

# Diversidad y estructura genética de accesiones de palma de aceite (*Elaeis guineensis* Jacq.) provenientes de Camerún

## Diversity and genetic structure of oil palm accession (*Elaeis guineensis* Jacq.) from Cameroon

Estiben Pacheco\*, Diana Arias\*\*, Zaida Ojeda\*\*\*, Hernán Romero\*\*\*\*

### Resumen

La palma de aceite *Elaeis guineensis* Jacq. posee gran importancia debido al aceite que se extrae de sus frutos, del cual se obtienen derivados refinados de gran valor comercial como el biodiesel, entre otros. Esta investigación buscó determinar la estructura y diversidad genética de 311 muestras de palma de aceite proveniente de la República de Camerún mediante 10 marcadores microsatélite. Los resultados mostraron valores promedio para el número medio de alelos por locus de  $N_a=8.433$  y un número efectivo de alelos por locus de  $N_e=4.756$ ; las diferencias entre estos valores permiten inferir que los 106 alelos encontrados para estas poblaciones son considerados alelos raros. Adicionalmente, el valor de diversidad genética fue alto (valor medio de  $H_e=0.781$ ) respecto a reportes de varios autores. La varianza molecular obtenida evidenció que el mayor porcentaje (80 %) se encuentra dentro de los individuos. Los análisis mostraron que no se definió ningún tipo de estructura poblacional, lo que permitió inferir un alto flujo genético entre las zonas geográficas, esto corroborado por los altos valores de diversidad genética obtenidos. Los 311 genotipos evaluados fueron definidos como una población natural heterogénea heterocigota, apta para favorecer el aumento de la base genética de las poblaciones cultivadas de palma de aceite.

**Palabras clave:** *Elaeis guineensis* Jacq., marcadores microsatélite, alelos, flujo genético.

### Abstract

Oil palm *Elaeis guineensis* Jacq. is of great importance because of the oil extracted from its fruits, whose refined derivatives are commercially valuable as biodiesel, among others uses. This study sought to determine the structure and genetic diversity of 311 oil palm samples from the Republic of Cameroon with 10 microsatellite markers. The results showed values for the average number of alleles per locus of  $N_a=8.433$  and effective number of alleles per locus of  $N_e=4.756$ ; from the differences between these values, it can be inferred that the 106 alleles found for these populations could be considered rare alleles. Additionally the value of genetic diversity was high (mean value of  $H_e=0.781$ ) compared to reports of several authors. The obtained molecular variance showed that the highest percentage (80 %) was found within the individuals. The analysis did not show any defined population structure, which allowed us to infer a high gene flow among the geographic zones, corroborated it by the high genetic diversity values obtained. The 311 genotypes were defined as a heterogeneous heterozygous natural population suitable to increase the genetic base of oil palm cultivated populations.

**Key words:** *Elaeis guineensis* Jacq, microsatellite markers, alleles, gene flow.

**Recibido:** diciembre 20 de 2013

**Aprobado:** octubre 20 de 2014

\* BSc. Universidad Pedagógica y Tecnológica de Colombia. estiby88@hotmail.com

\*\* BSc., MSc. Investigador Asociado. Programa de Biología y Mejoramiento de la Palma. Corporación Centro de Investigación en Palma de Aceite - Cenipalma. dianaris2002@gmail.com

\*\*\* Lic., MSc. Docente Asistente. Escuela de Ciencias Biológicas. Facultad de Ciencias. Laboratorio de Biotecnología Vegetal BIOPLASMA. Universidad Pedagógica y Tecnológica de Colombia, sede Tunja. zaida.ojeda@uptc.edu.co

\*\*\*\* BSc., MSc., PhD. Coordinador del Programa de Biología y Mejoramiento de la Palma. Corporación Centro de Investigación en Palma de Aceite - Cenipalma. Profesor Asociado. Departamento de Biología. Universidad Nacional de Colombia, sede Bogotá. hromero@cenipalma.org

## Introducción

La palma de aceite africana (*Elaeis guineensis* Jacq.) es una especie perenne originaria de África Central y Occidental; su cultivo es propio de las tierras bajas húmedas del trópico. La palma de aceite africana posee gran importancia comercial gracias al aceite que se extrae del mesocarpio, ya que de este tejido se obtiene el aceite para generar los productos para la industria alimenticia. Por otra parte se extraen otros ácidos grasos que están presentes en el endospermo de sus frutos, que también son importantes para la industria, por ejemplo para obtener la harina de coquito. Adicionalmente el aceite de palma, está siendo utilizado en la obtención de biodiesel a través de una reacción de trans-esterificación (FE-DEPALMA, 2009). Gracias al crecimiento de la industria, los programas de mejoramiento de la palma han cobrado importancia como una alternativa para contribuir al aumento de la rentabilidad de las plantaciones, mejorar la características como la calidad del aceite o la introducción de resistencia tanto a plagas como a enfermedades (Corley y Tinker, 2003).

