

4.1.2. Respuestas de las plantas a condiciones de Sequía

Durante las estaciones de sequía, las plantas sufren de estrés hídrico y generan diversas respuestas para su supervivencia que se incluyen dentro de dos principales estrategias: evasión del estrés o tolerancia al estrés [8, 25]. En la primera se mantienen los niveles de agua y turgencia estables aún cuando en el ambiente el recurso sea limitante; dentro de las adaptaciones fisiológicas y morfológicas incluidas en este mecanismo se encuentran el ajuste osmótico, crecimiento de raíces y la restricción de la apertura estomática [25]. Por su parte, la tolerancia al estrés implica mecanismos de tolerancia a la deshidratación, por lo tanto si bien la planta no logra mantener los niveles de agua y de turgencia, desarrolla mecanismos fisiológicos y/o morfológicos que le permiten sobrevivir al estrés. Esta estrategia se ejemplifica en los observados fenómenos de “resurrección de las plantas” que logran recuperarse después de perder aproximadamente el 90% de su contenido hídrico [8].

Los programas de mejoramiento genético van encaminados a escoger las características de evasión del estrés, en comparación con aquellas asociadas con tolerancia [25], si bien el término más usado se refiere a tolerancia. Otras de las respuestas fisiológicas que se presentan en periodos de estrés hídrico son la represión del crecimiento celular y fotosíntesis y la activación de la respiración celular. Además de estas respuestas fisiológicas, también se han elucidado respuestas a nivel celular y molecular debido a la acumulación de osmolitos y proteínas como las acuaporinas involucradas en la tolerancia al estrés abiótico [8].

4.1.3. Genes de tolerancia a estrés hídrico

Algunos autores [16, 26] sostienen la presencia de dos grupos de genes implicados en la respuesta a este tipo de estrés abiótico: el primer grupo conformado por los genes que codifican proteínas con función directamente relacionada con el estrés por sequía, también

llamadas proteínas efectoras; el segundo grupo lo integran aquellos genes codificadores de proteínas de regulación y transducción de señales.

En el primer grupo se ubicarán genes codificadores de proteínas con función enzimática o estructural conocida entre las que se encuentran enzimas productoras de osmoprotectores (como aminoácidos, poliaminas, azúcares y alcoholes), proteínas LEA, osmotinas, chaperonas, acuaporinas, ubiquitinas, proteasas y enzimas detoxificantes (catalasas, superóxido dismutasas, peroxidasas). Por otro lado se encuentran los genes reguladores que codifican osmosensores, quinasas, fosfatasas y factores de transcripción. [25].

Los proyectos de mejoramiento genético por medio de la biotecnología buscan modificar las características de algunas plantas de interés económico con genes que confieran tolerancia a estrés abiótico, sin embargo a veces la modificación de un gen no influye de manera significativa debido al gran número de vías de transcripción que se ven implicadas en los procesos de tolerancia a estrés hídrico, además de la tendencia natural de las plantas para recuperar la homeóstasis que oculta el efecto de la modificación de una única enzima en las vías bioquímicas [25].

Otra de las herramientas biotecnológicas consiste en la sobreexpresión de genes codificantes de factores de transcripción inducibles en condiciones de estrés; dentro de los reguladores transcripcionales estudiados que cumplen una función dentro de la respuesta al estrés hídrico se encuentran las familias: AP2/ERF, bZIP, NAC, MYB, MYC, Cys₂His₂, Zn finger, NFY y WRKY [26]. En la tabla siguiente (Tabla 1), se presenta una síntesis de la función de estas familias.

Tabla 1: Familias de factores de transcripción involucrados en respuesta/tolerancia a estrés hídrico.

<p>AP2/ERF</p> <p>APETALA2/ETHYLENE RESPONSE FACTOR</p>	<p>La familia de los factores de transcripción AP2 encontrados solo en plantas, incluyen genes codificadores de proteínas implicadas en la regulación de las vías de resistencia a estrés. Estos genes pertenecen a la subfamilia ERF (<i>ethylene response factor</i>), presentan un único dominio de unión al DNA lo cual los diferencia de la subfamilia DREB (<i>dehydration-responsive element binding</i>). [27].</p>
<p>bZIP</p> <p>Basic region/leucine zipper motif</p>	<p>Están implicados en la regulación de la respuesta a patógenos, a estrés abiótico y también en la señalización de la maduración de las semillas y el desarrollo de las flores. [28].</p>
<p>NAC</p> <p>(NAM,ATAF1,2, CUC2</p>	<p>Los factores de transcripción que conforman el complejo de regulación NAC incluyen microRNA y presentan además funciones de proteólisis dependientes de ubiquitina.. [29, 30].</p>
<p>MYB</p>	<p>Participan en la regulación del metabolismo secundario, regulación de la formación de meristemos y desarrollo floral y de las semillas [19].</p>
<p>MYC</p>	<p>Implicados en la señalización del ácido jasmónico. [32].</p>
<p>Cys₂His₂ Zn finger</p>	<p>Son represores de la transcripción bajo las condiciones de estrés por sequía, frío y alta salinidad [33].</p>
<p>NFY</p>	<p>Involucradas en el desarrollo embrionario y desarrollo de semillas [32].</p>
<p>WRKY</p>	<p>En plantas, estos factores de transcripción están involucrados en la respuesta a patógenos y en estrés abiótico resaltando su función en sequía y heladas. Cabe mencionar que algunos miembros de esta familia regulan la morfogénesis de tricomas. [34].</p>

Los genes de tolerancia a sequía evaluados en este proyecto pertenecen mayoritariamente a genes funcionales y metabólicos y se agrupan en diversas familias (Tabla 2).

Tabla 2: Genes de respuesta a estrés hídrico portadores de secuencias microsatelitales (SSRs) identificados en *Gmelina arborea* Roxb [1] y seleccionados en este estudio.

Descripción del Gen	ID	Función
PsbO Oxygen-evolving enhancer protein	CL3C1	La proteína PsbO es una subunidad extrínseca del fotosistema II (PS II), permite la lisis del agua en la fotosíntesis [35].
Splicing factor, arginine/serine- rich	CL214C2	Reguladores específicos del proceso de splicing, actúan en las etapas tempranas del spliceosoma. Algunos de estos factores permiten modular la selección del sitio de corte y empalme [36]
Chaperonin	CL260C2	Plegamiento y reparación de proteínas [26].
Glycerol-3- phosphate acyltransferase	CL571C2	Interviene en el control de la distribución de ácidos grasos en los cloroplastos [37].
Sodium/hydrogen exchanger Na ⁺ /H ⁺ exchanger 1 (NHX1);	CL543C1	Proteína de unión con actividad antiporte, función de transporte transmembranal de los iones de sodio. Está implicada en la respuesta a estrés osmótico. Intervienen además en el transporte de iones de litio y en el desarrollo de la hoja. Se ubica principalmente en membranas de vacuolas.
NAD-dependent malic enzyme	CL598C1	Oxidación de sustratos mitocondriales y respiración [38].
pectinesterase/pe ctinesterase	CL1054C1	Inhibe la catálisis de la hidrólisis de metil-ésteres [38].

inhibitor		
SUMO small ubiquitin- related modifie	CL430C1	Interviene en los mensajes núcleo-citosol, Se ha reportado que juegan importantes roles en la segregación cromosómica y en la división celular, además de la replicación y reparación del ADN, el importe de proteínas nucleares y en plantas regula el tiempo de la florescencia [39].
Actin- depolymerizing factor 4	CL275C2	Controla procesos de movilidad celular con base en la despolimerización de los filamentos de actina. [40].
adenosylmethioni ne synthetase	CL110C1	Precursor en la biosíntesis del etileno es requerido en la biosíntesis de los constituyentes a base de fenil-propanol que se encuentran en la pared celular [41].
Branched-chain- amino-acid aminotransferase- like protein (BCATs)	CL50C19	Juega un papel importante en el metabolismo de la leucina, isoleucina y valina, cataliza la síntesis y la degradación de estos aminoácidos. [42].
peroxiredoxin BAS1-like	CL493C1	Asegura el traspaso de electrones a partir de los residuos sulfhídricos a los peróxido, confiere a la planta la tolerancia a los hidroperóxidos alcalinos [43]
Beta-glucosidase	CL1C283	Involucrada en la síntesis del beta-glucano durante el desarrollo de la pared celular, además participa en el metabolismo de pigmentos, la maduración de los frutos y en algunos mecanismos de defensa. [43]
Metallothionein- like protein	CL691C1	Proteínas de unión a metales pesados incluyendo el cadmio, el Zinc y el níquel. Importantes en la homeóstasis y destoxificación de estos elementos. [44].
Heat shock protein 83	CL1C538	Intervienen en procesos de supervivencia y recuperación de la planta a estrés por calor además de estar implicadas en procesos específicos del desarrollo [45].

glutathione peroxidase	CL1879C1	Se expresa bajo condiciones de estrés por infección fúngica, metales y deficiencia de agua y disminuye su producción durante periodos de estrés fotooxidativo [43].
RING finger	CL421C1	Controla procesos de regulación transcripcional y traduccional; está implicada en procesos de proteólisis específica [46].
Phosphomannomutase	CL1948C1	Cataliza la conversión de manosa 1-fosfato (Man 1-P) a manosa 6-Fosfato (Man 6-P) necesarios para la síntesis de metabolitos secundarios vitales que juegan un papel crucial en la formación de polisacáridos y glicosilación de proteínas. [47].
Carbonic anhydrase	CL24C2	Catalizan la interconversión de CO_2 y HCO_3^- , importantes en el proceso de fotosíntesis [48]
Farnesylated protein	CL473C1	Regulación de procesos de señalización dependientes de GTP, regulación del ciclo celular, modificaciones de la pared celular, respuestas al estrés biótico y abiótico [49].
V-type proton ATPase subunit F	CL60C5	Tienen funciones <i>housekeeping</i> (metabolismo celular, crecimiento y homeostasis iónica). Además están involucradas en la retención de moléculas de NaCl durante la respuesta a estrés osmótico. [50].

4.1.4. Marcadores Moleculares

En los últimos años, se han desarrollado diversos marcadores moleculares asociados a genes de interés económico que permiten la selección indirecta de caracteres deseables desde las etapas iniciales del desarrollo de la planta. De esta forma, los marcadores moleculares y su uso no se ven limitados por los ciclos de vida muy largos propios de las especies arbóreas, ofreciendo una ventaja adicional para la selección de individuos. [51]. Se han convertido en una herramienta útil y común en estudios de variación genética, genética de poblaciones y filogenia [52]. Estas técnicas apoyan y aceleran programas de selección y

mejoramiento genético, permiten ahorrar energía y tiempo, minimizando los esfuerzos de mantener poblaciones enteras por varias generaciones, además de disminuir el error de confundir características que han sido generadas por el ambiente. [51].

Diferentes clases de marcadores moleculares han sido desarrolladas, sin embargo, algunas son más utilizadas debido a su eficiencia y a su óptima relación costo/beneficio, características que permiten su uso entre los cultivadores [51]. En la mayoría de los casos los marcadores desarrollados se basan en PCR (*Polymerase chain reaction* o Reacción en cadena de la polimerasa), entre los que se encuentran RAPDs, SCARs, CAPS, STS, STMS, AFLP, SSRs, entre otros [51].

Las herramientas moleculares de secuenciación de genomas y transcriptomas, que se incluyen dentro de las tecnologías de nueva generación de secuenciación (NGS: *Next Generation Sequencing*) han permitido la identificación y desarrollo de marcadores altamente polimórficos generalmente de tipo SSRs y SNP en especies para las cuales la información del genoma está disponible; en otras especies que no disponen de dicha información, las NGS pueden utilizarse para el ensamblaje del genoma por diferentes vías como la secuenciación de genotecas obtenidas en BACs o para la secuenciación de transcriptomas (RNAseq). Estas técnicas de secuenciación masiva y en paralelo son importantes en la generación de recursos genómicos, desarrollo de marcadores, análisis de perfiles de expresión, estudios de asociación genética a escala genómica (GWAs: Genome Wide Association), mapeo de QTL y genética de poblaciones [53].

