

■ Capítulo 10

Los marcadores moleculares en la investigación de los recursos genéticos y fitomejoramiento de la palma de aceite

Luz Stella Barrero, Gerardo Gallego, Harold Suárez-Barón, Myriam Duque, Jaime A. Osorio Guarín, Erick Hernández, Felix E. Enciso Rodríguez, Gina A. Garzón Martínez, Leidy Paola Delgadillo Durán, Silvio Bastidas

La palma de aceite *Elaeis guineensis* (G) es considerada la mayor fuente de aceite vegetal en el mundo (Turner et al., 2008), con un gran potencial en las industrias de alimentos y oleoquímica, y en la producción de combustible biodiésel. Esta última es una de las apuestas del sector palmicultor colombiano, como quedó expresado en el libro *Biodiésel de palma colombiano: de la ficción energética a la realidad de un negocio*, publicado por Fedepalma (2013). Dada la importancia que ha adquirido en Colombia, su característica perenne, sus usos y su potencial, el cultivo de la palma de aceite ha sido incluido por el MADR para su promoción dentro de la política estatal actual para la reactivación del sector agropecuario, lo cual incluye alianzas estratégicas entre grandes y pequeños productores. En Colombia, esta especie enfrenta problemas limitantes como la mancha anular, la PC, la marchitez sorpresiva y el anillo rojo, que afectan la sostenibilidad de su cultivo (Bastidas et al., 2007). La PC se considera el problema más limitante del cultivo, y para ella no hay medidas efectivas de control, por lo que la resistencia genética se perfila como una alternativa prometedora mediante la incorporación de resistencia aportada por la palma americana *E. oleifera* (O).

Las palmas *E. guineensis* y *E. oleifera* son especies alógamas con poblaciones heterogéneas heterocigotas, por lo cual los híbridos O × G surgidos de estas dos especies manifiestan signos de segregación bajo condiciones de campo. Esta observación sería un indicador de segregación de alelos parentales y, por

lo tanto, de contenidos genéticos diferenciales que pueden ser conocidos con mayor precisión mediante el uso de marcadores moleculares. Adicionalmente, el uso de estos marcadores permite: 1) conocer la variabilidad genética y la estructura de poblaciones híbridas y parentales; 2) identificar poblaciones híbridas con similitud diferencial a las parentales, lo cual puede llevarse a cabo en etapas de vivero, y 3) realizar estudios de asociación con fenotipos de interés para acelerar el proceso de mejoramiento y de este modo contribuir a reducir los costos y tiempos, dado que un ciclo genético en palma de aceite equivale a diez u once años.

El presente capítulo tiene como propósito resumir el estado del arte de la investigación con marcadores moleculares en palma de aceite a nivel internacional y nacional. Sobre la investigación nacional, se hace énfasis en los avances logrados por AGROSAVIA con poblaciones de mejoramiento genético provenientes de la iniciativa desarrollada en el C. I. El Mira de la misma corporación.

Contexto internacional

Los centros alrededor del mundo que adelantan investigaciones sobre la palma africana (*E. guineensis*) en la actualidad implementan un variado número de estudios moleculares basados en marcadores de ADN o regiones específicas del genoma de la palma, con diversas aplicaciones, como estudios de “huella genética”, análisis de diversidad genética, mapas de ligamiento y selección genómica (del inglés *genomic selection* [GS]). La huella genética o *fingerprinting* ha involucrado diferentes tipos de marcadores, entre los cuales se destaca un alto número de microsatélites o secuencias simples repetidas (*simple sequence repeats* [SSR]), disponibles en bases de datos (Billotte et al., 1999, 2001) y con un alto poder discriminativo y relativo bajo costo.

Varios estudios han priorizado el uso de SSR para el conocimiento de los recursos genéticos de la palma de aceite. Así, Bakoumé et al. (2007) determinaron la diversidad alélica dentro de cuatro poblaciones naturales, que representan diez países africanos, tres cultivares mejorados y un material semisilvestre. Con el uso de 16 locus SSR, se detectaron 209 alelos, de los cuales algunos pocos son comunes en poblaciones de áreas de clima seco. Por su parte, Singh et al. (2008) usaron etiquetas de secuencias expresadas

(*expressed sequence tags* [EST]) de la palma de aceite para buscar marcadores SSR. Los autores desarrollaron diez (EST-SSR) para evaluar 76 muestras de palma de aceite de siete países de África y una población estándar de Deli Dura. El número promedio de alelos observados y el contenido de información polimórfica (PIC) fueron de 2,56 y 0,53, respectivamente, con una diferenciación genética (F_{ST}) de 0,249 entre poblaciones, lo que indica un alto nivel de divergencia.

En otro estudio, Cochard et al. (2009) analizaron 318 individuos de 26 orígenes y ocho países con catorce marcadores SSR, con el propósito de describir la estructura de estos recursos genéticos involucrados en la historia del mejoramiento de la palma, en relación con los orígenes africanos subespontáneos. Los datos revelaron dos grupos originales: el grupo I, formado por África occidental, y el grupo II, formado por el conjunto Benín-Nigeria-Camerún-Congo-Angola. Un tercer grupo Deli, derivado del II, fue el resultado de selección artificial. Igualmente, Zaki et al. (2012) caracterizaron 20 SSR_s a partir de una librería genómica de *E. oleifera*, y confirmaron el potencial de estos marcadores para revelar la diversidad genética en el género *Elaeis*.

Además del alto número de análisis de diversidad genética llevados a cabo con bases de datos saturadas con SSR, otros estudios han utilizado diversos marcadores moleculares, incluidos los polimorfismos de longitud de fragmentos amplificados (*amplified fragment length polymorphisms* [AFLP]) (Kularatne et al., 2001); los polimorfismos de longitud de fragmentos de restricción (*restriction fragment length polymorphisms* [RFLP]) (Maizura et al., 2006; Mayes et al., 2000; Moretzsohn et al., 2002) y los marcadores basados en ADN polimórfico amplificado aleatoriamente (*random amplified polymorphic DNA* [RAPD]) (Shah et al., 1994). Así, los marcadores moleculares en palma de aceite han sido usados para avanzar en el fitomejoramiento de esta especie vegetal.

En lo que respecta a mapas de ligamiento, el primero de ellos fue construido por Mayes et al. (1997) a partir de RFLP, y posteriormente, con el uso otros marcadores se logró construir el primer mapa de ligamiento de alta densidad para la especie, con 255 SSR_s y 689 AFLP, distribuidos en 16 grupos de ligamiento correspondientes a los 16 pares de cromosomas homólogos de la palma africana (Billotte et al., 2005). El marcador más cercano al gen que codifica el

grosor de la cáscara fue un AFLP ubicado a 4,7 (centimorgan) del gen *shell* (Sh), el cual controla la presencia o ausencia de revestimiento en la fruta de la palma (Billotte et al., 2005). El gen que determina las tres diferentes formas del fruto en las palmas Dura, Ténera y Pisífera fue clonado gracias a metodologías genómicas, por lo que ahora se puede usar como marcador para distinguir estas formas antes del trasplante definitivo a campo y de este modo ahorrar tiempo y dinero en el fitomejoramiento (Singh et al., 2013). Además, Price et al. (2003) desarrollaron un sistema de marcadores moleculares para palma de aceite basados en retrotransposones, el cual permite su aplicación en propósitos de mejoramiento de la palma africana y de los híbridos que se obtienen mediante cruzamiento con la palma americana (*E. oleifera*). Estos marcadores pueden usarse como fuente de información, debido a su capacidad de integrarse dentro de un gran número de locus del genoma y, por tanto, de generar polimorfismos insercionales entre individuos.

El Malaysian Palm Oil Board (MPOB) desarrolló y evaluó en 2007 el potencial que los polimorfismos de un solo nucleótido (*single nucleotide polymorphisms* [SNP]) tienen como marcadores moleculares de nueva generación. Estos marcadores permiten realizar, con un alto rendimiento, análisis de mapeo genético y de diversidad, protección de germoplasma vegetal, investigaciones en genómica del aceite de palma y mejoramiento vegetal asistido por marcadores. En el estudio en mención se realizó un análisis de alrededor de 380.000 secuencias genómicas de las dos especies de palma, *E. guineensis* (Dura, Ténera y Pisífera) y *E. oleifera*. Este trabajo permitió elucidar también, a partir del análisis de secuencias polimórficas amplificadas (*cleaved amplified polymorphic sequences* [CAPS]), el potencial de diferenciación que estos SNP presentan a nivel intra e interespecífico (Ong et al., 2015).

Un consorcio que incluye el MPOB secuenció el genoma de *E. guineensis* (Singh et al., 2013) y adelantó estudios con marcadores SNP utilizando arreglos de Illumina en poblaciones híbridas interespecíficas O × G e intraespecíficas Dura × Pisífera (Ting et al., 2014). Adicionalmente, en un esfuerzo por aplicar GS en poblaciones de mejoramiento, el Centre de Coopération Internationale en Recherche Agronomique Pour le Développement (CIRAD) y AGROSAVIA han implementado recientemente el uso de SNP vía genotipificación por secuenciación (*genotyping by sequencing* [GBS]) (Cros et al., 2017; Osorio-Guarín et al., 2019), lo que ha permitido incrementar la precisión de GS, realizada

inicialmente con SSR (Cros et al., 2013) y más recientemente con marcadores SNP (Garzón-Martínez et al., 2022). Se espera que la investigación molecular y el desarrollo de los recursos genómicos en palma de aceite suministre a los fitomejoradores un medio poderoso para acelerar y conseguir ganancias por selección, de tal manera que la palma alcance su potencial de forma más acelerada como cultivo de aceite clave para el futuro (Mayes et al., 2008).