La historia muestra que el material comercial sembrado mundialmente proviene de cuatro plantas cultivadas en el Jardín Botánico de Bogor de Java (Indonesia) en 1884, hecho que ha provocado uniformidad en el material productivo, generándose un efecto de estrechez en la base genética (Bakoumé *et al.*, 2007; Cochard *et al.*, 2009). Esto ha provocado la homogenización del cultivo, evidenciando la necesidad de ampliar la base genética de la palma de aceite, garantizando la obtención y conservación de una amplia gama de recursos genéticos, lo cual justifica la exploración y estudio de diversidad genética de materiales de palma provenientes del continente africano, en especial de Camerún, uno de los países centro de origen de esta especie. El Centro de investigación en palma de aceite (Cenipalma) hacia mediados del año 2007 con la cooperación científica del *Institute for Agricultural Research and Development* (IARD) de Camerún y el Ministerio de Agricultura y Desarrollo Rural (MADR), llevó a cabo la prospección y colecta de accesiones silvestres de palma de aceite en 6 zonas geográficas de la república de Camerún, estos materiales fueron sembrados en el Campo Experimental Palmar de la Vizcaína (Arias *et al.*, 2011). Una de las herramientas utilizadas en el proceso de mejoramiento vegetal es el uso de marcadores moleculares, los cuales permiten dilucidar la diversidad genética de materiales silvestres, con el fin de acelerar la selección de individuos que van a ser incorporados en etapas de pre-mejoramiento.

La diversidad genética de una especie se puede analizar mediante el uso de marcadores moleculares, los cuales se definen como secuencias dentro de un genoma que permiten detectar las diferencias genotípicas (Mayes *et al.*, 2000). Diversas técnicas de biología molecular se encuentran disponibles para la detección de variabilidad genética, entre los cuales se encuentran los basados en electroforesis de proteínas (Aloenzimas), hi-

bridación de sondas "Southern Blot", RFLP (*Restriction Fragment Length Polymorphism*), minisatélites o VNTR (*Variable Number of Tandem Repeats*) y los basados en PCR como PCR-RFLP, RAPD (*Random Amplified Polymorphic DNA*), AFLP (*Amplified Fragment Length Polymorphism*) y los microsátélites que son Secuencias Simples Repetidas, conocidos como SSR. Estos últimos son los marcadores más utilizados para estudios de estructura y diversidad genética ya que presentan secuencias altamente conservadas (Rocha y Gómez, 2003), son codominantes, presentan multiallelismo, y su distribución es al azar dentro del genoma. Algunos de los más importantes estudios dentro del género *Elaeis* son los realizados por Billotte *et al.* (2001), Bokumé *et al.* (2007) y Cochard *et al.* (2009) entre otros, quienes proyectan sus resultados en la aplicabilidad al interior de los programas de fitomejoramiento de esta especie. En este contexto, los estudios de diversidad genética de poblaciones naturales, mediante la utilización de técnicas moleculares, representan una herramienta útil en los programas de mejoramiento genético, en la medida que se involucran materiales no explotados comercialmente y con posible potencial agronómico. Así, la finalidad de este proyecto fue determinar tanto la estructura como la diversidad genética de materiales de palma de aceite provenientes de Camerún, resaltando su importancia como centro de origen.