4.1.5. Selección asistida por marcadores moleculares (MAS)

Como se ha mencionado anteriormente, algunos caracteres relacionados con rasgos complejos de interés pueden encontrarse asociados a marcadores genéticos o moleculares. Esta es la base de la selección asistida por marcadores moleculares (MAS) que se ha constituido en una de las herramientas más efectivas de los últimos tiempos para el mejoramiento genético en plantas [54]. La diversidad genética en un cultivo es esencial para iniciar programas de mejoramiento genético, pues a pesar de que se escogen determinados genotipos, es necesario disponer de un pool genético amplio que permita escoger características deseadas [55].

En los cultivos convencionales, para seleccionar características deseables de los individuos, los cultivadores se enfrentaban a diversos problemas: 1) Debían mantener poblaciones completas de la especie para poder seleccionar el rasgo de interés, 2) Era necesario esperar varias generaciones (F5 o F6) para poder seleccionar las características de manera correcta, (una selección en la segunda o tercera generación no presentaba los resultados óptimos esperados). 3) Era muy difícil, sino imposible, escoger rasgos de interés comercial cuando estos habían sido afectados por el ambiente. 4) A pesar de obtener las semillas de parentales con fenotipo favorable, no era posible distinguir en la etapa de semilla, aquellos individuos que en realidad presentaban las características de interés, en cambio, era necesario esperar hasta etapas adultas 5) Se presentaba la dificultad de seleccionar genes adicionales a los genes relacionados con características de interés. Así, el desarrollo de técnicas de selección por marcadores moleculares permitió ahorrar tiempo, esfuerzos y costos en los cultivos [51].

Las ventajas de la selección asistida por marcadores se resumen entonces en los siguientes ítems: 1) Evitan los riesgos de error producidos por condiciones ambientales, 2) Pueden ser examinados en juveniles antes de que los rasgos de interés se expresen, 3) Pueden ser implementados en un solo individuo o en una población, 4) Es mucho más económica que la selección por rasgos fenotípicos [55].

4.1.6. Microsatélites (SSRs: Simple sequence repeats)

Los microsatélites son secuencias repetidas en tándem, clasificados en tres familias: puros, compuestos e interrumpidos [56].

Puro: CACACACACACACA

Compuesto: CACACACACAGAGAGAGAGA

Interrumpido: CACATTCACACATTCATCA

Son de presencia frecuente en el genoma de los eucariotas, dicha frecuencia varía significativamente de un organismo a otro. En plantas se presenta en repetidas ocasiones dinucleótidos (GA)_n y (AC)_n, aunque se han encontrado en menor proporción secuencias de tri y tetranucleótidos (AAG)_n, y (AAT)_n [51]. Sin embargo estas relaciones varían en el análisis de microsatélites intragénicos; dentro de los motivos trinucleótidos, la subclase

AAG representa del 28,3 al 42,1% de los microsatélites en especies como *Arabidopsis thaliana*, sin embargo en algunas especies de cereales el motivo CCG es el más común [17]. La densidad de microsatélites en el genoma vegetal es aproximadamente un SSR cada 64,4kb en monocotiledóneas y uno cada 21,2kb para el caso de dicotiledóneas [15].

Estudios recientes han demostrado que las secuencias microsatelitales no están agrupadas en un sitio específico del genoma sino que se distribuyen uniformemente en diferentes regiones [51]. Los microsatélites son altamente utilizados debido a su reproducibilidad, heredabilidad codominancia, multiallelismo natural, alta abundancia relativa y alto porcentaje de cobertura del genoma [15]. Estos han sido utilizados en estudios de variación fenotípica y genotípica con éxito [55]. Además han sido utilizados en mapeo de QTLs (loci de carácter cuantitativo) así como en estudios filogenéticos realizados por homología.

Los microsatélites son quizá las secuencias más variables en el genoma [3], sin embargo la mayor parte de los polimorfismos de estas secuencias derivan de las diferencias en las longitudes. Además la variación genética en diversos loci microsatelitales se debe a la alta heterocigosis entre los individuos [56].

4.1.7. EST-SSRs

Los EST (*Expressed sequence tags*) son fragmentos de secuencias génicas expresadas del genoma, que se secuencian aleatoriamente a partir de genotecas de cDNA, [3]. Diversos estudios han demostrado la presencia de microsatélites localizados en regiones transcritas del genoma o secuencias génicas [17]. Los SSR como marcadores moleculares, han aumentado progresivamente su interés debido a las ventajas mencionadas anteriormente; sin embargo, desarrollar SSRs anónimos a partir del ADN genómico de una especie llega a ser costoso y no siempre útil. Por ello se ha popularizado el desarrollo de SSRs a partir de ESTs, los cuales son altamente polimórficos no sólo dentro de la especie, sino además en los taxa relacionados [36] y pueden ser utilizados en mapeos y estudios evolutivos [15]. Los EST-SSRs en comparación con los SSRs genómicos presentan ventajas como la rápida obtención mediante recursos electrónicos (bases de datos públicas de ESTs y herramientas bioinformáticas de libre acceso) y su presencia dentro de genes de interés [15].

Los SSRs génicos pueden afectar la función de los genes debido a que las expansiones o contracciones en las secuencia de ADN, afectan directamente los productos transcritos muy probablemente los productos traducidos (proteínas) y, de esta forma, pueden generarse cambios fenotípicos [17]. La localización de los EST-SSRs puede variar entre exones, 3'UTR y 5'UTR, la presencia en cada una de estas regiones pudiendo afectar la expresión génica y fenotípica. Los microsatélites ubicados en 5'UTR son requeridos para la expresión del gen, estudios han demostrado que la eliminación de estas regiones disminuye la expresión del gen en el 45% de los casos. También influyen en la adaptación de las proteínas debido a los cambios en el número de repeticiones, las cuales pueden afectar la secuencia de amino ácido en el caso de microsatelites ubicados en exones [17].

Se ha demostrado que si la secuencia microsatelital se encuentra en la región 3'UTR afecta la longitud del transcrito debido a que modifica las secuencias del codón de parada, dando como resultado secuencias más largas (incluso en términos de kb) en comparación con los tamaños esperados, lo cual influye sobre la expresión del gen [17].

4.1.8. Polimorfismos de DNA en microsatélites

Los polimorfismos de DNA, debido a las diferencias de tamaño de las secuencias microsatelitales, son detectados por medio de PCR utilizando *primers* específicos de locus microsatelitales adyacentes o también oligonucleótidos sintéticos, seguida de una electroforesis de los productos en geles de poliacrilamida para poder visualizar diferencias de bandas entre 1 a 5pb[51]. A pesar de que inicialmente la clonación y secuenciación requerida para el diseño de los primers es costosa y laboriosa, una vez obtenidos los resultados, se aumenta su relación costo/beneficio [51].

Las técnicas de tinción posteriores a la amplificación y electroforesis pueden basarse en el uso de bromuro de etidio, sin embargo, para aumentar la calidad de resolución de las imágenes se prefiere generalmente realizar geles de poliacrilamida teñidos con nitrato de plata [51]. Así, la electroforesis en geles de poliacrilamida (PAGE) desnaturizante ha permitido la visualización de las diferencias de tamaño entre los alelos [51].

4.1.9. *Slippage* y Bandas *stutter*

Algunos cambios en el tamaño de los productos de amplificación esperados para marcadores microsatelitales se deben al fenómeno denominado *slippage*, el cual ocurre tanto durante la replicación *in vivo* (responsable del polimorfismo de longitud de las repeticiones) como durante la PCR (*in vitro*), y en el que la cadena de DNA en replicación se desliga momentáneamente y se reasocia en diferentes regiones, el resultado de este proceso es una cadena con inserciones o deleciones con respecto al DNA molde, lo cual genera, *in vitro*, la presencia de bandas adicionales de tamaños diferentes en torno a al tamaño mayoritario o bandas *stutter* (Figura 1). Con lo cual, estas bandas no deben ser tenidas en cuentas en la generación de matrices genéticas a analizar.

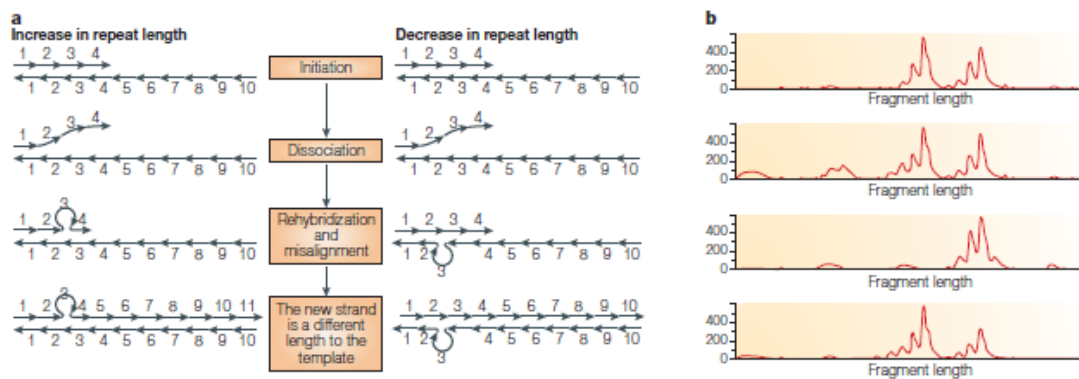


Figura 1: Fenómenos de *slippage* durante la replicación de secuencias microsatelitales, tomado de Ellegren, 2004 [57]. a. La desligación puede producir *loops* en la cadena naciente lo cual generara fragmentos de mayor peso molecular, la presencia de *loops* en la cadena molde disminuirá el tamaño de los fragmentos. b. La presencia de bandas *stutter* puede evidenciarse durante la visualización de los productos de PCR, observando picos menores en comparación con las concentraciones de las bandas microsatelitales.

4.1.10. Análisis bioinformático

Para la identificación de EST-SSRs y el diseño de los primers es necesario reconocer los patrones microsatelitales en el genoma o secuencias a evaluar, para ello se han generado diferentes programas de dominio público utilizando diferentes algoritmos como MISA (Micro SAtellite) o Sputnik, entre otros [15]. Sin embargo, a pesar de la posibilidad de

diseñar *primers* basados en ESTs, el porcentaje de éxito se estima entre un 60 y 90%, debido a diferentes razones: 1) Uno o los dos primers se encuentran en uniones exón-exón logradas durante el mecanismo de *splicing* pero que no existen a nivel de DNA genómico, 2) La presencia de largos intrones en la secuencia de DNA genómico que pueden impedir la amplificación o alterar la visualización de polimorfismos en geles de agarosa, 3) El uso de información cuestionable en el diseño de los primers, 4) Los primers se desarrollaron a partir de cDNA quiméricos [15].

Para el análisis de los perfiles de electroforesis de marcadores microsatelitales es necesario realizar regresiones logarítmicas utilizando el marcador de peso molecular para establecer los pesos de los productos amplificados con mayor precisión [58]. Una vez realizados estos procedimientos debe procederse al análisis de los datos obteniendo número de alelos por locus, porcentaje de loci polimórficos, heterocigosis observada y esperada, análisis de equilibrio de Hardy-Weinberg si la población cumple las condiciones de apareamiento al azar, ausencia de selección natural o artificial y ausencia de flujo de genes. También se evalúan las distancias génicas entre otros posibles análisis que ofrecen software libres como lo son PopGen, Arlequín, TFGPA, GenAlEx y GDA [59].