Contexto nacional

En lo que respecta a la investigación realizada a nivel molecular en palma de aceite en Colombia, son varios los estudios que se han llevado a cabo hasta el momento. Estas caracterizaciones responden a la necesidad de proporcionar a los cultivadores de esta especie vegetal un material con mayor potencial genético, a fin de satisfacer la creciente demanda de aceite de palma, dado que Colombia es el primer productor de Suramérica y ha pasado a ocupar el quinto en el mundo (Cook, 2023).

Es así como se destacan inicialmente los estudios de Ochoa et al. (1997), quienes reportan la identificación de marcadores moleculares tipo RAPD; los de Villegas et al. (2000), quienes caracterizaron molecularmente la palma de aceite tipo Dura; y los de Arias y Rocha (2004), los cuales analizaron a nivel molecular la diversidad genética en cultivares resistentes y susceptibles a la PC.

Posteriormente, Montoya et al. (2005) estimaron la diversidad y estructura genética de 48 entradas pertenecientes a seis poblaciones procedentes de Angola con 20 combinaciones de marcadores SSR, y lograron identificar 69 alelos y una heterocigosidad esperada (H_e) de 0,45. A su vez, Galeano (2005) realizó una investigación cuyo objetivo fue estandarizar el uso de AFLP en materiales Dura procedentes del programa de fitomejoramiento de Cenipalma, y realizó de manera preliminar la caracterización molecular de doce palmas de los tipos Dura, Ténera y *E. oleifera*. Según los agrupamientos hechos con los métodos fenéticos y multivariados, se encontró que la mayoría de palmas tipo Dura se agrupan de acuerdo con el lugar de colecta,

mientras que la palma de origen americano se ubica de forma distante al compararla con la de origen africano.

Barrera (2011), por su parte, evaluó la diversidad y estructura genética de una colección de 455 entradas de palma de aceite colectadas en cinco zonas geográficas de Angola y nueve híbridos interespecíficos, con 19 locus SSR, a fin de aportar información a la colección de germoplasma de Cenipalma y su programa de mejoramiento genético.

Contexto con materiales de mejoramiento genético producidos por AGROSAVIA

El programa de mejoramiento de AGROSAVIA dio lugar al híbrido O × G a partir de parentales de las especies *E. oleifera* (palma nolí) y polen de palmas *E. guineensis* de la serie La Mé de Costa de Marfil. La semilla del híbrido O × G presenta bajo porcentaje de germinación, lo cual incrementa los costos de la semilla (USD 3/unidad), en comparación con *E. guineensis* Ténera (USD 0,75). Por su parte, los híbridos interespecíficos O × G presentan alta esterilidad (alto porcentaje de frutos partenocárpicos) y requieren polinización asistida obligatoria, cuyo costo varía entre USD 300 y 350 ha/año, lo que restringe su uso entre los pequeños y medianos productores. Los híbridos O × G Corpoica Elmira en la actualidad se encuentran en pruebas de evaluación agronómica ante el ICA, en las cuales se está demostrando que no requieren polinización asistida obligatoria, porque heredan de sus madres *E. oleifera* la característica de inflorescencias expuestas, lo que ahorra hasta USD 300/ha/año, y además presentan niveles de incidencia de PC inferiores a 1% (acumulado de tres años).

AGROSAVIA cuenta también con una colección de trabajo mejorada compuesta por 20 progenies (descendientes) *E. guineensis* Dura de segunda generación (F2) y diez progenies (descendientes) Ténera/Pisífera de tercera generación (F3). Las características detectadas en los híbridos O × G Corpoica Elmira y la colección de trabajo justifican su caracterización molecular con fines de protección y asistencia al programa de mejoramiento genético para características de interés, acelerando en el tiempo la producción de cultivares.

Debido al alto número de materiales producto de los diferentes cruces ha sido necesario realizar diversos estudios de caracterización, con el fin de generar insumos para seleccionar las mejores progenies.

Marcadores bioquímicos

Los primeros marcadores utilizados en la población de mejoramiento de AGROSAVIA datan de 1995, cuando se utilizaron muestras foliares de parentales de *E. guineensis* Dura y Pisífera y de *E. oleifera*, en un esfuerzo inicial por evaluar la introgresión recíproca entre las dos especies en híbridos F1 obtenidos por polinización controlada. Se utilizaron marcadores bioquímicos mediante análisis electroforético de isoenzimas, y se encontraron polimorfismos asociados exclusivamente con el parental *E. guineensis* Pisífera para cuatro isoenzimas (esterasa, fosfatasa ácida, polifenol oxidasa, fosfoglucomutasa) y relacionados con el parental *E. oleifera* en los patrones de dos isoenzimas (fosfogluconato deshidrogenasa y fosfoglucomutasa), los cuales segregaron diferencialmente en los individuos del híbrido F1. Perfiles electroforéticos de nuevas isoenzimas (por ejemplo, glutamato oxalacetato transaminasa) y proteínas totales también permitieron diferenciar de forma preliminar las dos especies y su progenie híbrida (Hernández et al., 1997; Vives et al., 1997).

Marcadores AFLP

Posteriormente, se inició una estandarización de marcadores moleculares tipo AFLP y SSR como base para la futura identificación molecular de semilla híbrida (Arias & Núñez, 2006). Se analizó una muestra de ocho palmas proveniente de dos parentales y seis híbridos del proceso de mejoramiento adelantado por AGROSAVIA. Para AFLP se evaluaron seis combinaciones de cebadores previamente utilizados por varios investigadores (Barcelos, 1998; Purba et al., 2000; Rajinder et al., 1998). La combinación de cebadores E-ACT/M-CAC presentó el mayor número de polimorfismos (18 bandas polimórficas), seguida de la combinación E-AGC/M-CAA (16 bandas polimórficas), y la E-ACC/M-CAC (seis bandas polimórficas). Para los SSR se usaron cuatro pares de cebadores reportados por Billotte et al. (2001), y se obtuvieron amplicones polimórficos con los cebadores mEgCIR0008, mEgCIR0465 y mEgCIR0134, en los que fue posible distinguir parentales de progenie. En ambos casos se encontraron grupos genéticos diferenciales

conformados por la palma materna, la paterna y los hijos, con semejanzas a uno u otro parental. El estudio fue un acercamiento preliminar, que evidenció la importancia de incrementar el número de marcadores y el tamaño de la población por analizar.

Marcadores microsatélites (SSR)

En el año 2011, AGROSAVIA, en convenio con el Centro Internacional de Agricultura Tropical (CIAT), realizó una caracterización molecular con 46 marcadores tipo SSR (incluidos los tres reportados por Arias & Núñez [2006]) en 577 palmas provenientes del programa de mejoramiento genético de AGROSAVIA (Gallego et al., 2012) (tabla 10.1). El proceso fue realizado de forma conjunta por el Laboratorio de Biotecnología del CIAT y el Laboratorio de Genética Molecular Vegetal de AGROSAVIA.

Los 46 cebadores SSR fueron seleccionados de acuerdo con diferentes índices relacionados con diversidad genética, como el PIC, la heterocigosidad observada (H_o) y el número de alelos observados (A_o), previamente publicados en palma de aceite (Billotte et al., 2001; Ting et al., 2010; Zaki et al., 2010). Los marcadores fueron analizados en una muestra inicial de doce individuos, y se identificaron un total de 27 SSR polimórficos (tabla 10.2).

Del set de 27, se seleccionaron diez SSR con base en variación interespecífica y amplificación en la mayor parte de la población muestreada. Estos marcadores corresponden a seis SSR genómicos reportados por Billotte et al. (2001) (mEgCIR008 y mEgCIR230) y Zaki et al. (2010) (sMo00018, sMo00027, sMo00134 y sMo00138). Los otros cuatro son marcadores de naturaleza génica derivados de secuencias EST, de las cuales dos (que contienen los marcadores sEg00009 y sEg00076) no presentaron secuencias similares en las bases de datos del National Center for Biotechnology Information (NCBI), mientras que otros dos (sEg00125 y sEg00126) fueron homólogos con una proteína hipotética reportada para la especie *Vitis vinifera* y con una proteína determinada por predicción para la especie *Populus trichocarpa*. El marcador mEgCIR230 reportado por Billotte et al. (2001) ha sido previamente ubicado en el cromosoma 15 de la especie *E. guineensis*. La figura 10.1 muestra un ejemplo del tipo de variabilidad alélica observada con uno de los marcadores SSR seleccionados.

Tabla 10.1. Descripción de muestras de tejido vegetal tomadas en centros de investigación de AGROSAVIA

Especie	N.º accesiones	N.º plantas/ accesión	N.º total de muestras	Lugar de colecta
<i>E. guineensis</i> (G) Dura × Dura	20	3-6	117	C. I. Nataima (Espinal, Tolima)
<i>E. oleifera</i> (O)	43	1-8	154	C. I. El Mira (Tumaco, Nariño)
Híbridos O × G	20		197	C. I. El Mira (Tumaco, Nariño)
<i>E. guineensis</i>	11	10	109	C. I. Palmira (Valle del Cauca)
Total	94		577	

Fuente: Gallego et al. (2012)

Tabla 10.2. Características de los 27 marcadores SSR identificados como polimórficos

Nombre SSR	Secuencia cebadores F/R (5'-3')	Tipo de repetición	Temperatura anillamiento	Alelos observados (Ao)	Tamaño observado	Tamaño reportado
sEg00113	GTCACCGAACCCCTAATAAAAT	(CT) 15	54	11	150-170	n/a
	ATGCAGTTGAGGACAAAAAG					
sEg00009	TCCACTGACAACAGGACTCA	(AG) 12	52	8	198-208	n/a
	AAAAACGCATCTCAGAGAGA					
sEg00066	ACTGATGCAGGAAAGAGGAA	(AT) 8	52	10	192-220	192-215
	GAAGTACACAAGGTAAGTTCATAG					
sEg00036	GGACCCTTTTTGTACTGTTT	(AG) 9	52	10	150-170	n/a
	AGCCTACCACAACCTCCTTT					
sEg00076	GTGCTATATACTGTCACTTAAGATT	(TG) 9	52	5	133-150	n/a
	GGTAGAATATCTTCGTTTCGATT					
sEg00067	GATTAAGTCCCAACCGTCTC	(TGTA) 6	52	5	152-174	152-174
	TAAGAGAGCACGCAGTTCAG					
sEg00140	TAGAAAGTGAGACGGTGGAT	(GA) 10	53	6	224-230	216-220
	GTAATATTCTCAAGCTGGCAGT					
sEg00125	TACCCTTTCCCTCCCTCCATA	(GCG) 6	52	6	120-180	152-170
	CATCATCTCCGTTGCCAGTATT					

(Continúa)

(Continuación tabla 10.2.)