4.2. ANTECEDENTES

A partir de 1964, la FAO estableció como prioridad mejorar los métodos empleados para la utilización y conservación de *G. arborea* debido a que estaba catalogada, junto con otras especies arbóreas, como especie promisoría por sus características como rápido crecimiento, fácil establecimiento en diferentes ambientes y las propiedades industriales de su madera [60]. Diferentes estudios empezaron a interesarse en la adaptabilidad de la especie en respuesta a su introducción en países africanos [60]. A partir de 1980, en América latina, el Centro Agronómico de Investigación y Enseñanza (CATIE), inició planes de mejoramiento genético con el apoyo del Instituto tecnológico de Costa Rica, evaluando la altura de los individuos y el DAP [18]. Desde entonces diferentes técnicas biotecnológicas han sido aplicadas para el mejoramiento de esta especie a nivel mundial, aunque en forma aún muy incipiente, incluyendo cultivo de tejidos *in vitro*, estudios de

diversidad genética en poblaciones clonales a partir de marcadores moleculares, y primeros intentos de transformación genética [2].

Sin embargo, se han visto varios factores que pueden afectar el grado de mejoramiento y el uso de esta especie, como lo son la susceptibilidad de melina a los insectos y otras enfermedades, el alto costo de la mano de obra, la baja densidad de la madera de los árboles en época de tala (8 años aproximadamente) con respecto a especies de eucalipto y acacia, y por supuesto, la sensibilidad de esta especie a los cambios medioambientales que se vienen presentando [22]. El principal elemento requerido para solucionar estos problemas, es el acceso a una amplia base genética y el desarrollo de marcadores moleculares destinados a apoyar programas de mejoramiento genético [22].

Con base en lo anterior, la Unidad de Biotecnología Vegetal de la PUJ, en asocio con la empresa Pizano S.A., desarrolló un programa de investigación encaminado al desarrollo de un programa de mejoramiento genético de melina, mediante un primer proyecto de identificación de variabilidad genética en una colección de 33 clones élite. En este proyecto se logró evidenciar la presencia de 5 clusters, claramente diferenciados en dos grupos principales distantes genéticamente (figura 2).

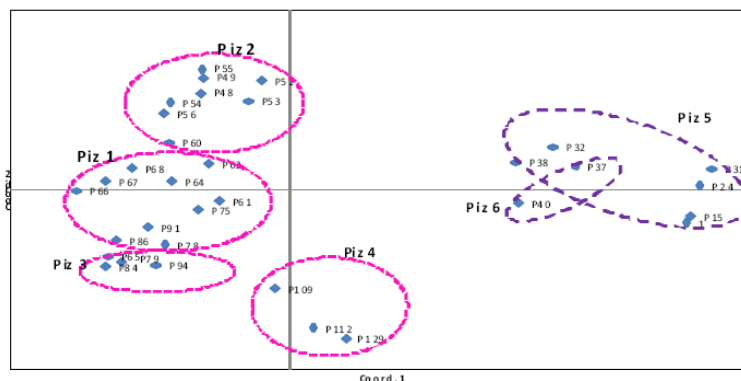


Figura 2 Análisis de la variabilidad genética en la población de 33 clones de *Gmelina arborea* de la colección de Pizano S.A. [2]

Un segundo proyecto buscó generar el primer recurso genómico para la especie por medio de secuenciación de última generación del transcriptoma (RNAseq) de hojas y raíces de plantas del clon 66, (escogido por presentar las mejores respuestas a condiciones de sequía)

sometidas a estrés hídrico [1]. El proyecto fue realizado debido a que la sequía se presenta como una de las principales limitaciones para la distribución de la especie en los municipios de la región Caribe donde se encuentran las plantaciones clonales de la empresa Pizano S.A. [1] [2]. Se logró generar un primer catálogo de genes representado por 18913 unigenes, de los cuales un 60% presentaron homología mediante Blast con secuencias génicas de especies como *Arabidopsis thaliana*, *Vitis vinífera*, *Populus trichocarpa* y *Ricinus communis* [1] (W. Terán, comunicación personal). Se identificaron en este proyecto un gran número de genes relacionados con la respuesta a estrés hídrico, desde osmosensores, genes de transducción de señales, factores de transcripción y proteínas efectoras. Asimismo, se logró identificar en estas secuencias génicas, un total de 629 motivos microsatelitales (EST-SSRs), permitiendo el diseño de 344 parejas de *primers*, entre las que se escogió, para este estudio, un primer juego de 28 *primers* por tener anotación bibliográfica que daba cuenta de su importancia en la respuesta a estrés hídrico [1]. Así, este trabajo, si bien arroja resultados muy preliminares, representa el primer estudio de validación de secuencias microsatelites tipo EST-SSRs para el género *Gmelina*.

Objetivo general

Analizar y validar secuencias microsatelitales polimórficas asociadas a genes expresados bajo condiciones de estrés hídrico en una colección de 33 clones élite de melina (*G. arborea* Roxb).

Objetivos específicos

- Identificar la ubicación de los EST-SSRs respecto a la secuencia codificante.
- Estandarizar la amplificación de ADN microsatelital para 28 EST-SSRs seleccionados.
- Determinar y caracterizar el polimorfismo de los 28 EST-SSRs en la colección de 33 clones élite de *melina*..
- Obtener las distancias genéticas entre los 33 clones élite de la colección por medio de los EST-SSRs polimórficos validados.

6. Metodología

Métodos por objetivo

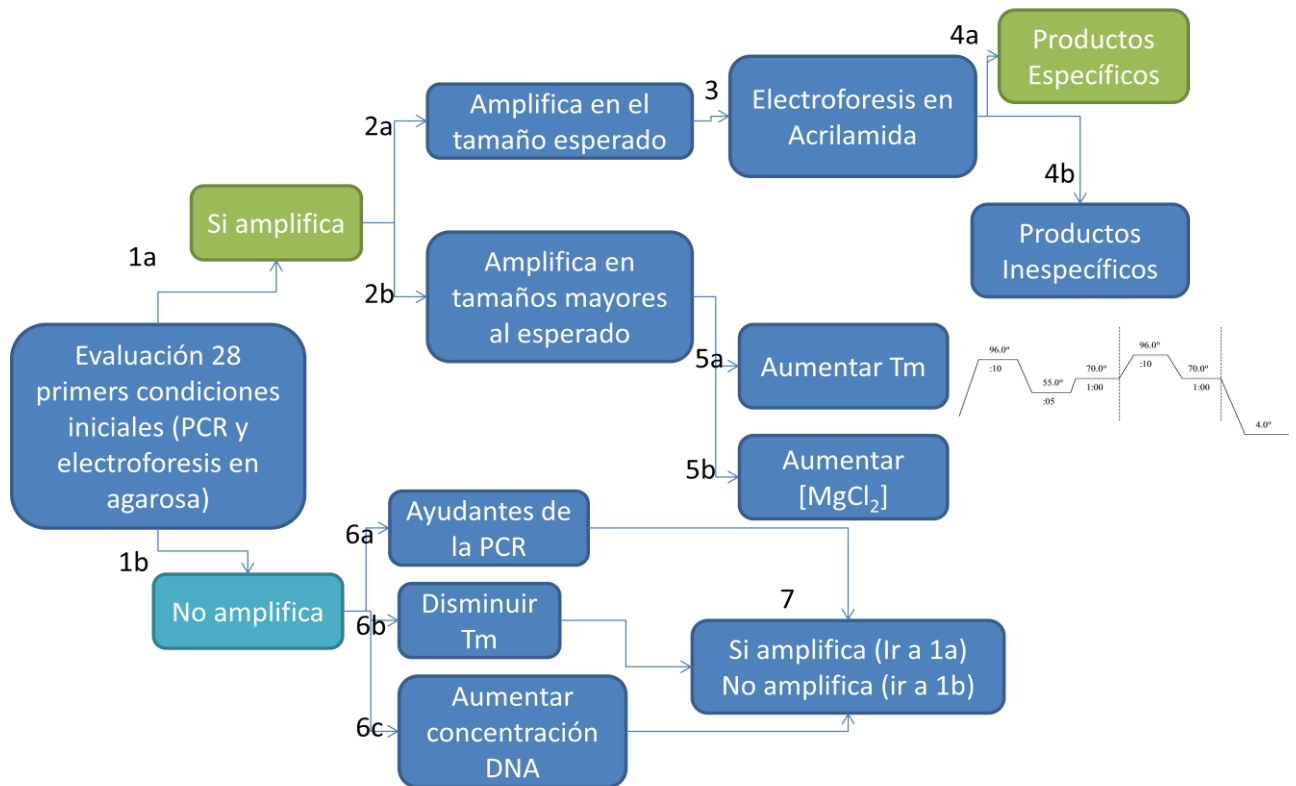
6.1. Analisis de la ubicación de los EST-SSRs respecto a la secuencia codificante.

Se anotó la ubicación de los microsatélites respecto a la secuencia codificante utilizando la base de datos GmelinaDB (www.gmelinadb.co) y se realizó un análisis de BLAST-X para comparar las ubicaciones localizando el microsatélite en las regiones: 5'UTR, 3'UTR y CDS

6.2. Estandarización de la amplificación del ADN microsatelital

Se extrajo ADN genómico a partir de individuos de melina pertenecientes a los 33 clones élite del banco de germoplasma de la empresa Pizano (S.A) mantenido en las instalaciones de Monterrey Forestal Ltda. Se utilizó el protocolo de extracción de ADN estandarizado por Solano (2010) [68]. Posteriormente se realizó la evaluación de la pureza y cuantificación del ADN extraído por medio de la medición de las absorbancias a 260 nm, 280 nm y 230 nm, utilizando un espectrofotómetro Nanodrop 1000. Para la evaluación de la integridad del ADN se obtuvieron perfiles de electroforesis en gel de agarosa al 0.8% (p/v). [61].

Una vez verificadas la calidad y pureza de los ADNs extraídos se procedió a la estandarización de la amplificación de los microsatelites siguiendo el siguiente flujo de trabajo:



Se obtuvieron productos de amplificación por PCR en 9 individuos de la población clonal genéticamente más diversos utilizando las 28 parejas de primers de secuencias microsatelitales diseñadas automáticamente a partir de los EST-SSRs identificados.

La composición PCR por reacción utilizada fue: Buffer 1X (Invitrogen); MgCl₂ 2,0 mM; dNTPs 0,2 mM; Primer 0,25 μM; Taq 0,5 unidades; DNA 200 ng; en un volumen final de 20μl.

Programa PCR: 95° por 5 minutos, 95° por 30 segundos, Tm por 30 segundos (Tabla 3), 72° por 30 segundos. 72° por 5 minutos. 32 ciclos

Posteriormente se evaluaron los resultados de la PCR a través de electroforesis en gel de agarosa al 2.0 % (p/v) y a continuación se realizaron modificaciones de la temperatura de hibridación y/o concentración de reactivos con aquellos *primers* que no hayan permitido amplificación o que amplifiquen productos inespecíficos.

En algunos casos se purificaron las bandas de geles de agarosa con el tamaño esperado y se reamplificaron para el posterior análisis en geles de poliacrilamida al 6% (p/v). Finalmente

se aplicó el protocolo estandarizado al resto de clones de la colección a analizar (33 clones) analizando los productos de PCR en geles de electroforesis de poliacrilamida.

6.3. Identificación de polimorfismos

El análisis de polimorfismo de los microsatélites amplificados por PCR se realizó por medio de electroforesis en geles de poliacrilamida (PAGE) seguida de tinción con nitrato de plata (Anexo 2) para cada uno de los loci microsatelitales. Solo se prosiguió con el análisis estadístico para los microsatélites que presentaron mayor cantidad de polimorfismos.

6.4. Análisis de las secuencias polimórficas y validación de microsatélites como marcadores moleculares

Se construyeron matrices de presencia/ausencia a partir de los perfiles de electroforesis PAGE. Para obtener el peso exacto de los fragmentos se realizaron regresiones logarítmicas a partir de las distancias de migración confrontadas las distancias de migración obtenidas para patrones de peso molecular conocidos (10 bp DNA ladder, invitrogen). La estadística para cada microsatélite polimórfico fue analizada utilizando el software GenAlEx (Índices y valores Anexo 3) y por último se validaron las secuencias microsatelitales.

7. RESULTADOS

7.1. Caracterización de secuencias y ubicación de secuencias microsatelitales respecto al gen.

A partir de la identificación de los *primers* por medio de la base de datos gmelinaDB (gmelinadb.co) y ubicación de los microsatélites en las secuencias génicas por BLAST-X, se obtuvieron los resultados presentados en la Tabla 3.

Tabla 3: Identificación de la ubicación de las secuencias microsatelitales y condiciones de amplificación utilizadas

Identificación	Nº Accesión	Región	[MgCl ₂]	<i>T_m</i>	Motivo
CL3C1	Q40459	5'UTR	2.0	64	(AG)12
CL214C2	Q9R0U0	5'UTR	1.5	53	(CT)9
CL260C2	O65282	CDS	1.5	55	(AG)11
CL571C2	Q43307	3-UTR	2.0	54	(AC)10
CL543C1	Q68KI4	CDS	1.5	51	(TG)9
CL598C1	P37225	CDS	2.3	56	(AG)10
CL298C6	Q8GYN5	5'UTR	1.5	57	(CT)8
CL50C19	Q9ASR4	3-UTR	1.5	52	(AC)12
CL1054C1	P17407	CDS	2.0	55.5	(AT)9
CL691C1	P43390	CDS	2.0	53	(AT)8
CL1948C1	Q57842	CDS	1.6	54	(AG)17
CL24C2	P16016	5'UTR	1.5	53	(AT)10
CL60C5	Q9ZQX4	CDS	1.5	54	(GA)9
CL430C1	Q28CQ4	5'UTR	1.6	55	(AT)8
CL897C1	O23052	5'UTR	1.6	52	(AT)9
CL275C2	Q9ZSK3	5'UTR	1.5	52	(CT)8
CL13C5	Q6ICZ8	5'UTR	1.5	55	(AC)8
CL473C2	A5C6S4/Q 8LGG1	5'UTR	2,0	62	(AC)11
CL110C1	Q96552	5'UTR	1.5	54	(TC)8
CL493C1	Q9C5R8	CDS	2,0	62	(TC)13
CL1C283	Q9M1C9	CDS	1.8	53	(AT)9
CL1C364	Q8W378 / B6GVA2	CDS	1.6	52	(AC)9
CL1C538	P51819	5'UTR	2,0	56	(CA)12
CL1C853	Q9LQ55	CDS	2.6	52	(CT)10
CL1C1084	Q38850 / P27898 / P10290	5'UTR	1.5	55	(TC)11
CL1C1342	P51819	5'UTR	1.5	53	(GT)12
CL1879C1	Q8L910	CDS	1.5	53	(AC)12
CL421C1	B6T4N6	5'UTR	1.8	53	(AG)10

*Por motivos de confidencialidad no se presentan las secuencias de los primers

7.2. Estandarización de la amplificación y visualización de ADN microsatelital

Para la estandarización de la amplificación del ADN microsatelital se extrajo ADN vegetal (Protocolo de Stefenon *et al.* 2004, utilizado por Solano 2010) [2]. Las concentraciones y valores de absorbancia a 260, 280 y 230 nm se presentan en el Anexo 1. Además, se comprobó la integridad del ADN por medio de electroforesis en gel de agarosa al 0,8% (ver anexo 1). Se decidió realizar las reacciones de PCR con 2µl de ADN con concentración de 100 ng/µl debido a que con concentraciones menores no se obtuvo producto de amplificación. Se realizaron dos extracciones independientes para cada clon, obteniendo iguales patrones de bandas en los perfiles de electroforesis con las parejas de primers evaluadas.

De acuerdo con los trabajos previos de diversidad genética en la misma colección de clones de *Gmelina arborea*, se escogieron 9 clones que presentan mayor distancia genética: 15, 31, 40, 53, 56, 66, 78, 94 y 112, como subpoblación con fines de estandarización de la PCR, tomando el clon 66 como control positivo debido a que fue el clon cuyo transcriptoma fue secuenciado. Para realizar las reacciones de PCR con las 28 parejas de primers se tomaron los valores iniciales en cuanto a las temperaturas de hibridación y se utilizó una concentración de MgCl₂ de 1,5 mM, sin embargo no se obtuvo producto de amplificación con todos los *primers* y en algunos casos se presentaron productos inespecíficos de más de 1000 pb (Figura 3). En estos casos, no se evidenciaron bandas con el tamaño molecular esperado; teniendo en cuenta que los valores de peso molecular para los productos de amplificación fluctúan entre 150 y 400 pb, sólo se escogieron 13 parejas de *primers* que presentaron un solo producto de amplificación en el rango de tamaño esperado.

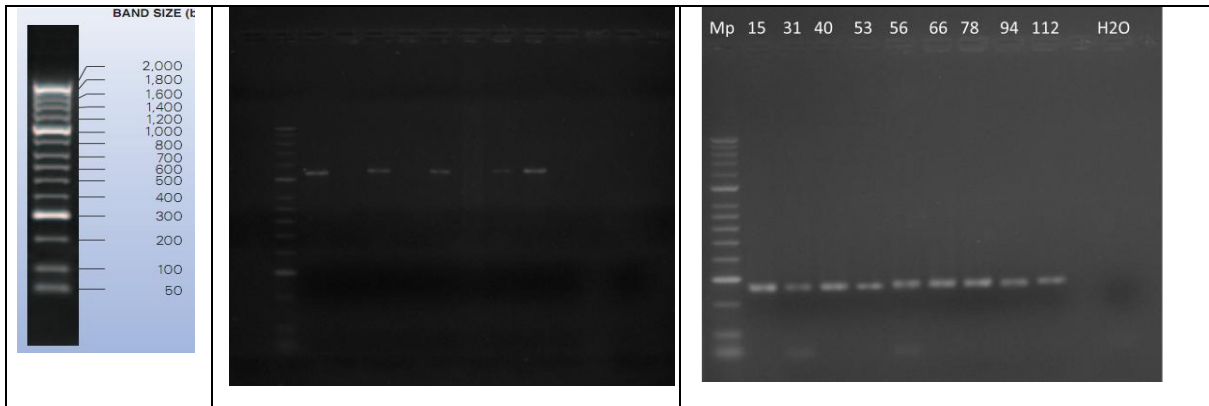


Figura 3: Patrones gels de electroforesis en agarosa 2% V=90 Tiempo de corrida 60 minutos. 3a) Marcador de peso molecular HypperLadder II 3b) Primer CL260C2, bandas amplificadas entre 1000 y 1200pb 3c) Primer CL571C2 bandas amplificadas entre 200 y 300 pb.

Debido a la presencia de productos inespecíficos se realizaron modificaciones de la temperatura de hibridación y la concentración de $MgCl_2$ en gradientes desde 1,6 hasta 2,9 mM para algunas parejas. Paralelamente se realizó la búsqueda en la base de datos gmelinadb.co para comprobar la presencia del microsatélite en las secuencias FASTA y la ausencia de gaps entre las secuencias microsatelitales y las secuencias de los primers.

Las modificaciones con las nuevas condiciones de amplificación se comprobaron nuevamente mediante PAGE.

El aumento de la temperatura de hibridación para muchos *primers* especificó productos de mayor tamaño molecular en comparación con el esperado. En otras ocasiones se observan más de dos bandas aún a pesar del aumento de la temperatura de hibridación (Figura 4a) lo cual no permitió identificar su homocigosis o heterocigosis. En la mayoría de los individuos se presentaron bandas sombra o *stutter* con tamaños moleculares menores a los esperados (figura 4b).

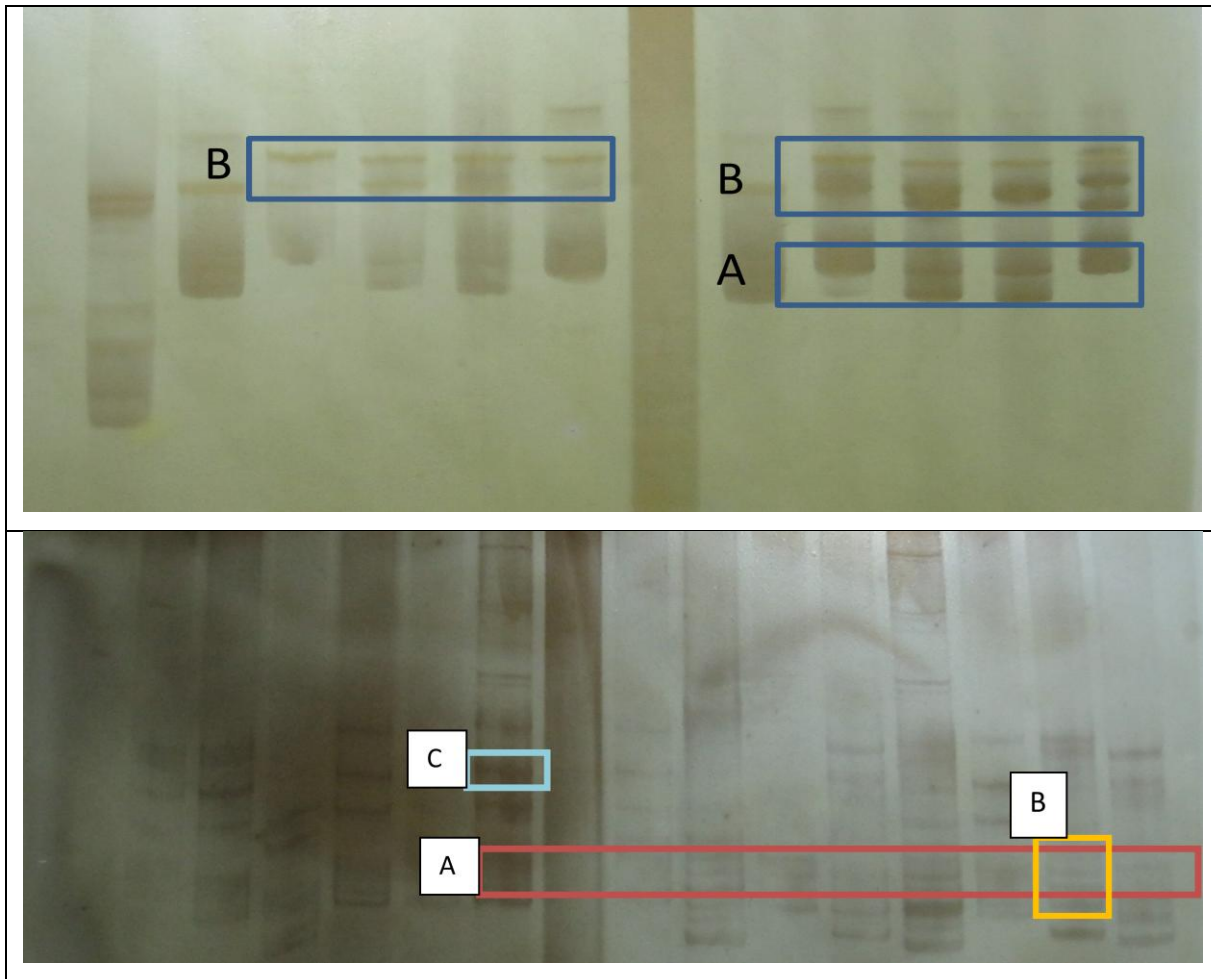


Figura 4: Análisis electroforético PAGE de los microsatélites amplificados. A) Locus CL3C1 a) Tamaño esperado, alta intensidad b) Bandas por encima del tamaño esperado, igual intensidad. B) Locus CL3C1. a) Rango alélico esperado, b) Bandas *stutter*, c) Bandas más intensas con mayor peso molecular.