Nombre SSR	Secuencia cebadores F/R (5'-3')	Tipo de repetición	Temperatura anillamiento	Alelos observados (Ao)	Tamaño observado	Tamaño reportado
sEg00126	CCGTCTCAAAAGCCCTAAAC	(CGC) 7	52	2	212-215	212-215
	TTGTTGTCCCACTCCCTCTT					
sMo00130	TAAGCAAAAGATCAGGGCACTC	(AAG) 11	56	11	200-220	173-214
	GGCTGGTGAAAATAGGTTTACAAAG					
sMo00020	CCTTTCTCTCCCTCTCCTTTTG	(AG) 15	58	9	175-208	182-208
	CCTCCCTCCCTCTCACCATA					
sMo00027	TTACAGTTGAGGCAGTATGTCAAT	(TC) 14	50	6	200-210	n/a
	CTGTATGTCAAACCTTCTGCAC					
sMo00018	TAAATGAGAGAGAGACGAGGAC	(CT) 14	54	3	245-255	n/a
	TGGAGCCATGAGAAAGAGTA					
sMo00152	GGAACAGAGGACAAGAAAGAAA	(AC) 6 (AG) 11	56	2	220-260	245-256
	TGTATCAAGCCTCAAGTATCTGG					
sMo00134	TCCAATAGTCGTTACAAACCAG	(ATTA) 6	56	2	252-256	252-256
	GATTAGCAAAAGGGCAAAAAGG					
sMo00138	AGGGTTGTCGCTCCAATTTAT	(TTTTTC) 6	56	3	170-205	180-194
	GGCATCTTTTTGACCTGTAGAAG					
mEgCIR0008	CGGAAAGAGGGAAGATG	(GA) 18	52	11	194-218	195-220
	ACCTTGATGATTGATGTGA					
mEgCIR0221	TGCCATGTTCCAGAGAGC	(GA) 11	52	9	210-240	195-213
	TTCAGATTTTTCCGACTTC					
mEgCIR0230	CCCTGGCCCCGTTTTTC	(TA) 6 GAG (GA) 19	52	11	260-265	n/a
	AGCGCTATATGTGATTCTAA					
mEgCIR0304	CCACAAACAATCCAAGCAAGT	(GT) 4 (GC) 2 (GT) 2	52	3	125-130	275-281
	TGGCATAACGAAAGCATAA					
mEgCIR0326	GCTAACCACAGGCAAAAACA	(GT) 9	52	4	248-265	131-147
	AAGCCGCACTAACATACACATC					

(Continúa)

(Continuación tabla 10.2.)

Nombre SSR	Secuencia cebadores F/R (5'-3')	Tipo de repetición	Temperatura anillamiento	Alelos observados (Ao)	Tamaño observado	Tamaño reportado
mEgCIR0332	ATTCGTAAGGTGGGTGT	(GT) 8	52	4	140-155	269-281
	CCTCCAAACTCCTCTGT					
mEgCIR0350	AAATCCTAAATCCTAAACTC	(GT) 8	52	5	252-258	80-102
	TCTACCTGTACTGGTGACAA					
mEgCIR0353	AGAGAGAGAGAGTGCATATG	(GT) 11 (GA) 15	52	12	250-265	n/a
	GTCCCTGTGGCTGCTGTTTC					
mEgCIR0377	GTCTGCTAAAACATCAACTG	(GT) 6 (GC) 4	52	3	155-160	n/a
	GAGGAGGAGGGAACGATAA					
mEgCIR0465	TCCCCACGACCCATTC	(CCG) 6	52	5	120-144	n/a
	GGCAGGAGAGGCAGCATTC					
mEgCIR0134	AGTTTGGGGCTTACCTG	(GA) 15	52	5	130-148	n/a
	CTCCACGCACCTCTC					

Fuente: Gallego et al. (2012)

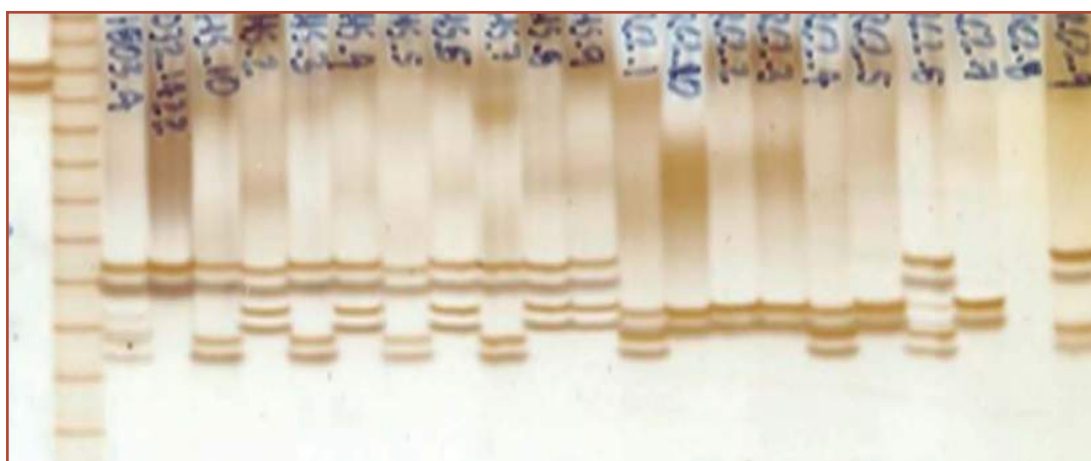


Figura 10.1. Variabilidad alélica observada para el marcador sMo00138 en parte de la población bajo estudio, correspondiente a híbridos de *E. guineensis* × *E. oleifera*. Los productos de reacción en cadena de la polimerasa (PCR, por sus siglas en inglés) fueron visualizados en gels de acrilamida al 6% (urea 5 M) de acuerdo con el protocolo de Bassam et al. (1991).

Fuente: Gallego et al. (2012)

De las 577 muestras de la población bajo estudio, 20 de *E. guineensis* procedentes del C. I. Nataima y 43 de *E. oleifera* provenientes del C. I. El Mira fueron manejadas de forma combinada (*bulks*) con individuos por accesión, ya que están constituidas por grupos de hermanos, para un total de 369 muestras analizadas. De estas, siete fueron eliminadas del análisis (entre ellas dos *bulks* de *E. oleifera*), debido al considerable número de marcadores que no lograron ser amplificados y dada la imposibilidad de confirmarlos como alelos nulos. Los datos se codificaron como tamaño en pares de bases de cada una de las bandas (alelos), las cuales fueron registradas para los dos alelos que conforman cada locus. Lo anterior se dio gracias a la naturaleza diploide de las especies del estudio, y de esta manera se pudieron identificar individuos homocigotos o heterocigotos. Los análisis se desarrollaron con el programa SAS versión 9 (Rajinder & Suan-Choo, 1998; SAS Institute Inc., 1999), en el ambiente Linux.

De los diez marcadores y las 362 muestras incluidas en el análisis estadístico, se obtuvieron 45 bandas diferentes, con un promedio de 4,5 bandas por marcador. El tamaño de los alelos varió entre 150 y 330 pares de bases, en los marcadores sEg00125 y mEgCIR230, respectivamente. La tabla 10.3 muestra la comparación de estos resultados con la información reportada en diferentes estudios.

Tabla 10.3. Comparación entre el rango de tamaño de los alelos en pares de bases (PB) reportados por otros autores y los encontrados por Gallego et al. (2012)

Locus SSR	Reportados	Referencia	Obtenidos
sEg0009	-*	Ting et al. (2010)	172-198
sEg0076	-	Ting et al. (2010)	152-158
sEg00125	154-156	Ting et al. (2010)	150-160
sEg00126	214-216	Ting et al. (2010)	214-216
sMo00018	242-252	Zaki et al. (2010)	244-258
sM0027	184-213	Zaki et al. (2010)	200-212
sM00134	252-256	Zaki et al. (2010)	252-256
sM00138	180-194	Zaki et al. (2010)	178-194
mEgCIR008	195-220	Billotte et al. (2001)	200-225
mEgCIR230	326-354	Billotte et al. (2001)	325-375

* No existen pesos en pares de bases reportados para estos marcadores.

Fuente: Gallego et al. (2012)

A partir de las frecuencias alélicas se calculó H_e , la cual representa un indicador de diversidad genética. La H_e se define como la probabilidad de tomar dos individuos al azar, y que estos no compartan ningún alelo. El valor de H_e se calculó para cada uno de los marcadores microsatélites y a nivel global, y también para los cuatro grupos principales de muestras: híbridos intraespecíficos de *E. guineensis* (Dura \times Dura), híbridos interespecíficos *E. guineensis* \times *E. oleifera*, *E. oleifera* \times *E. guineensis*, parentales *E. guineensis* y *E. oleifera* (tabla 10.4). El valor más alto se obtuvo cuando el cálculo se hizo para el total de muestras evaluadas (0,57); el más alto al interior de los grupos definidos por su estatus biológico se obtuvo en los híbridos interespecíficos (0,55), y el valor más bajo fue de 0,28 para el grupo de muestras del parental *E. oleifera*, y 0,34 para el parental *E. guineensis*. Los valores de H_e para las muestras evaluadas en este estudio demuestran una diversidad genética y un polimorfismo moderado.