En algunos casos, se logró la estandarización de 5 clones, pero al aplicar las nuevas condiciones de PCR a la población completa se observó la presencia de productos inespecíficos especialmente entre los clones 75 a 129.

La purificación de las bandas a partir de geles de poliacrilamida productos del clon 66 utilizando el primer CL473C1 resultó siempre en 4 bandas de igual intensidad con tallas moleculares entre 120 y 180pb, lo cual dificultó el análisis de los polimorfismos.

7.3. Identificación de polimorfismos

Para las parejas de *primers* CL473C2, CL493C1, y CL1054 se obtuvieron perfiles de electroforesis propios de secuencias microsatelitales, que permitieron la generación de matrices, y para lo cual sólo se tomaron las bandas más intensas (una a dos bandas por locus e individuo). La medición del tamaño de los fragmentos amplificados se realizó utilizando regresiones logarítmicas de las distancias de migración de patrones de peso molecular conocido y utilizando curvas estándar y ecuaciones de regresión con valores R^2 superiores a 0.99.

Se presentaron 9 alelos por cada locus analizado. Se evaluó cada secuencia (Tabla 4) donde se ubicaron las secuencias microsatelitales buscadas y algunas secuencias dinucleotídicas adicionales de 4 a 6 repeticiones.

Tabla 4: Resultados productos amplificados EST-SSRs

Marcador	Motivo	Tamaño esperado	Tamaño observado (Rango alélico)	Tm	Nº de Alelos	Genes relacionados
CL473C2	(AC)11	219	204 a 219 No presente en toda la población 227-250 Bandas más intensas 257-265 No presente en toda la población	60.5	9	Farnesylated protein
Rango alélico: 227, 230, 234, 240, 242, 248, 255, 259, 263						
CL493C1	(TC)13	217	196 – 217 Se presenta en los individuos 1 a 68 225-253 Bandas más intensas presentes en toda la población 260-287 Presentes en algunos individuos de la población	62.5	9	2-Cys peroxiredoxin BAS1-like, chloroplastic
Rango alélico: 229, 234, 240, 243, 249, 253, 256, 260 y 263						

CL1054	(AT)9	240	Las bandas del tamaño esperado se presentan en todos los clones con la mayor intensidad	55. 5	9	pectinesterase/pectinesterase inhibitor
Rango alélico: 222, 227, 229, 232, 236, 240, 243, 246 y 255						

*Por motivos de confidencialidad no se presentan las secuencias de los contigs

Para las parejas de primers CL473C2 y CL493C1 se presentaron bandas más intensas de pesos mayores al esperado, se intentó realizar purificación de cada una de las bandas en el clon 66 para comprobar la presencia del microsatélite en tamaños más altos, sin embargo la reamplificación reveló un patrón de 4 bandas de pesos de 258, 230, 220 y 208pb idéntico al presentado por primera vez. Se descartó contaminación debido a que la reacción de PCR con agua da resultados negativos, por lo cual se atribuyeron estas bandas a fenómenos de slippage in vitro, imposibilitando el envío para secuenciación contemplado.

7.4. Análisis de secuencias polimórficas y validación de microsatélites

Para los loci positivos CL473C2 y CL493C1, se realizó el análisis con el programa TFPGA de distancias genéticas de Nei (1972) distancias que permitieron generar un dendrograma (Figura 5), y una agrupación de los clones de la colección en 7 *clusters* diferenciados. Adicionalmente, los valores para alelos efectivos, índice de información, heterocigosis observada y esperada y valores F se calcularon con el mismo programa (Tabla 5).

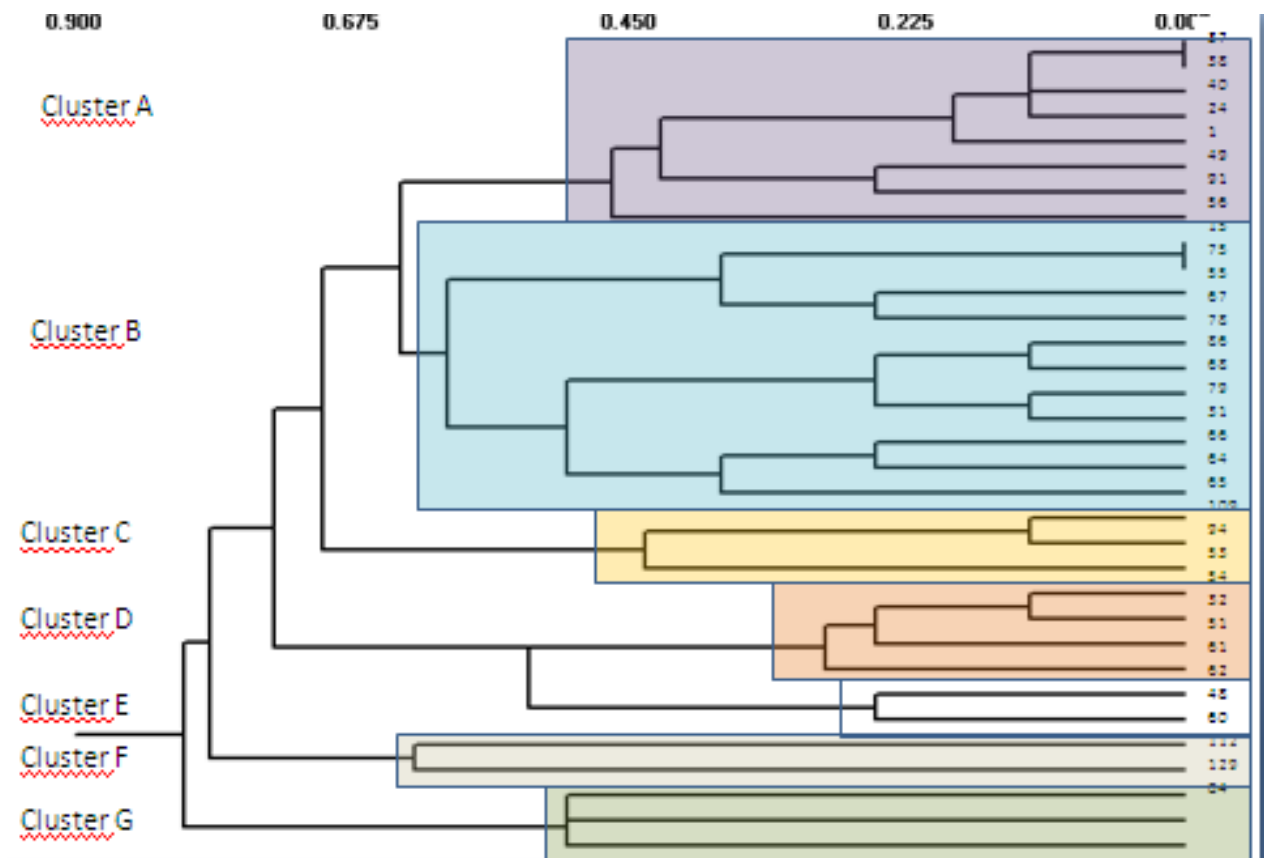


Figura 5: Dendrograma basado en distancias genéticas Nei (1972)

Tabla 5: Clusters generados a partir del análisis de UPGMA con distancias mínimas de Nei (1972)

Cluster	Individuos	Distancia
A	1,3,6,7,8,10,15,29	0.4643
B	2,4,14,19,21,22,23,24,25,26,28	0.5982
C	20,30,31	0.4375
D	5,11,12,13	0.2917
E	17,18	0.25
F	9,16	0.62
G	27,32,33	0.5

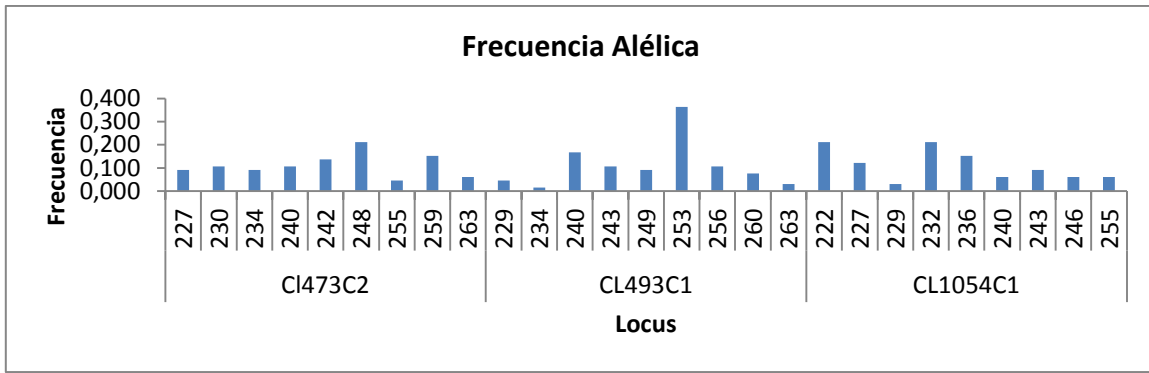


Figura 6: Frecuencias alélicas observadas

Tabla 6: Índices de heterocigosis y polimorfismo. (N: tamaño población, Na: número de alelos por locus, Ne: Alelos efectivos, Ho: Heterocigosis observada, He: Heterosigosis esperada, UHe: Heterocigosis Imparcial. F: Índice de Fijación

Locus	N	Na	Ne	Ho	He	UHe	F
CL473C2	33	9.000	7.615	0.242	0.869	0.882	0.721
CL493C1	33	9.000	5.007	0.485	0.800	0.813	0.394
CL1054C1	33	9.000	6.764	0.000	0.852	0.865	1.000

Para cada pareja de primers tomadas en el análisis como locus independientes los siguientes valores de polimorfismo y promedio de heterocigosis.

CL473C2 % loci polimórficos: 100.0000

CL493C1% loci polimórficos (criterio 99%): 100.0000

CL1054C1% loci polimórficos (Criterio 95%): 100.0000

Promedio Heterocigosis: 0.8404

Promedio Heterocigosis (UHe): 0.8533

Promedio Heterocigosis (cuenta directa): 0.2424

8. ANÁLISIS DE RESULTADOS

8.1 Estandarización amplificación y visualización en electroforesis

La cantidad de DNA recomendado en la amplificación por PCR para el análisis de secuencias microsatelitales es alrededor de 20 a 50 ng por reacción [20], sin embargo, solo a partir de cantidades de 100ng se observaron bandas tenues, el aumento de la concentración del DNA permitió una mayor amplificación y por lo tanto la visualización de bandas, pero se pudo correr el riesgo de aumentar la presencia de productos inespecíficos y bandas *stutter* que se observaron con frecuencia en geles de poliacrilamida. [12].

Aún a pesar del aumento de la concentración de DNA y adición de ayudantes a la PCR, 15 de las 28 parejas de *primers* no permitieron obtener amplificación. Las diversas posibles causas han sido esbozadas previamente (ver marco teórico) tratándose de EST-SSRs, entre ellas la posibilidad de que los primers diseñados se hibriden en unione exón-exón del cDNA, no pudiéndose hibridar a ADN genómico.

La frecuencia de variaciones intraclonales en plantas de melina de la colección de Pizano SA es relativamente alta [2], incluso otros autores afirman que es posible encontrar variación genética entre distintos órganos o tejidos de un mismo individuo [62]. Sin embargo los patrones en electroforesis en geles de agarosa y poliacrilamida utilizando ADN de dos extracciones de diferentes hojas del mismo individuo son consistentes, indicando la ausencia de variabilidad genética dentro de un mismo individuo, tal como fue igualmente corroborado por Solano (2010) [2].

Es necesario tener en cuenta que debido a la utilización de marcadores microsatelitales dinucleótidos se aumenta la probabilidad de presentarse artefactos de la PCR como *slippage* generando bandas *stutter* [57]. Este fenómeno aumenta su probabilidad en la amplificación de repeticiones de los dinucleótidos CA y CT. Se ha observado que una extensión mínima de 5 a 6 repeticiones es requerida para que ocurra *slippage* y por ello es posible que pueda presentarse incluso en replicas de la PCR en un mismo individuo [63].

Los resultados en PAGE para el clon 66 evidenciaron bandas más intensas con mayores pesos moleculares en comparación con las bandas esperadas; a pesar de la revisión de la

base de datos de gmelina (gmelinadb.co) en la que no se encontraron anomalías en las secuencias ensambladas, es posible que existan diferencias en el número de repeticiones para los microsatélites evaluados o para microsatélites adyacentes, debido a la posibilidad de existencia de variabilidad intraclonal.

La presencia de bandas de 30 a 60 pb adicionales al peso molecular esperado podría también explicarse por la existencia de regiones afines a los *primers* presentes en otras regiones del genoma, de la misma forma la hipótesis podrá verificarse por medio de la secuenciación de la región de ADN genómico completa de la especie de interés. [17].

Una mayor cantidad de bandas en los últimos individuos de la colección (a partir del clon 75) en comparación de los primeros puede deberse a que estos clones parecen ser muy distanciados genéticamente con respecto a los primeros, como fue evidenciado mediante análisis de diversidad genética usando AFLP's [2]. Sin embargo se ha observado que los resultados de la tinción con nitrato de plata en los geles de poliacrilamida, no son completamente reproducibles para algunos de los microsatélites evaluados, obteniéndose algunas diferencias entre las dos réplicas realizadas para toda la población.

A pesar de que los marcadores moleculares tipo SSR presentan una alta confiabilidad en comparación con otros marcadores como AFLP's, se debe tener en cuenta que solo deben tomarse las bandas más intensas [4], debido a que las bandas menos visibles corresponden generalmente a artefactos de la PCR productos del *slippage* o de mutaciones, generando típicamente resultados de tipo microvariantes, patrones trialelicos o alelos marginados que dificultan el análisis [64].

8.2 Análisis de Polimorfismos en las secuencias microsatelitales

Los polimorfismos en las secuencias microsatelitales pueden presentarse por inserciones o deleciones durante el proceso de replicación. Debido a que el estudio se realizó con secuencias de dinucleótidos se esperaría que el rango alélico de cada primer sea en orden de +2 ó -2, sin embargo la presencia de alelos +5, +7 puede indicar la no pertenencia de los alelos al mismo locus. Aún así, el proceso de medición de los alelos, por tratarse de una regresión logarítmica y extrapolación de los datos, puede generar errores en los tamaños exactos de las bandas. Los eventos de mutaciones en otras regiones de la secuencia en

estudio que también presentan repeticiones de dinucleótidos, pueden igualmente ser responsables de los resultados obtenidos. Se observó que en algunos casos las diferencias entre un alelo y otro eran de ± 4 ó ± 6 lo que corresponde a alelos faltantes dentro de la población pero que pueden existir en una población mayor. Es necesario tener en cuenta que debido al proceso de selección artificial al que han sido sometidas las plantaciones de la empresa Pizano, es posible que no se puedan encontrar los alelos faltantes aún en tamaños poblacionales mayores.

El análisis de Clusters para los 33 individuos de la colección de Pizano S.A. está basado en 2 loci microsatelitales polimórficos heterocigotos que fueron utilizados para medir las distancias genéticas por medio del algoritmo de distancias mínimas de Nei (1972) (figura 5), el cual permite distinguir 7 grupos principales. El análisis se realizó solamente con dos loci debido al exceso de homocigosis en el tercer locus, además no se distinguían claramente los clusters y se presentaban individuos compartidos. La procedencia de los clones de este estudio es desconocida, por lo cual no es posible relacionar las distancias genéticas con datos geográficos, tampoco se cuenta con información sólida acerca de características fenotípicas claras con respecto a tolerancia a estrés hídrico o sequía, con lo cual es muy preliminar poder inferir cualquier asociación. Si bien, los datos obtenidos a partir de microsatélites no corresponden a las matrices obtenidas en el estudio de diversidad genética con AFLP's debido a la naturaleza diferente de los marcadores y el número reducido de loci empleados, los altos valores de distancia reflejados en el dendrograma demuestran una alta diversidad genética a nivel de clones para estas dos secuencias génicas asociadas con una respuesta a estrés hídrico, posiblemente explicada por la amplia distribución de la especie en diferentes regiones y condiciones climáticas [2], y que podrían ser causa de variaciones fenotípicas con respecto a la actividad de los genes correspondientes.

Las secuencias microsatelitales amplificadas con los primers CL473C2, CL493C1 y CL1054C1 son polimórficas, debido a que presentan más de dos alelos y el alelo más común tiene una frecuencia menor a 0,95 (Figura 6) [65]. Además para los locus CL493C1 y CL1054C1 la posibilidad de polimorfismos funcionales o fenotípicos se ve favorecida por encontrarse dentro de las secuencias codificantes (CDS) (Tabla 3) [66]. Sin embargo la figura demuestra un comportamiento monomodal esperado en el locus CL473C2 y

comportamientos bimodales no esperados en los locus CL493C1 y CL1054C1, estos comportamientos se deben a la composición de la población estudiada, la cual no se encuentra en equilibrio de Hardy-Weinberg debido a que no cumplen con las condiciones de una población natural o de una población con parentales comunes: son individuos seleccionados artificialmente en regiones muy distantes geográficamente, y reproducidos de forma asexual. Por ello, se debe también las diferencias entre la heterocigosis observada y esperada que se presenta en la Tabla 3: a pesar de que se presenta un nivel de H_e alto (0,840) que demuestra la variabilidad genética en la población, esta medida puede tener un sesgo debido a que se analiza un número menor a 30 loci [67], además los valores de heterocigosis para cada locus son menores a 0.5, y en el locus CL1054C1 no se presentan individuos heterocigotos lo cual disminuye el nivel de polimorfismo. El exceso de homocigotos es una consecuencia de alelos nulos en donde las mutaciones pueden haber ocurrido en regiones de unión de los *primers* inhibiendo la amplificación [68]. Otra explicación para la ausencia de heterocigotos es el efecto Wahlund, en donde se disminuye la heterocigosis de una población por causa de la estructura de las subpoblaciones [68]. También los valores de fijación (Tabla 5), índice relacionado con la homocigosis de la población cuyo rango se distribuye entre 0 a 1 [69], presentan altos niveles lo cual representa un menor porcentaje de polimorfismo.

A pesar de los valores de heterocigosis y fijación, los 3 loci presentan altos niveles de polimorfismos, especialmente el locus CL493C1 el cual presentó el mayor nivel de polimorfismo, además de alta heterocigosis. Si bien los resultados de este trabajo arrojan datos preliminares sobre un set limitado de microsatélites EST-SSRs, logran demostrar que pueden ser una fuente valiosa de genes candidatos a convertirse en marcadores moleculares asociados a respuesta a sequía en futuros programas de selección asistida.

9. CONCLUSIONES

Se obtuvieron 13 parejas de *primers* que permiten la amplificación de secuencias microsatelitales en el rango de tamaños esperados dentro de la población.

Se localizaron las secuencias microsatelitales en las regiones 5'UTR, 3'UTR y regiones CDS con una menor frecuencia en la región 3'UTR.

Se logró la estandarización de las condiciones de PCR para 3 de las parejas de *primers* trabajadas obteniendo que el locus microsatelital CL493C1 es altamente polimórfico y presenta el mayor nivel de heterocigosis en el estudio. Las parejas de *primers* CL473C2, y CL1054C1 son polimórficas debido a que presentan 9 alelos por locus y las frecuencias alélicas corresponden a la definición de polimorfismo [65]. Sin embargo presentan un nivel de heterocigosis bajo o nulo, probablemente debido a que algunos alelos no tuvieron representación en la población.

10. RECOMENDACIONES

Es necesario utilizar cámaras de electroforesis vertical que permitan el análisis de la población completa (33 individuos) en un solo vidrio y de esta forma garantizar las mismas condiciones de migración/ tinción para todas las muestras en forma homogénea.

Una posibilidad para disminuir las bandas inespecíficas consiste en realizar PCR solamente del segmento amplificado en el peso molecular esperado, por tanto debe realizarse la purificación del DNA a partir de los geles de agarosa. Como esta amplificación no asegura que el segmento de dicho peso molecular corresponda a la secuencia microsatelital esperada, también se hace necesario realizar una secuenciación de las bandas obtenidas.

Debido a que la causa del alto número de bandas puede ser la utilización de más 200 ng de ADN molde en las reacciones de PCR, puede realizarse reamplificación a partir del producto de PCR obtenido con 20 ng de ADN molde

Para comprobar la presencia del microsatélite en las bandas de mayor peso molecular y mayor intensidad se recomienda realizar clonación de la purificación para posteriores procesos de secuenciación.

Se recomienda construir *ladders* alélicos para las secuencias de primers CL473C2, CL493C1 y CL1054C1 con el fin de comprobar los tamaños reales de las bandas obtenidas.

Finalmente, aumentar el tamaño poblacional puede permitir incluir alelos producto de deleciones o inserciones de dinucleótidos que no se presentan en este estudio.

10. BIBLIOGRAFÍA

1. ROSERO, C. ARGOUT, X. RUIZ, M. & TERÁN, W. A drought stress transcriptome profiling as the first genomic resources for white teak –Gamhar- (*Gmelina arborea* Roxb) and related species. BMC Proceedings 5 (Suppl 7): P178
2. SOLANO, G. Caracterización molecular de clones de *Gmelina arborea* mediante el uso de marcadores moleculares AFLP's (Amplified fragment length polymorphism). Tesis de grado. (2010)
3. XUE, Z. WEI, C. YINGPENG, H. & WENBIN, L. Review for the methods of developing SSR Molecular Markers. Journal of Northeast Agricultural University (2008) 15(1): 62-66
4. WESING, K. NYMBOM, H. WOLFF, K & KAHL, G. DNA fingerprinting in plants. Principles, Methods and Applications. Second Edition. CRC Press – Taylor and Francis Group (2005)
5. ASHRAF, M. Inducing drought tolerance in plants: Recent advances. Biotechnology Advances 2010. 28 169–183
6. SALEKDEH, G. REYNOLDS, M. BENNETT, J & BOYER, J. Conceptual framework for drought phenotyping during molecular breeding. Trends in Plant Science (2009) 14(9): 488-496
7. SEKI, M. UMEZAWA, T. URANO, K. SHINOZAKI, K. Regulatory metabolic networks in drought stress responses. Plant Biology (2007) No 10, Pp 296 - 302
8. CHAVES, M. MAROCO, J. PEREIRA, J. Understanding plant responses to drought- From gene to the whole plant. Functional Plant Biology. (2003), 30:229-264
9. OGBONNAYA, C. NWALOZIE, M & NWAIGBO, L. Growth and wood properties of *Gmelina arborea* (Verbenaceae) seedlings grown under five soils moisture regimes. American Journal of Botany 79(2). 1992 Pp 128-132