En el estudio desarrollado por Ting et al. (2010) se encontraron valores de H_e entre 0,34 y 0,53 para la especie *E. guineensis*, y entre 0,21 y 0,32 para *E. oleifera*, resultados posiblemente relacionados con el uso de microsatélites génicos (EST-SSR). En contraste, los resultados derivados del trabajo de Billotte et al. (2001), en el que se emplearon marcadores exclusivamente genómicos, mostraron valores de diversidad genética para *E. guineensis* de 0,68 y 0,63, que en el caso de *E. oleifera* representan polimorfismos altos en las poblaciones estudiadas. Los valores de heterocigosidad obtenidos aparecen como intermedios entre los estudios citados antes. Estos resultados son la suma de diferentes factores presentes en la naturaleza de las poblaciones estudiadas, así como la combinación de microsatélites génicos (40 %) y genómicos (60 %) usados en la evaluación molecular.

Dentro de los resultados de caracterización genética de los materiales de palma se encontraron algunos alelos con una frecuencia de 100 % para determinados grupos biológicos (sMo00027, sEg00126 y sMo00138). Por ejemplo, para el marcador sEg00126 se encontró que el alelo 214 estuvo presente como homocigoto en todas las muestras de la especie *E. oleifera*. También se halló lo mismo para el alelo 152 (marcador sEg0076), el cual se encontró con una frecuencia de 100 % en las muestras de *E. guineensis*, así como en los híbridos interespecíficos, pero no se identificó en ningún material de *E. oleifera*. Los alelos exclusivos para ambas especies se observaron en nueve de los diez

marcadores evaluados, y estuvieron presentes en 51,1% (23) de los alelos hallados en las muestras de las poblaciones de híbridos interespecíficos.

Tabla 10.4. Heterocigosidad esperada (He)* para los diez marcadores microsatélites evaluados en 362 materiales de palma de aceite por Gallego et al. (2012)

Estatus biológico	Frecuencia	sEg00126	sMo00134	sMo00138	sEg0076	sMo00027	sEg00125	mEgCIR230	sMo00018	mEgCIR008	sEg0009	He
Total individuos	362	0,51	0,19	0,62	0,24	0,57	0,56	0,71	0,70	0,74	0,83	0,57
Híbrido <i>E. guineensis</i> (Dura × Dura)	20	0,47	0,04	0,42	0,00	0,00	0,05	0,71	0,32	0,54	0,40	0,29
Híbrido <i>E. guineensis</i> × <i>E. oleifera</i>	191	0,51	0,12	0,57	0,08	0,63	0,58	0,71	0,68	0,74	0,86	0,55
Parental <i>E. guineensis</i>	110	0,49	0,29	0,48	0,01	0,20	0,24	0,51	0,50	0,64	0,63	0,40
Parental <i>E. oleifera</i>	41	0,00	0,28	0,00	0,02	0,51	0,41	0,26	0,63	0,15	0,53	0,28

* He calculada a nivel global y por subgrupos.

Fuente: Gallego et al. (2012)

Estos resultados reflejan la posibilidad de establecer un sistema de identificación molecular de plantas de palma de aceite por medio de un grupo específico de marcadores microsatélites que exhiba un alto poder de discriminación, como el observado en algunos marcadores del presente estudio. Es importante anotar que también se identificaron nuevos alelos que no habían sido reportados previamente y que servirán como punto de partida para la elaboración de perfiles genéticos más detallados y específicos de cada especie y sus respectivos híbridos intra- e interespecíficos.

Para determinar la estructura poblacional y las características diferenciales de cada uno de los grupos evaluados, así como las posibles relaciones genéticas de los individuos y su dispersión, se utilizó un análisis de correspondencia múltiple (ACM) con el procedimiento CORRESP de SAS (SAS Institute Inc., 1999). El análisis de correspondencia múltiple relaciona diferentes perfiles moleculares de los microsatélites, para ayudar a entender la tendencia principal de los individuos observados con respecto a su genotipo. De este modo, los individuos que aparezcan muy cercanos en la gráfica compartirán un alto

número de alelos dentro de sus respectivos perfiles moleculares y, por consiguiente, los que estén muy distantes presentarán perfiles genéticos muy diferenciados en relación con su composición alélica.

Así, el método multivariado de ACM permitió diferenciar de manera eficiente las muestras de la especie *E. oleifera*, así como las correspondientes a la especie *E. guineensis* y los híbridos interespecíficos (figura 10.2a). Estos tres grupos principales, ubicados espacialmente dentro del plano tridimensional, mostraron subagrupaciones que pueden estar relacionadas con características específicas que permiten establecer diferencias dentro de un determinado estatus biológico o especie.

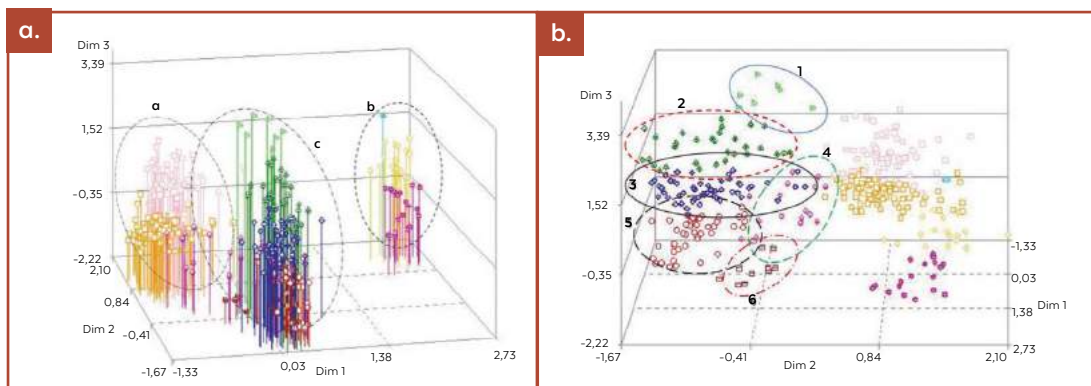


Figura 10.2. ACM de datos genotípicos que muestran la distribución espacial de 362 muestras de palma de aceite. a. Distribución espacial de las 362 muestras: a) individuos de *E. guineensis*, b) individuos de *E. oleifera* y c) materiales híbridos interespecíficos y retrocruces; b. Comportamiento de los híbridos interespecíficos. Los subgrupos de cada grupo se representan en diferentes colores

Fuente: Gallego et al. (2012)

En el grupo b, correspondiente a las muestras de *E. oleifera*, se encontraron tres subgrupos (figura 10.2a), uno de los cuales estaba constituido por una única muestra (*bulk* 46, señalado en azul), que mostró características muy diferenciadas a nivel de perfil molecular. El grupo a, conformado por los individuos de *E. guineensis*, se dividió en dos subgrupos bien diferenciados (figura 10.2a), y un grupo adicional (señalado en morado en el grupo b) apareció como puente genético entre los individuos híbridos (grupo c) y los materiales de la especie *E. guineensis* (grupo a). Estos individuos corresponden a los retrocruces que se desarrollaron con el parental *E. guineensis*, y el resultado refuerza la validación

de los datos ya que indica que estas muestras se acercan más al progenitor elegido. En los híbridos interespecíficos, se observó un nivel de subdivisión considerable, con la diferenciación de seis grupos (figura 10.2b).

Dentro del mismo análisis de ACM se pudo visualizar la relación entre la información molecular generada y los datos fenotípicos que habían sido recopilados por AGROSAVIA para cada uno de los materiales (figura 10.3). Las características analizadas fueron producción de aceite linoleico, oleico, palmítico, carotenos, color del polen, crecimiento de la planta, grupo de polen, olor del polen, partenogénesis, resistencia a PC, tipo de cuesco y viabilidad del polen. Los resultados de tres características —crecimiento de la palma (figura 10.3a), producción de aceite (figura 10.3b) y resistencia a PC (figura 10.3c)— indican la distribución espacial derivada de los datos moleculares y el modo como esas agrupaciones se comportan con cada variable fenotípica, lo cual permite inferir la razón por la cual se separan o agrupan las poblaciones del estudio. Este ejercicio no obedece a un análisis de correspondencia entre datos moleculares y datos fenotípicos, ya que la posición de la muestra en la figura 10.3 está dada en todos los casos por la información molecular.

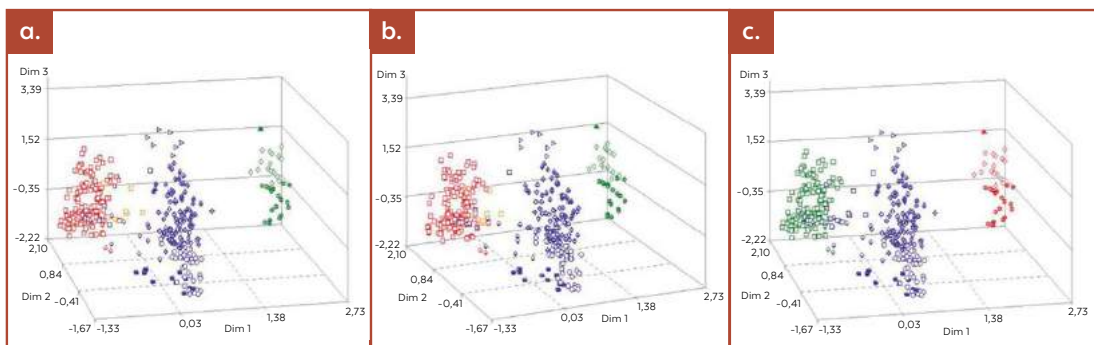


Figura 10.3. ACM de datos moleculares en el que se marcaron las muestras por colores para las variables (*E. guineensis* se muestra a la izquierda de cada figura [color rojo en a y b, y verde en c], *E. oleifera* a la derecha [color verde en a y b, y rojo en c] y los híbridos y retrocruzas en el intermedio [colores azul y naranja, respectivamente]).
a. Crecimiento; b. Producción de aceite; c. Pudrición del cogollo (PC).