10. CATTIVELLI, L. *et al.* Drought tolerance improvement in crop plants: An integrated view from breeding to genomics. *Field Crops Research* 2008. 105 PP 1–14
11. BUDAK, H. *et al.* Potential uses of molecular markers in crop improvement. *Journal of science and engineering*. 2004 - 7(1).
12. DE VIENNE, D. 2003. *Molecular markers in plant genetics and biotechnology*. Science Publishers, Inc. USA.
13. VARSHNEY, R *et al.* A comprehensive resource of drought and salinity responsive ESTs for gene discovery and marker development in chickpea (*Cicer arietinum* L.) *BMC Genomics*. (2009) 10(523)
14. **Varshney, R. Nayak, S. May, G & Jackson, S.** Next-generation sequencing technologies and their implications for crop genetics and breeding. [Trends in Biotechnology](#) (2009) 27(9): 522-530
15. VARSHNEY, R, GRANER, A &SORRELLS, M. Genic microsatellite markers in plants: features and applications. *TRENDS in Biotechnology* (2005) Vol.23 No.1
16. SELKOE, K & TOONEN, R. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters* (2006), 9: 615–629
17. LI, Y. KOROL, A. FAHIMA, T. & NEVO, E. Microsatellites within genes: Structure, function and evolution. *Molecular biology and Evolution* (2004) 21(6): 991-1007
18. ARAYA, E. *et al.* Relaciones genéticas en una colección de clones de *Gmelina arborea* (Roxb) reveladas con marcadores AFLP. Kurú: *Revista Forestal* (Costa Rica) 2005. 2(6)
19. BERTRAND, C. *et al.* Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. *Trans. Royal Society. B* 2008 - 363, 557-572
20. AVISE, J.2004. *Molecular Markers, Natural history and Evolution*. Second edition. Sinauer Associates, Inc. Publishers. Massachusett.
21. DUBKOVSKY, J.. Marker-assisted selection in public breeding programs: The wheat experience. University of California, Davis. (2004) 1895-1897
22. DVORAK, W. World view of *Gmelina arborea*: opportunities and challenges. *New Forest* 28 (2004): 111–126,

23. MOYA, R. Effect of management treatment and growing regions on wood properties of *Gmelina arborea* in Costa Rica. *New Forest* (2004) 28: 325–330
24. NAIK, D. SINGH, D. VARTAK, V. PARANJPE, S. & BHARGAVA, S. Assessment of morphological and genetic diversity in *Gmelina arborea* Roxb. *New Forest* (2009). 38: 99-115
25. COMINELLI, & TONELLI. Transgenic crops coping with water scarcity - In: *New Biotechnology*. (2010) - 27:5, pp. 473-477.
26. SHINOZAKI & YAMAGUCHI. Gene networks involved in drought stress response and tolerance. *Journal of Experimental Botany* (2007) Vol 58, No 2, pp 221-227
27. GUTTERSON & REUBER. Regulation of disease resistance pathways by AP2/ERF transcription factors. (2004). (4):465-71.
28. GUILTINAN, MARCOTTE & QUATRANO. A plant leucine zipper protein that recognizes an abscisic acid response element. *Science*. (1990) Vol. 250 no. 4978 pp. 267-271
29. OLSEN, ERNST, LEGGIO & SKRIVER. NAC transcription factors: structurally distinct, functionally diverse. *TRENDS in Plant Science* (2005) Vol.10 No.2 Pp 79-87
30. HONGHONG, DAI, YAO, XIAO, LI, ZHANG & XIONG. Overexpressing a NAM, ATAF, and CUC (NAC) transcription factor enhances drought resistance and salt tolerance in rice. *PNAS*. (2006) vol. 103 no. 35 12987–12992
31. DERBYSHIRE, P. DREA, S. SHAW, P. DOONAN, J. DOLAN, L. Proximal–distal patterns of transcription factor gene expression during *Arabidopsis* root development. *Experimental Botany*. (2008) 59: 235-245
32. BOTER, M. RUIZ, O. ASHRAF, A & PRAT, S. Conserved MYC transcription factors play a key role in jasmonate signaling both in Tomato and *Arabidopsis*. *Genes and Development*. (2004). 18: 1577-1591
33. SAKAMOTO, H. MARUYAMA, K. SAKUMA, Y. MESHI, T. IWABUCHI, M. SHINOZAKI, K AND YAMAGUCHI-SHINOZAKI, K. *Arabidopsis* Cys2/His2-Type Zinc-Finger Proteins Function as Transcription Repressors under Drought, Cold, and

- High-Salinity Stress Conditions. *Cell Biology and Signal Transduction – Plant Physiology*. (2004) 136:2734-2746
34. [ZHANG Y, WANG L](#). The WRKY transcription factor superfamily: its origin in eukaryotes and expansion in plants. [BMC Evol Biol](#). 2005 3;5:1.
35. DE LAS RIVAS, J & BARBER, J. Analysis of structure of the PsbO protein and its implications. *Photosynthesis Research*.(2004) 81: 329-343
36. FU. [The superfamily of arginine/serine-rich splicing factors](#). *RNA*. (1995). Cambridge University Press. P 663 – 680
37. FRENTZEN, *et al*. Specificities and Selectivities of Glycerol-3-Phosphate Acyltransferase and Monoacylglycerol-3-Phosphate Acyltransferase from Pea and Spinach Chloroplasts. *European Journal of Biochemistry* (1983) [Volume 129, Issue 3](#), pages 629–636,
38. VALENTI, V AND PUPILLO, P. Activation Kinetics of NAD-Dependent Malic Enzyme of Cauliflower Bud Mitochondria. **Plant Physiol**. 1981 November; 68(5): 1191–1196.
39. DOGHMEN, R. Sumo Protein modifications. [Biochim Biophys Acta](#). (2004);1695(1-3):113-31
40. CARLIER, M. et al. Actin depolymerizing factor (ADF/Cofilin) enhances the rates of filament turnover: Implications in actine-based motility. *Journal of Cell Biology*. (1997) 136(6): 1307-1323
41. ESPARTERO, J. PINTOR J & PARDO, J. Differential accumulation of S-adenosylmethionine synthetase transcripts in response to salt stress. [PLANT MOLECULAR BIOLOGY Volume 25, Number 2](#), 217-227
42. DIEBOLD *et al*. The Branched-Chain Amino Acid Transaminase Gene Family in Arabidopsis Encodes Plastid and Mitochondrial Proteins1 *Plant Physiology*, June 2002, Vol. 129, pp. 540–550

43. NAVROT, N. *et al.* Plant Glutathione Peroxidases Are Functional Peroxiredoxins Distributed in Several Subcellular Compartments and Regulated during Biotic and Abiotic Stresses. *Plant Physiology* (2006) 142:1364-1379
44. HUDSPETH *ET AL.* Characterization and expression of Metallothionein-like genes in cotton. [PLANT MOLECULAR BIOLOGY](#). (1996). [Volume 31, Number 3](#), 701-705
45. WATERS, E. LEE, G & VIERLING, E. Evolution, structure and function of the small heat shock proteins in plants. *J. Exp. Bot.* (1996) 47 (3): 325-338
46. KOSAREV, P. MAYER, K & HARDTKE, C. Evaluation and classification of RING-finger domains encoded by the Arabidopsis genome. *Genome Biology* (2002), 3(4)
47. YU *ET AL.* Molecular analysis of phosphomannomutase (PMM) genes reveals a unique *PMM* duplication event in diverse *Triticeae* species and the main *PMM* isozymes in bread wheat tissues. *BMC Plant Biology* (2010), 10:214.
48. BADGER & PRICE. The Role of Carbonic Anhydrase in Photosynthesis. *Annual Review of Plant Physiology and Plant Molecular Biology* 1994. Vol. 45: 369-392
49. GALICHET, A & GRUISSEM, W. Protein farnesylation in plants — conserved mechanisms but different targets. *Current Opinion in Plant Biology* 2003, 6:530–535
50. LEHR *et al.* cDNA and genomic cloning of sugar beet V-type H⁺-ATPase subunit A and c isoforms: evidence for coordinate expression during plant development and coordinate induction in response to high salinity. *Plant Molecular Biology* (1999) 39: 463–475
51. GUPTA, P & VARSHNEY, R.. The development and use of microsatellite markers for genetic analysis and plant breeding with emphasis on bread wheat. *Euphytica*. (2000) 113. 163-185.
52. WESING, K. NYMBOM, H. WOLFF, K & KAHL, G. DNA fingerprinting in plants. Principles, Methods and Applications. Second Edition. CRC Press – Taylor and Francis Group (2005)

53. **Varshney, R. Nayak, S. May, G & Jackson, S.** Next-generation sequencing technologies and their implications for crop genetics and breeding. [Trends in Biotechnology](#) (2009) 27(9): 522-530
54. SALEKDEH, G. REYNOLDS, M. BENNETT, J & BOYER, J. Conceptual framework for drought phenotyping during molecular breeding. *Trends in Plant Science* (2009) 14(9): 488-496
55. PRASAD, M · VARSHNEY R, ROY, J, BALYAN H, GUPTA P. The use of microsatellites for detecting DNA polymorphism, genotype identification and genetic diversity in wheat. *Theor Appl Genet* (2000) 100:584–592
56. JARNE, P & LAGODA, P. Microsatellites, from molecules to populations and back. *Tree* (1996) 11:424 429
57. ELLEGREN, H. Microsatellites: Simple sequence with complex evolution. *Nature Reviews*. (2004). 5: 435 – 445
58. [VIDYBIDA AK](#), [ERKIZAN V](#), [ALTUNGÖZ O](#), [CELEBI G](#). Testing of nonlinear electrofrictiophoresis in agarose gel. [Bioelectrochemistry](#). (2000). 52(1): 91-101.
59. TING, N. ZAKI, N. ROSLI, R. LOW, E. ITHNIN, M. CHEAH, S. TAN, S. & SING, R. SSR mining in oil palm EST database: application in oil palm germplasm diversity studies. *Journal of Genetics*. (2010). 89.
60. LAURIDSEN, E & KJAER, E. Provenance research in *Gmelina arborea* Linn., Roxb. A summary of results from three decades of research and a discussion of how to use them. *International Forestry Review* (2002) 4(1),
61. STEWART, C. 2008. Plant biotechnology and genetics, principles techniques and applications: Plant Breeding. Wiley publications. Pp47-83
62. KLIMES, L. KLIMESOVA, J. HENDRICKS, R & GROENENDAEL, J. Clonal plant architecture: a comparative analysis of form and function. *The Ecology and Evolution of clonal plant*. Backbuys Publishers. (1997) 1-29

63. BEAKER, S. SINKE, R. & PEARSON, P. Differences in stutter intensities between microsatellites are related to length and sequence of the repeat. En línea. <http://igitur-archive.library.uu.nl/dissertations/2005-0531-200105/c8.pdf>
64. BUTLER, J. Forensic DNA Typing, Biology, Technology and Genetics of STR markers. Second Edition. ELSERVIER. (2005)
65. HARTL & CLARK. Population genetics. Third Edition. Sinauer Associates. (1997)
66. ELLIS, J & BURKE, J. EST-SSRs as a resource for population genetic analyses. *Heredity* (2007) 99:125 – 132
67. IPGRI & CORNELL UNIVERSITY. Measures of genetic diversity. (2003)
68. OOSTERHOUT, C. HUTCHINSON, W. WILLS, D & SHIPLEY, P. MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* (2004). 4: 535-538
69. GRIFFITHS, A. WESSLER, S. LEWONTIN, R & CARROLL, S. Introduction to genetic analysis. 9th Edition. W.H. Freeman and Company (2008).