Fuente: Gallego et al. (2012)

Así, a partir del análisis de ACM generado únicamente con los datos moleculares fue posible identificar cómo están distribuidas las muestras en relación con diferentes variables fenotípicas y datos de pasaporte, por asignación de colores. Los parentales *E. guineensis* exhiben elevada tasa de crecimiento y altos porcentajes de producción de aceite (color rojo) y susceptibilidad a PC (color verde); los híbridos interespecíficos presentan tasa de crecimiento y producción de aceite medias, así como resistencia a PC (color azul); los individuos de las retrocruzas (RC2) muestran tasa de crecimiento y porcentaje de aceite medio-alto (color naranja), y los parentales *E. oleifera* presentan tasa de crecimiento y producción de aceite bajas (color verde), así como resistencia a PC (color rojo) (figura 10.3).

Como parte de un convenio de agenda interna de AGROSAVIA, se continuó la caracterización molecular en una nueva población del parental *E. oleifera* (seis individuos) y de los híbridos O × G (90 individuos) colectados en el C. I. El Mira, con cinco marcadores tipo SSR (Barrero et al., 2013a) adicionales a los diez descritos anteriormente en el estudio de Gallego et al. (2012). A continuación, se describen brevemente los resultados generados con estos cinco marcadores.

Para las pruebas de estandarización se evaluaron inicialmente 45 SSR en tres individuos híbridos y tres parentales mediante corrido electroforético en gel de poliacrilamida. A partir de esto, se seleccionaron nueve SSR polimórficos, que fueron evaluados en el total de la población de 96 individuos; de estos, cinco fueron seleccionados para análisis posteriores por representar alto polimorfismo en la población híbrida (sEg00032, sEg00090, sEg00140, sMo00129 y sMo00132). Un marcador adicional evaluado previamente por Gallego et al. (2012) (sEg00009) (tabla 10.4) fue usado como control. El número de alelos por locus presentó un rango entre 2 y 3, con un promedio de 2,67 (tabla 10.5). El PIC presentó un rango entre 0,57 para sMo00129 y 0,23 para sEg0140, lo cual concuerda con la presencia de un menor número de alelos en este último. La heterocigosidad observada (H_o) en la población presentó un valor alto, con promedio de 0,59, mientras que la H_e presentó un valor promedio de 0,50, y fue el marcador sMo00129 el que obtuvo los valores más altos ($H_e = 0,64$).

Tabla 10.5. Resumen del análisis de variabilidad de cinco nuevos SSR y un control, de acuerdo con Barrero et al. (2013a)

Marcador	Secuencia del cebador dirección 5-3	Referencia	Frecuencia alélica	N.º alelos	Heterocigosidad esperada (He)	Heterocigosidad observada (Ho)	PIC+
sEg0009	F-TCCACTGACAACAGGACTCA R-AAAAACGCATCTCAGAGAGA	Ting et al. (2010)	0,53	2	0,50	0,94	0,37
sMo00129	F-TTAGTATTGGGTGTGCATAAGTGG R-GCTTCCAGCTCCTCTTTCTACC	Zaki et al. (2010)	0,47	3	0,64	0,95	0,57
sMo00132	F-ATAGCCAGAGGGCAAACACTGT R-GCAACACACGGACTCAAAACTA	Zaki et al. (2010)	0,48	3	0,54	0,08	0,43
sEg00090	F-TCAGAAATGCCTACATCAAAC R-AGGGACACGAGAATACATACA	Ting et al. (2010)	0,65	3	0,50	0,58	0,44
sEg00032	F-CTGTTGAGCTGGAGAGACCC R-CCAACCAGGATCAGTTTGGT	Ting et al. (2010)	0,50	3	0,54	0,74	0,44
sEg0140	F-TAGAAAGTGAGACGGTGGAT R-GTAATATTCTCAAGCTGGCAGT	Ting et al. (2010)	0,84	2	0,26	0,31	0,23
Promedio			0,58	2,67	0,50	0,60	0,41

* Ho y He calculadas con el software PowerMarker (Liu & Muse, 2005) + PIC calculado de acuerdo con Botstein et al. (1980).
Fuente: Barrero et al. (2013a)

Al comparar estos resultados con el estudio de Gallego et al. (2012), se encuentra que el número promedio de alelos por locus es menor (2,67 versus 4,5) y que la He para la población total es ligeramente menor (0,50 versus 0,57), lo cual posiblemente se relaciona con el menor tamaño poblacional y con las diferencias inherentes a los individuos y marcadores analizados. Al comparar con otros estudios, como el de Billotte et al. (2001), se encuentra que el número promedio de alelos por locus es también menor (2,67 versus 5,25). En este último trabajo se evaluó un alto número de materiales de diferentes procedencias y especies relacionadas como *Bactris gasipaes*. Al comparar estos resultados con los reportados por Bakoumé et al. (2007),

quienes encontraron un número total de 209 alelos, un promedio de 3,3 alelos por locus en 16 SSR y 494 individuos procedentes de diez países africanos, se observa que los valores son similares (2,67 versus 3,3). Igualmente, el número promedio de alelos es parecido al obtenido por Singh et al. (2008), con un valor de 2,6, el cual es consistente con el tipo de población evaluada, ya que los cultivares mejorados tienden a presentar un bajo número de alelos.

Por otro lado, el valor promedio de H_e (0,50) es similar al obtenido por Montoya et al. (2005) en la exploración preliminar del germoplasma de Angola llevada a cabo en Cenipalma ($H_e = 0,45$). Además, estudios efectuados con isoenzimas (Hayati et al., 2004) y RFLP (Maizura et al., 2006), que incluyen germoplasma de varios países africanos, reportan valores de H_e menores (0,18 y 0,19, respectivamente), al igual que en el estudio de Singh et al. (2008), en el que se obtiene un $H_e = 0,36$ con el uso de marcadores EST-SSR.

Por el contrario, con locus microsatélites, Cochard et al. (2009) reportaron una H_o promedio de 0,63, comparable al valor obtenido en el estudio de Barrero et al. (2013a) ($H_o = 0,60$, tabla 10.5). De lo anterior se puede inferir que el número de alelos detectados, así como la diversidad genética, dependen de la procedencia del germoplasma, el tamaño poblacional, el tipo de marcadores y el estado de mejoramiento.

Para conocer las relaciones genéticas de la población de estudio, se generó un dendrograma a partir de las frecuencias alélicas, que muestra cuatro grupos. El grupo 1 (rojo) está conformado por los seis genotipos maternos *E. oleifera*, mientras que los otros tres están compuestos por los individuos híbridos: el grupo 2 (azul), de 35 genotipos con mayor similitud al parental materno; el grupo 3 (negro), de siete genotipos con similitud intermedia, y el grupo 4 (verde), de 48 genotipos, posiblemente más relacionados con el padre *E. guineensis* (figura 10.4a). Estos resultados fueron soportados por el análisis de componentes principales (ACP), en el que los primeros componentes explican el 35,71 % de la variación presente en la población de estudio (figura 10.4b).