ANEXO 1

Extracción de DNA de *Gmelina arborea*

Tabla 1: Pureza y concentración de las muestras de DNA de *Gmelina arborea* según el protocolo de Stefenon *et al.* 2004. Primera extracción

Individuo	260/280	260/230	ng/μl
1 ^a	1,81	1,08	187,1
1b	1,79	1,37	298,2
15 ^a	1,67	0,79	231,6
15b	1,78	1,14	228,7
24 ^a	1,73	0,96	273,1
24b	1,73	0,96	140,8
31 ^a	1,52	0,53	181,8
31b	1,75	0,98	98,2
32	1,61	0,70	298,1
37 ^a	1,63	0,76	227,0
37b	1,71	1,03	428,0
38 ^a	1,77	1,09	297,0
38b	1,74	1,03	187,9
40 ^a	1,93	1,67	273,8
40b	1,87	1,40	504,5
49 ^a	1,74	0,89	135,7
49b	1,78	1,14	175,9
51 ^a	1,84	1,11	347,1
51b	1,85	1,52	441,1
53 ^a	1,39	0,60	85,7
53b	1,05	0,30	48,1
54 ^a	1,89	1,62	568,3
54b	1,91	1,79	1210,2
55 ^a	1,41	0,44	86,3

55b	1,64	0,84	40,9
56 ^a	1,36	0,44	168,1
56b	1,49	0,54	59,2
60 ^a	1,73	0,92	186,9
60b	1,64	0,68	96,9
61 ^a	1,77	1,12	290,1
61b	1,57	0,60	218,2
62 ^a	1,82	1,42	411,2
62b	1,56	0,54	61,3
64 ^a	1,51	0,57	125,3
64b	1,50	0,52	105,1
65 ^a	1,75	0,96	111,9
65b	1,84	1,50	558,4
66 ^a	1,80	1,08	216,6
66b	1,77	0,96	155,0
67 ^a	1,90	1,64	464,1
67b	1,89	1,63	449,2
75 ^a	1,57	0,63	96,9
75b	1,69	0,92	234,4
78 ^a	1,82	1,38	508,6
78b	1,85	1,47	505,4
79	1,69	0,83	218,3
84 ^a	1,72	1,12	204,4
84b	1,25	0,30	123,7
86	0,85	0,14	19,2
91 ^a	1,64	0,79	215,2
91b	1,52	0,63	222,9
94 ^a	1,71	0,96	373,6
94b	1,73	1,02	355,5
109 ^a	1,41	0,43	73,4

109b	1,38	0,44	92,4
112 ^a	1,84	1,39	297,9
112b	1,74	1,02	237,2
129	1,85	1,63	851,0

Tabla 2: Pureza y concentración de las muestras de DNA de *Gmelina arborea* según el protocolo de Stefenon *et al.* 2004. Segunda extracción

Individuo	260/280	260/230	ng/μl
31a	2,01	1,62	188,7
31b	1,90	1,50	124,2
32a	1,92	1,56	498,9
32b	2,11	2,13	91,3
38a	2,01	1,62	77,0
38b	2,01	1,62	97,1
40a	1,96	1,74	194,1
40b	1,85	1,61	557,2
48a	2,01	1,85	99,4
48b	1,98	1,72	1356,6
53a	1,87	1,45	51,8
53b	1,86	1,45	195,3
55a	2,05	1,87	277,0
55b	2,04	2,01	407,8
56a	2,07	1,93	722,3
56b	2,06	2,02	439,3
60a	2,09	1,60	74,8
60b	2,08	1,83	86,5
64a	1,99	1,70	144,4
64b	2,02	1,60	79,0

66a	2,05	1,81	236,7
66b	2,06	1,84	416,1
67a	2,09	2,04	190,3
67b	2,02	1,73	526,1
68a	1,90	1,66	548,7
68b	2,05	2,04	704,8
75a	1,93	1,44	436,3
75b	2,01	1,74	180,0
79a	2,11	2,11	796,1
79b	2,14	2,10	972,0
86a	1,98	1,66	297,9
86b	1,98	1,79	271,6
91a	2,08	2,03	235,2
91b	2,08	1,99	380,2
109a	2,06	2,12	230,6
109b	2,06	2,05	312,9
112a	1,95	1,46	160,7
112b	2,06	1,96	125,6
129a	2,14	2,25	1189,7
129b	2,16	2,22	1140,3

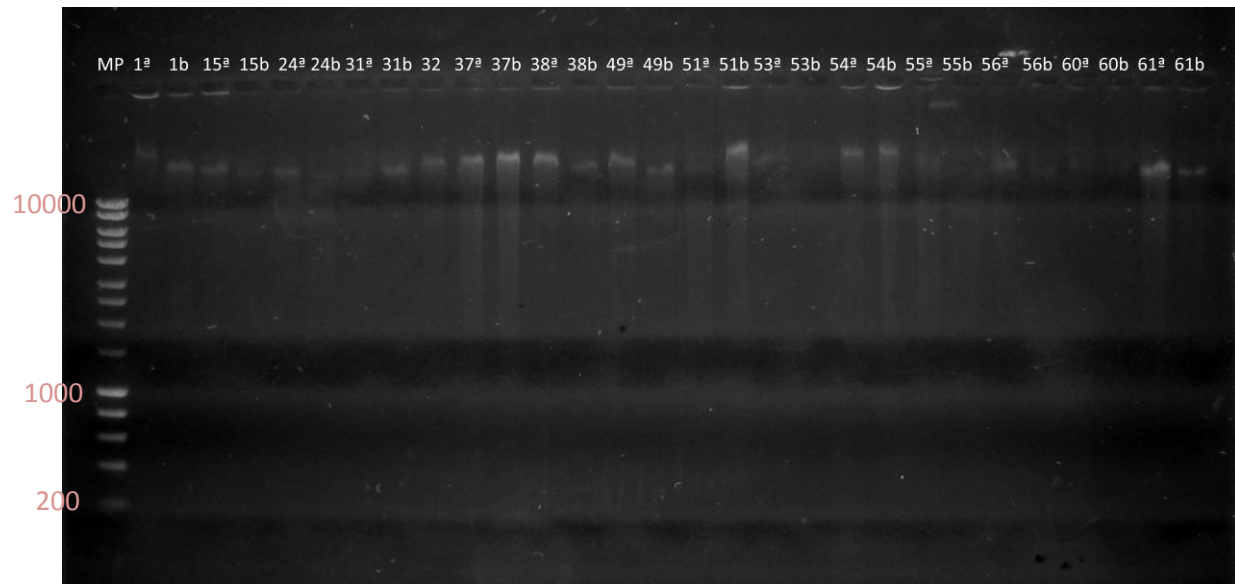


Imagen 1: Prueba electroforesis ADN total de *Gmelina arborea* en geles de agarosa al 0,8%. A. Clones 1 al 61 Primera extracción. Se utilizó marcador de peso molecular HypperLadder I, el peso molecular del ADN total de *Gmelina* supera los 10.000 pb

ANEXO 2

Protocolo de Tinción con Nitrato de Plata propuesto por Radojkovic* & Kus̃ic 2000

1. Realizada la electroforesis en acrilamida se coloca el vidrio en solución fijadora (etanol y ácido acético) durante mínimo 10 minutos, a temperatura ambiente y preferiblemente en shaker.
2. Lavar el vidrio dos veces durante 3 minutos con agua destilada
3. Disponer el vidrio en solución de tinción (Nitrato de plata 0,1% w/v, Formaldehído 0,05%)
4. Realizar un lavado con agua destilada para remover el exceso de plata durante 15 segundos
5. Realizar un lavado en solución de revelado (NaOH 1,5%, NaBH₄ 0,01%, Formaldehído 0,05%, Tiosulfato 185,7 g/L)
6. Incubar en la misma solución con agitación constante durante 20 minutos o hasta ver bandas.
7. Colocar en solución de fijación durante 10 minutos a temperatura ambiente
8. Lavar el vidrio con agua destilada durante 5 minutos.

ANEXO 3

Índices y valores utilizados en el análisis estadístico (Peakall & Smouse 2010)(Hoelzel & Dover 1991)

Índice/Valor	Descripción	Ecuación	Rango
Nº de alelos (N _a)	Número total de alelos en el locus, determinado por cuenta directa		0 a ∞
Frecuencia alélica (p _i)	Medida de la frecuencia de un alelo con respecto a todos los alelos de la población.	$\frac{N^{\circ} i *}{Na}$	0 a 1
Polimorfismo o Tasa de Polimorfismo (P _j)	La frecuencia de un alelo es menor o igual a 0.95 o 0.99. Debe tenerse en cuenta que estos valores son arbitrarios.	P _j =p _i ≤0.95 ó P _j =q≤0.99	0-100%
Proporción de loci polimórficos	Número de loci polimórficos sobre número total de loci. Expresa el porcentaje de loci variables en una población.	P=n _{p_i} /n total	0-1
Nº de alelos efectivos (N _e)	Provee un número estimado de las frecuencias alélicas en equilibrio en una población ideal.	$\frac{1}{1 - He}$	0 a n alelos
Heterocigosis observada (H _o)	Calculada por cada locus, donde el número de heterocigotos está determinado por la cuenta directa y el número de muestras	$\frac{N^{\circ} Heterocigotos}{N}$	0 a 1
Heterocigosis esperada (H _e)	Nº de heterocigotos esperado teniendo en cuenta las	$1 - \sum p_i^2$	0 a 1

	frecuencias alélicas. Establece la probabilidad de que en un único locus, un par de alelos escogidos al azar sean diferentes entre sí. La H_e promedio de todos los loci es una medida del grado de variabilidad genética de una población		
Heterocigosis equilibrada (U_{H_e})	Nº de heterocigotos esperados empleando El algoritmo de Levene.	$\frac{2n}{n-1}(1 - \sum p_i^2)$	0 a 1
Indice de fijación	Cuando existe apareamiento aleatorio se esperan valores cercanos a 0 mientras que valores positivos indican endogamia o alelos nulos no detectados. Valores negativos indican exceso de heterocigosis	$\frac{H_e - H_o}{H_e}$	-1 a 1
Distancia genética Nei	Distancia entre los individuos de una población o entre poblaciones asumiendo que existe diferencias genéticas debido a mutaciones y deriva genética. Nei definió la distancia genética como una función de la frecuencia de los genotipos	-Ln I Donde $I = J_{xy}/J_x J_y$ Así J_{xy} es la probabilidad de que se	0 a

*i: Nº de repeticiones Del alelo i en La población

Anexo 4

Análisis UPGMA - Distancias Genéticas de Nei (1972)

UPGMA Cluster using Nei's (1972) minimum distance

Nodos de Distancia

1	0.2500	27	28								
2	0.2917	1	27	28							
3	0.5000	1	27	28	29						
4	0.5000	1	14	27	28	29					
5	0.5500	1	14	27	28	29	32				
6	0.6250	1	14	16	27	28	29	32			
7	0.2500	7	15								
8	0.4583	6	7	15							
9	0.2500	3	25								
10	0.3750	3	10	25							
11	0.4167	3	10	22	25						
12	0.5625	3	10	22	25	30					
13	0.3333	13	17								
14	0.4583	13	17	23							
15	0.5000	13	17	19	23						
16	0.5833	6	7	13	15	17	19	23			
17	0.4167	5	9								
18	0.5833	5	9	21							
19	0.4167	8	18								
20	0.7381	5	6	7	9	13	15	17	19	21	23

21 0.4167 26 31
22 0.5000 11 26 31
23 0.3333 2 24
24 0.7494 5 6 7 9 11 13 15 17 19 21 23 26 31
25 0.7852 1 5 6 7 9 11 13 14 15 16 17 19 21 23 26 27 28 29 31 32
26 0.8656 1 5 6 7 9 11 12 13 14 15 16 17 19 21 23 26 27 28 29 31 32
27 0.9010 1 5 6 7 9 11 12 13 14 15 16 17 19 21 23 26 27 28 29 31 32 33
28 0.3333 4 20
29 0.6389 2 11 24 26 31
30 0.6750 2 4 11 20 24 26 31
31 0.7262 2 3 4 10 11 20 22 24 25 26 30 31
32 0.6528 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23