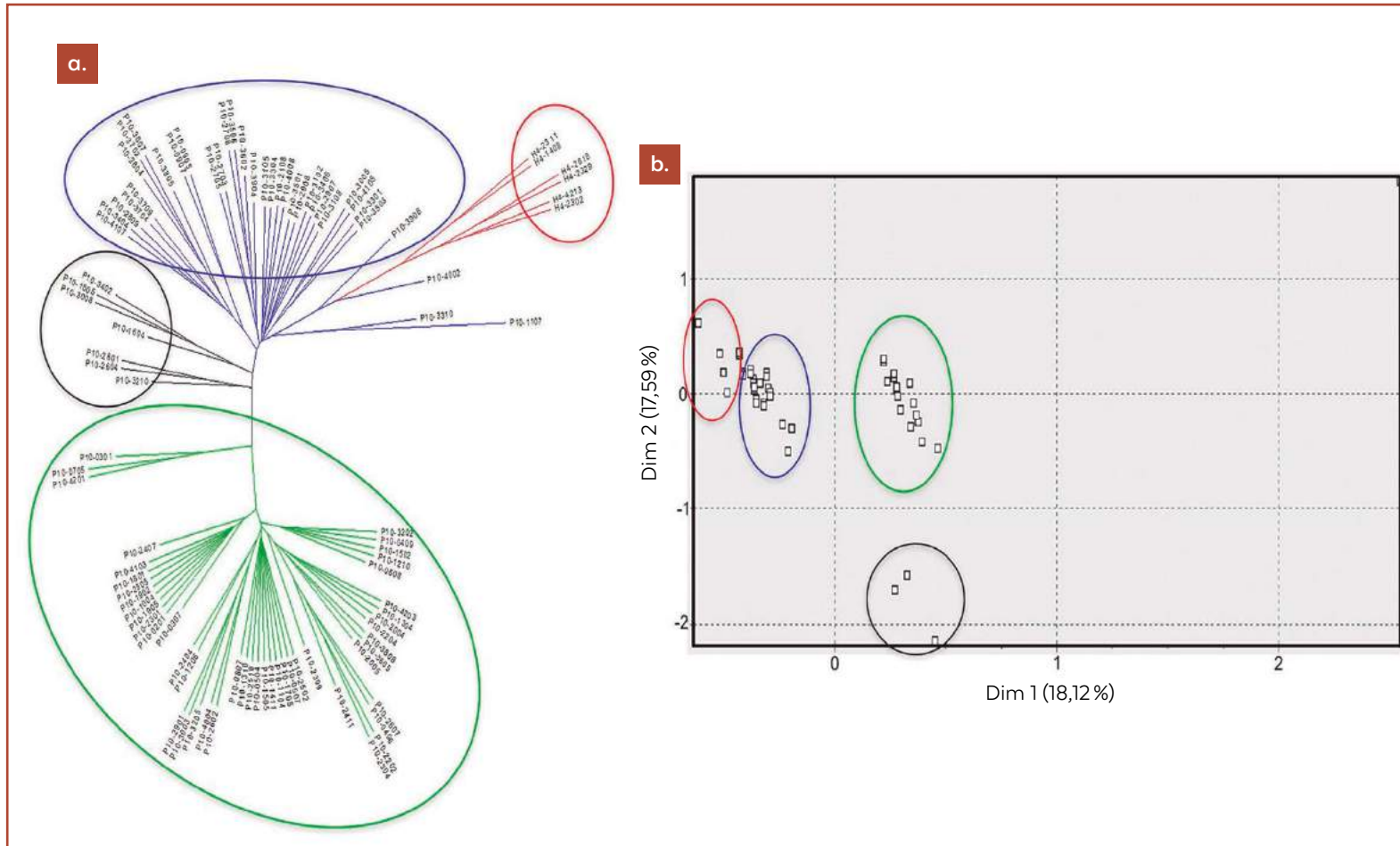


Figura 10.4. Análisis *neighbor-joining* y de componentes principales (ACP) en 96 genotipos de palma de aceite. a. Dendrograma aplicado al algoritmo *neighbor-joining*, basado en la matriz de distancias de Nei (1972) y en un remuestreo de 1.000 réplicas (PHYLIP, versión 3.5c) (Felsenstein, 1989); b. ACP, que muestra las dos primeras coordenadas (Genetix, versión 4.05) (Belkhir et al., 2004). En rojo se muestran los individuos de *E. oleifera*, mientras que los híbridos se presentan en tres grupos de otros colores.

Fuente: Barrero et al. (2013a)

De esta forma, los cinco nuevos locus microsatélites permitieron discriminar tres poblaciones con similitud intermedia al parental materno *E. oleifera* y con posible similitud al padre *E. guineensis*. Así, con base en los estudios anteriormente mencionados (Barrero et al., 2013a; Gallego et al., 2012), se puede concluir que fue posible identificar un total de quince marcadores SSR adecuados para discriminar las poblaciones híbridas generadas por el programa de mejoramiento de AGROSAVIA.

Marcadores SNP

Los anteriores estudios indican que es posible establecer un sistema de identificación molecular de las plantas de palma de aceite para futuros estudios de mejoramiento, diversidad genética y conservación, mediante el uso de marcadores altamente informativos como los SSR. A futuro también se espera incluir poblaciones con un mayor número de individuos y con marcadores moleculares más abundantes en el genoma, como los SNP, para la selección de poblaciones diferenciales en etapa de vivero y la asociación con el fenotipo de interés (por ejemplo, selección genómica o mapeo por asociación), como apoyo al mejoramiento del cultivo.

En relación con lo anterior, AGROSAVIA, en convenio con la Universidad de Cornell, EE. UU. (Barrero et al., 2013b), procesó nuevas poblaciones F1 O×G en fase de vivero (380 individuos preliminares) y posibles parentales (27 individuos), mediante GBS, lo que permitió obtener miles de marcadores SNP, ideales para estudios de diversidad genética y de asociación genotipo (variante alélica o SNP) con fenotipo (Elshire et al., 2011).

Como producto de la genotipificación vía GBS, se identificaron 131,825 SNP en un panel de 407 muestras, los cuales fueron filtrados por desequilibrio de ligamiento (D_L) para eliminar marcadores redundantes y frecuencia mínima de alelos ($MAF = 0,05$) y así se obtuvo un set de 3,506 SNP de alta calidad, con un porcentaje de datos perdidos de 18,75 %. Con estos marcadores se realizó un análisis de componentes principales en el paquete estadístico R (Dalgaard, 2010), en el que los tres primeros componentes explicaron 15,47 % de la variación presente en la población de estudio (figura 10.5).

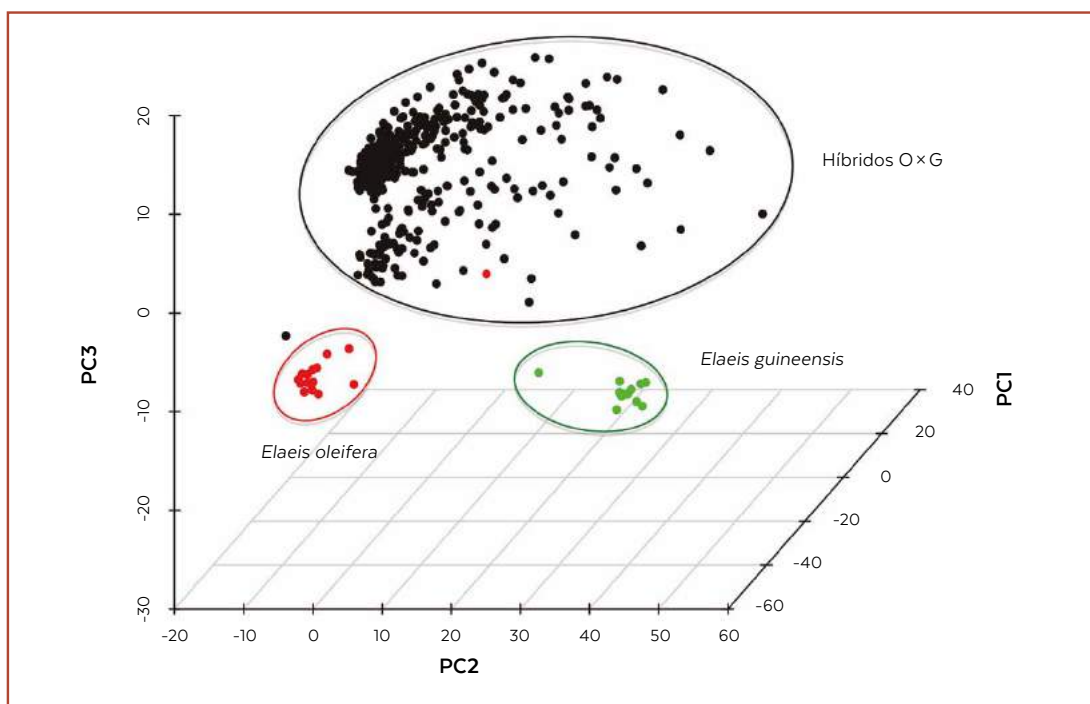


Figura 10.5. Distribución espacial de los tres primeros componentes principales en la población de híbridos O × G y parentales. Se observa una clara separación entre los parentales y los híbridos.

Fuente: Elaboración propia

Adicionalmente, en el análisis de distancias genéticas, con el método de agrupación del vecino más cercano (*neighbor joining*), implementado en el software TASSEL V3 (Bradbury et al., 2007), se observó un claro patrón de tres grandes grupos. Dos estaban relacionados por su afinidad con cada uno de los parentales, y dentro de ellos también pudieron diferenciarse claramente varios subgrupos de los híbridos O × G (figura 10.6). El gran número de marcadores obtenido permite diferenciar, con un alto grado de confianza, los grupos de poblaciones segregantes, en relación con su afinidad a los parentales. A futuro, en la medida en que estas poblaciones segregantes se evalúen para fenotipos de interés como la PC, será posible realizar estudios de asociación rasgo-marcador para identificar regiones del genoma relacionadas con estos rasgos y seleccionar progenies con atributos de interés en etapas de previvero.

Un ejemplo de esto se ilustra en el estudio de asociación de genoma completo, genotipo (GBS)-fenotipo de Osorio-Guarín et al. (2019), en el que se usó una población de 378 individuos del híbrido O × G generado por AGROSAVIA y

se identificaron doce regiones genómicas (SNP), localizadas cerca de genes candidatos y asociadas con la arquitectura de la planta y la producción, como potenciales marcadores para selección asistida del programa de mejoramiento. Además, recientemente Garzón-Martínez et al. (2022) evaluaron 431 híbridos O x G y 444 retrocruces (BC1) generados por AGROSAVIA para características morfológicas y de producción, mediante estudios de predicción genómica. Se encontraron altos grados de predicción para el área foliar, la altura del tallo y la producción —esta última con el uso de modelos multicares—, lo cual confirmó el potencial de GS para selección de parentales y predicción de valores genéticos.

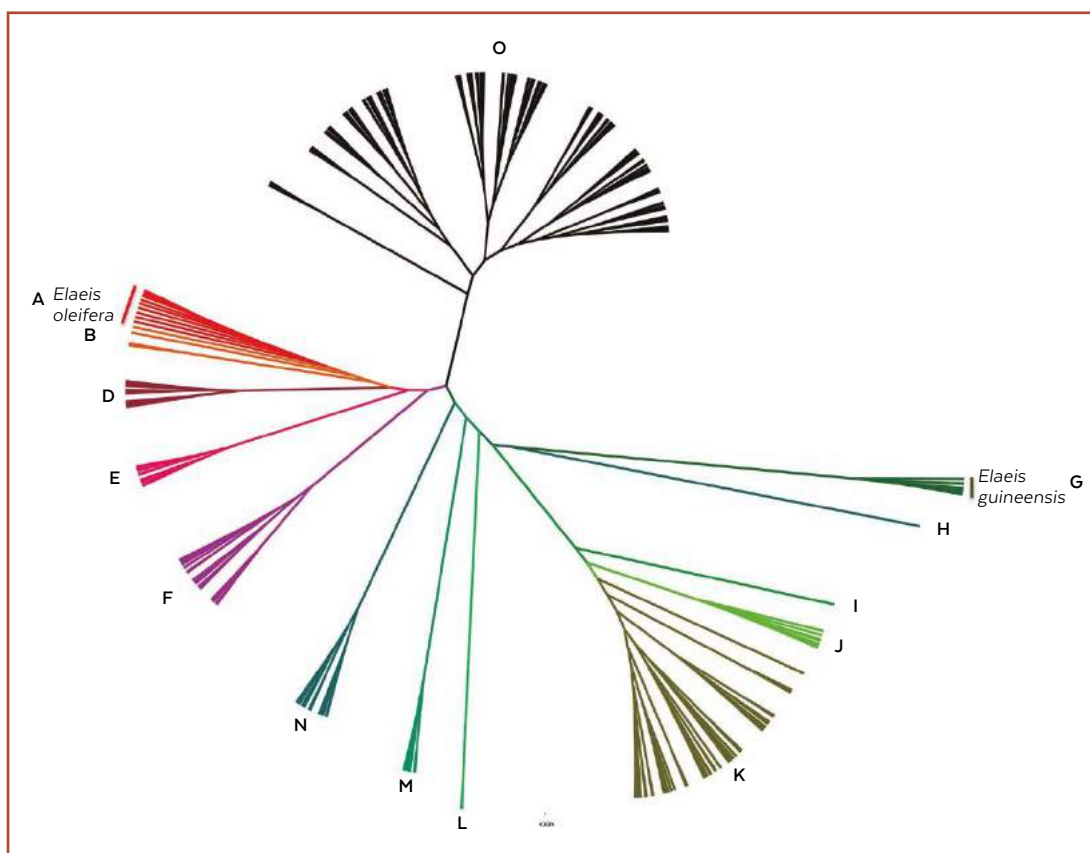


Figura 10.6. Árbol de distancias genéticas generado con el método de agrupamiento de *neighbor joining*, en el que se observan tres grupos, dos que conforman subgrupos cercanos al parental resistente *E. oleifera* (A [subgrupos B, D, E y F]) y al parental susceptible *E. guineensis* (G [subgrupos H, I, J, K, L M y N]), y un tercer grupo intermedio, situado entre los dos parentales (O).

Fuente: Elaboración propia con base en Barrero et al. (2013a)

Estas plataformas pueden alinearse con sistemas de marcadores de última generación recientemente implementados o en proceso de implementación, para el eficaz mejoramiento genético de la especie. Dado lo anterior, se espera que en un futuro cercano se puedan establecer acercamientos entre instituciones nacionales e internacionales para avanzar eficientemente en el mejoramiento genético de este cultivo en Colombia.

Referencias

- Arias, D., & Rocha, P. (2004). Análisis de diversidad genética en materiales tolerantes y susceptibles a la pudrición de cogollo en palma de aceite mediante marcadores moleculares. *Palmas*, 25(3), 11-27. <https://publicaciones.fedepalma.org/index.php/palmas/article/view/1004>
- Arias, D., & Núñez, V. M. (2006). *Caracterización molecular de genotipos de palma mejorados, utilizados en el programa de mejoramiento de palma de aceite del centro experimental El Mira* [Informe final, Programa de Recursos Genéticos Vegetales y Biotecnología]. Corporación Colombiana de Investigación Agropecuaria (Corpoica).
- Bakoumé, C., Wickneswari, R., Rajanaidu, N., Kushairi, A., Amblard, P., & Billotte, N. (2007). Allelic diversity of natural oil palm (*Elaeis guineensis* Jacq.) populations detected by microsatellite markers: Implications for conservation. *Plant Genetic Resources*, 5(2), 104-107. <https://doi.org/10.1017/S1479262107710870>
- Barcelos, E. (1998). *Etude de la diversité génétique du genre Elaeis (E. oleifera [Kunth] Cortés et E. guineensis Jacq.) par marqueurs moléculaires (RFLP et AFLP)* [Thèse de docteur] Université de Montpellier II.
- Barrera, S. N. (2011). *Estructura genética de poblaciones naturales de palma de aceite (Elaeis guineensis Jacq.) procedentes de Angola* [Tesis de maestría, Universidad Nacional de Colombia, sede Bogotá]. <https://repositorio.unal.edu.co/handle/unal/7835>
- Barrero, L. S., Bastidas, S., Duque, L. M., Pérez, J. I., & Delgadillo, L. P. (2013a). *Nuevos (adicionales) marcadores moleculares desarrollados y tres poblaciones de trabajo identificadas del híbrido O x G con contenido genético similar o intermedio a los parentales* [Informe final, Convenio Agenda Interna 211, Corpoica-MADR].

- Barrero, L. S., Bastidas, S., Duque, L. M., Pérez, J. I., & Delgadillo, L. P. (2013b). *Población seleccionada y procesada de 95 individuos del híbrido O×G y sus parentales, pertenecientes al programa de mejoramiento genético en palma de aceite de Corpoica para genotipificación* [Informe final, Convenio 211 Corpoica-Universidad de Cornell-MADR].
- Bassam, B. J., Caetano-Anollés, G., & Gresshoff, P. M. (1991). Fast and sensitive silver staining of DNA in polyacrylamide gels. *Analytical biochemistry*, 196(1), 80-83. [https://doi.org/10.1016/0003-2697\(91\)90120-1](https://doi.org/10.1016/0003-2697(91)90120-1)
- Bastidas, S., Peña, E., Reyes, R., Pérez, J., & Tolosa, W. (2007). Comportamiento agronómico del cultivar híbrido RC1 de palma de aceite (*Elaeis oleifera* × *Elaeis guineensis*) × *Elaeis guineensis*. *Revista Corpoica Ciencia y Tecnología Agropecuaria*, 8(1), 5-11. https://doi.org/10.21930/rcta.vol8_num1_art:77
- Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N., & Bonhomme, F. (2004). *GENETIX 4.05, logiciel sous Windows™ pour la génétique des populations*. Université de Montpellier II. <https://kimura.univ-montp2.fr/genetix/>
- Billote, N., Lagoda, P. J. L., Risterucci, A. M., & Baurens, F. C. (1999). Microsatellite-enriched libraries: Applied methodology for the development of SSR markers in tropical crops. *Fruits*, 54(4), 277-288.
- Billotte, N., Marseillac, N., Risterucci, A. M., Adon, B., Brottier, P., Baurens, F. C., Singh, R., Herrán, A., Asmady, H., Billot, C., Amblard, P., Durand-Gasselin, T., Courtois, B., Asmono, D., Cheah, S. C., Rohde, W., Ritter, E., & Charrier, A. (2005). Microsatellite-based high density linkage map in oil palm (*Elaeis guineensis* Jacq.). *Theoretical and Applied Genetics*, 110, 754-765. <https://doi.org/10.1007/s00122-004-1901-8>
- Billotte, N., Risterucci, A. M., Barcelos, E., Noyer, J. L., Amblard, P., & Baurens, F. C. (2001). Development, characterisation, and across-taxa utility of oil palm (*Elaeis guineensis* Jacq.) microsatellite markers. *Genome*, 44, 413-425. <https://doi.org/10.1139/g01-017>
- Botstein, D., White, R. L., Skolnick, M., & Davis, R. W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet*, 32(3), 314-331. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1686077/?page=1>
- Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., & Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics*, 23(19), 2633-2635. <https://doi.org/10.1093/bioinformatics/btm308>

- Cochard, B., Adon, B., Rekima, S., Billotte, N., Chenon, R. B., Koutou, A., Nouvy, B., Omoré, A., Razak Purba, A., Glazsmann, J. C., & Noyer, J. L. (2009). Geographic and genetic structure of African oil palm diversity suggests new approaches to breeding. *Tree Genetics & Genomes*, 5(3), 493-504. <https://doi.org/10.1007/s11295-009-0203-3>
- Cook, R. (2023). *Ranking of countries that produce the most oil palm fruit (FAO)*. <https://beef2live.com/story-ranking-countries-produce-oil-palm-fruit-fao-204-215761>
- Cros, D., Bocs, S., Riou, V., Ortega-Abboud, E., Tisné, S., Argout, X., Pomiès, V., Nodichao, L., Lubis, Z., Cochard, B., & Durand-Gasselín, T. (2017). Genomic preselection with genotyping-by-sequencing increases performance of commercial oil palm hybrid crosses. *BMC Genomics*, 18, 839. <https://doi.org/10.1186/s12864-017-4179-3>
- Cros, D., Denis, M., Sanchez, L., Cochard, B., Edyana-Suryana, H., Omoré A., Durand-Gasselín, T., & Bouvet, J.M. (2013). *Factors controlling accuracy of genomic selection in oil palm (Elaeis guineensis)* [Conferencia]. Plant and Animal Genome Conference, PAG XXI. San Diego, Estados Unidos.
- Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., & Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One*, 6, 319-379. <https://doi.org/10.1371/journal.pone.0019379>
- Federación Nacional de Cultivadores de Palma [Fedepalma]. (2013). *Biodiésel de palma colombiano: de la ficción energética a la realidad de un negocio*.
- Felsenstein, J. (1989). PHYLIP: Phylogeny inference package: Versión 3.2. *The Quarterly Review of Biology*, 64(4), 164-166. <https://doi.org/10.1086/416571>
- Galeano, C. H. (2005). Estandarización de la técnica molecular de AFLP en palma de aceite tipo Dura (*Elaeis guineensis* Jacq.) y estudio preliminar de caracterización. *Agronomía Colombiana*, 23(1), 42-49. <http://www.redalyc.org/articulo.oa?id=180316951006>
- Gallego, G., Barrero, L. S., Suárez, H., Bastidas, S., & Hernández, E. G. (2012). *Aplicación de herramientas biotecnológicas para potenciar el desarrollo regional: parentales e híbridos O x G de palma de aceite caracterizados molecularmente (SSR)* [Informe final, Convenio 072 Corpoica-CIAT-MADR].
- Garzón-Martínez, G. A., Osorio-Guarín, J. A., Moreno, L. P., Bastidas, S., Barrero, L. S., López-Cruz, M., & Enciso-Rodríguez, F. E. (2022). Genomic selection for morphological and yield-related traits using genome-wide SNPs in oil palm. *Molecular Breeding*, 42, 71. <https://doi.org/10.1007/s11032-022-01341-5>

- Hayati, A., Wickneswari, R., Maizura, I., & Rajanaidu, N. (2004). Genetic diversity of oil palm (*Elaeis guineensis* Jacq.) germplasm collections from Africa: Implications for improvement and conservation of genetic resources. *Theoretical and Applied Genetics*, 108, 1274-1284. <https://doi.org/10.1007/s00122-003-1545-0>
- Hernández, V., Barrero, L. S., & Narváez, J. (1997). Estandarización de una metodología para la electroforesis de proteínas totales de palmas de aceite. En *V Congreso de la Sociedad Colombiana de Fitomejoramiento y Producción de Cultivos* (p. 105). Produmedios.
- Kularatne, R. S., Shah, F. H., & Rajanaidu, N. (2001). The evaluation of genetic diversity of Deli dura and African oil palm germplasm collection by AFLP technique. *Tropical Agricultural Research*, 13, 1-12. <http://192.248.43.153/bitstream/1/1596/2/PGIATAR-13-1.pdf>
- Liu, J., & Muse, S. V. (2005). PowerMarker: An integrated analysis environment for genetic marker data. *Bioinformatics*, 21(9), 2128-2129. <https://doi.org/10.1093/bioinformatics/bti282>
- Maizura, I., Rajanaidu, N., Zakri, A. H., & Cheah, S. C. (2006). Assessment of genetic diversity in oil palm (*Elaeis guineensis* Jacq.) using Restriction Fragment Length Polymorphism (RFLP). *Genetic Resources and Crop Evolution*, 53, 187-195. <https://doi.org/10.1007/s10722-004-4004-0>
- Mayes, S., Hafeez, F., Price, Z., MacDonald, D., & Billotte, N. (2008). Molecular research in oil palm, the key oil crop for the future. In P. Moore & R. Ming (eds), *Genomics of tropical crop plants* (pp. 371-404). Springer. https://doi.org/10.1007/978-0-387-71219-2_15
- Mayes, S., Jack, P. L., & Corley, R. H. V. (2000). The use of molecular markers to investigate the genetic structure of an oil palm breeding programme. *Heredity*, 85(3), 288-293. <https://doi.org/10.1046/j.1365-2540.2000.00758.x>
- Mayes, S., Jack, P. L., Corley, R. H., & Marshall, D. F. (1997). Construction of a RFLP genetic linkage map for oil palm (*Elaeis guineensis* Jacq.). *Genome*, 40, 116-122. <https://doi.org/10.1139/g97-016>
- Montoya, C., Arias, D., Rey, L., & Rocha, P. J. (2005). Caracterización molecular de materiales *Elaeis guineensis* Jacq. procedentes de Angola. *Fitotecnia Colombiana*, 5, 1-10. https://www.researchgate.net/publication/236091128_CHARACTERIZACION_MOLECULAR_DE_MATERIALES_Elaeis_guineensis_Jacq_PROCEDENTES_DE_ANGOLA
- Moretzsohn, M. C., Ferreira, M. A., Amaral, Z. P. S., Coelho, P. J. A., Grattapaglia, D., & Ferreira, M. E. (2002). Genetic diversity of Brazilian oil palm (*Elaeis oleifera* H. B. K.) germplasm collected in the Amazon Forest. *Euphytica*, 124, 35-45. <https://doi.org/10.1023/A:1015606304653>

- Nei, M. (1972). Genetic distance between populations. *The American Naturalist*, 106(949), 283-292. <https://doi.org/10.1086/282771>
- Ochoa, I., Villegas, V., & Beebe, S. (1997). Identification of RAPD molecular markers associated with resistance to the bud rot complex in oil palm (*Elaeis guineensis* Jacq.). *Revista Palmas*, 18, 33-38. <https://publicaciones.fedepalma.org/index.php/palmas/article/download/581/581>
- Ong, P. W., Maizura, I., Abdullah, N. A. P., Rafii, M. Y., Ooi, L. C. L., Low, E. T. L., & Singh, R. (2015). Development of SNP markers and their application for genetic diversity analysis in the oil palm (*Elaeis guineensis*). *Genetics and Molecular Research*, 14(4), 12205-12216. <https://doi.org/10.4238/2015.October.9.9>
- Osorio-Guarín, J. A., Garzón-Martínez, G. A., Delgadillo-Durán, P., Bastidas, S., Moreno, L. P., Enciso-Rodríguez, F. E., Cornejo, O. E., & Barrero, L. S. (2019). Genome-wide association study (GWAS) for morphological and yield-related traits in an oil palm hybrid (*Elaeis oleifera* × *Elaeis guineensis*) population. *BMC Plant Biology*, 19, 533. <https://doi.org/10.1186/s12870-019-2153-8>
- Price, Z., Schulman, A. H., & Mayes, S. (2003). Development of new marker methods—an example from oil palm. *Plant Genetic Resources*, 1(2-3), 103-113. <https://doi.org/10.1079/PGR200317>
- Purba, A. R., Noyer, J. L., Baudouin, L., Perrier, X., Hamon, S., & Lagoda, P. J. L. (2000). A new aspect of genetic diversity of Indonesia oil palm (*Elaeis guineensis* Jacq.) revealed by isoenzyme and AFLP markers and its consequences for breeding. *Theoretical and Applied Genetics*, 101(5-6), 956-961. <https://doi.org/10.1007/s001220051567>
- Dalgaard, P. (2010). *The R Project for Statistical Computing*. <http://www.R-project.org/>
- Rajinder, S., & Suan-Choo, A. R. (1998). Generation of molecular markers in oil palm (*Elaeis guineensis*) using AFLPTM analysis. *Plant Biotechnology*, 20, 26-27.
- SAS Institute Inc. (1999). *SAS/STAT User's Guide 9.1*. SAS Institute Inc. https://support.sas.com/documentation/onlinedoc/91pdf/sasdoc_91/stat_ug_7313.pdf
- Shah, F. H., Rashid, O., Simons, A. J., & Dunsdon, A. (1994). The utility of RAPD markers for the determination of genetic variation in oil palm (*Elaeis guineensis*). *Theoretical and Applied Genetics*, 89, 713-718. <https://doi.org/10.1007/BF00223710>

- Singh, R., Ong-Abdullah, M., Leslie Low, E. T., Manaf, M. A. A., Rosli, R., Nookiah, R., Ooi, L. C. L., Ooi, S. E., Chan, K. L., Halim, M. A., Azizi, N., Nagappan, J., Bacher, B., Lakey, N., Smith, S. W., He, D., Hogan, M., Budiman, M. A., Lee, E. K. ... Sambanthamurthi, R. (2013). Oil palm genome sequence reveals divergence of interfertile species in old and new worlds. *Nature*, 500, 335-339. <https://doi.org/10.1038/nature12309>
- Singh, R., Zaki, N., Ting, N. C., Rosli, R., Tan, S. G., Leslie Low, E. T., Ithnin, M., & Cheah, S. C. (2008). Exploiting an oil palm EST database for the development of gene-derived SSR markers and their exploitation for assessment of genetic diversity. *Biologia*, 63, 227-235. <https://doi.org/10.2478/s11756-008-0041-z>
- Ting, N. C., Zaki, N. M., Rosli, R., Low, E. T. L., Ithnin, M., Cheah, S. C., Tan, S. G., & Singh, R. (2010). SSR mining in oil palm EST database: Application in oil palm germplasm diversity studies. *Journal of genetics*, 89, 135-145. <https://doi.org/10.1007/s12041-010-0053-7>
- Ting, N.-C., Jansen, J., Mayes, D., Massawe, F., Sambanthamurthi, R., Cheng-Li Ooi, L., Weng Chin, C., Arulandoo, X., Seng, T.-Y., Syed Alwee, S. S. R., Ithnin, M., & Singh, R. (2014). High density SNP and SSR-based genetic maps of two independent oil palm hybrids. *BMC Genomics*, 15, 309. <https://doi.org/10.1186/1471-2164-15-309>
- Turner, E. C., Snaddon, J. L., Fayle, T. M., & Foster, W. A. (2008). Oil palm research in context: Identifying the need for biodiversity assessment. *PLoS One*, 3, e1572. <https://doi.org/10.1371/journal.pone.0001572>
- Villegas, V., Durán, C., & Beebe, S. (2000). Caracterización molecular de materiales Dura. *Revista Palmas*, 21(número especial), 35-40. <https://publicaciones.fedepalma.org/index.php/palmas/article/view/758>
- Vives, A. M., Barrero, L. S., & Narváez, J. (1997). Análisis comparativo de los perfiles electroforéticos de isoenzimas de las especies aceiteras *Elaeis guineensis* y *E. oleifera*. En *V Congreso de la Sociedad Colombiana de Fitomejoramiento y Producción de Cultivos* (p. 104). Produmedios.
- Zaki, N. M., Ismail, I., & Rosli, R. (2010). Development and characterization of *Elaeis oleifera* microsatellite markers. *Sains Malaysiana*, 39(6), 909-912. <http://journalarticle.ukm.my/7432/>
- Zaki, N. M., Singh, R., Rosli, R., & Ismail, I. (2012). *Elaeis oleifera* genomic-SSR markers: Exploitation in oil palm germplasm diversity and cross-amplification in Arecaceae. *International Journal of Molecular Sciences*, 13, 4069-4088. <https://doi.org/10.3390/ijms13044069>