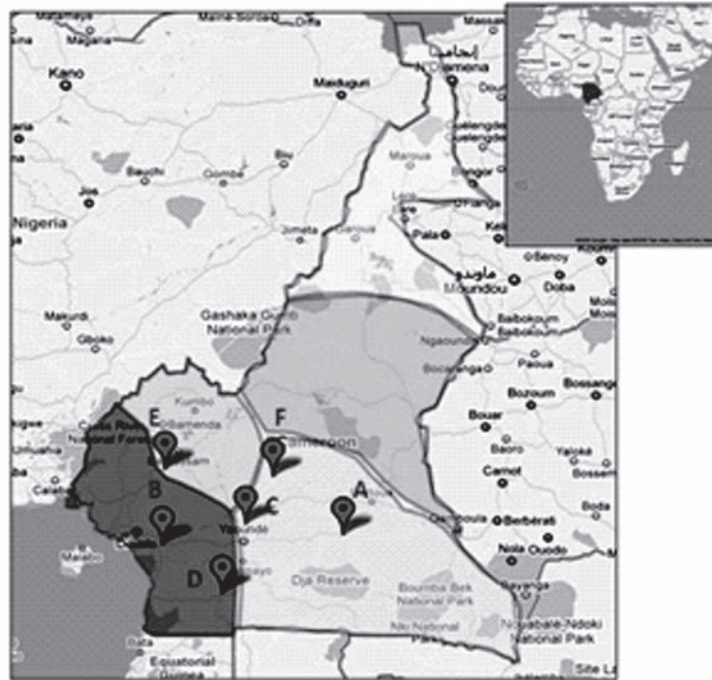
## Materiales y Métodos

### Material vegetal

El material vegetal utilizado consistió de 311 genotipos de palma de aceite *Elaeis guineensis* (en estado de vivero), colectado en seis zonas geográficas de la República de Camerún (figura 1). Actualmente estas accesiones se encuentran sembradas en el Centro Experimental Palmar de la Vizcaína (Barrancabermeja-Santander) y hacen parte de la colección *ex situ* de Cenipalma. La estrategia de prospección de las palmas colectadas en Camerún, fue llevada a cabo inicialmente mediante un recorrido de 9700 km por tierra, para la identificación y selección de regiones geográficas con base a las condiciones climáticas, vegetación, suelos y altitud. En las zonas escogidas se tomaron muestras de racimos maduros, los cuales fueron organizados de acuerdo con las poblaciones espontáneas encontradas. Cada población espontánea fue catalogada como población independiente, si se delimitaba y se verificaba que por lo menos a más de 100 metros no hubiera más palmas. Posteriormente dentro de cada área se colectó un racimo por palma, al cual se le realizó una caracterización morfo-agronómica (Arias *et al.*, 2011).

### Extracción y cuantificación del ADN

El tejido foliar colectado en cada uno de los 311 genotipos, fue triturado utilizando nitrógeno líquido y posteriormente la extracción del ADN se realizó uti-



**Figura 1.** Mapa de localización de las seis zonas geográficas de la república de Camerún, en donde se realizó la colecta de las accesiones de palma de aceite. **A.** Abong-Mbang **B.** Edea **C.** Obala **D.** Ebolowa **E.** Bafang **F.** Mankim.

lizando el kit DNeasy Plant Mini kit (250) ref: 69106 de QIAGEN, según las indicaciones del fabricante. El proceso para verificar la calidad y estimar la cantidad del ADN obtenido se realizó mediante dos metodologías; la primera fue la visualización de la muestra en gel de agarosa y su cuantificación por comparación de la fluorescencia de las bandas, observadas en el Fotorodocumentador BioRad- Universal Hood II-. El segundo método utilizado, fue un procedimiento de espectrofotometría UV-visible a través de un espectrofotómetro JENWAY 6405 Uv/vis, utilizando una longitud de onda de 260 nm. Para verificar la calidad del ADN se observó la relación entre la longitud de onda de 260nm y 280nm. Las lecturas obtenidas por espectrofotometría se usaron para calcular la concentración de ADN obtenida para cada muestra y a partir de estas se realizaron las diluciones para uniformar la concentración de ADN de cada muestra a 5 ng/μl.

### **Amplificación SSR**

La reacción de amplificación de SSR contenía una concentración final de 25 ng/μl de ADN, empleando 10 marcadores microsatélites reportados por Billotte *et al.* (2001), Billotte *et al.* (2005) y Singh *et al.* (2008). La amplificación de microsatélites mediante PCR se llevó a cabo en un volumen final de 15μl que contenía una solución amortiguadora de PCR 1X (20mM de Tris HCl, pH8.4 y 50 mM de KCl); 1.5 mM de MgCl<sub>2</sub>, 0.2μM de cada dNTP y 1U de Taq DNA polimerasa

(Invitrogen Life Technologies). Las condiciones de amplificación fueron: 1 minuto a 95°C (desnaturalización inicial), seguido por 35 ciclos de amplificación, cada uno compuesto por intervalos de temperatura a 94 °C durante 2 min, 52 °C por 30 seg, y 72 °C por 90 seg. Un paso de extensión final durante 8 min a 72 °C. La verificación de los productos amplificados, se realizó mediante electroforesis en gel denaturalizante de poli-acrilamida al 6 % y 5M de urea. Posteriormente, el gel fue teñido con nitrato de plata.

### **Análisis de datos**

Para cada locus se determinó el número de alelos amplificados, utilizando como referencia los patrones de amplificación reportados por los autores. Para estimar la diversidad genética se tuvieron en cuenta los siguientes parámetros: frecuencia alélica (*f*), Número medio de alelos por locus (*N<sub>a</sub>*), número efectivo de alelos por locus (*N<sub>e</sub>*), heterocigosidad observada (*H<sub>o</sub>*), heterocigosidad esperada (*H<sub>e</sub>*) (parámetro utilizado para cuantificar la diversidad genética) bajo el índice de Nei (1978), porcentaje de loci polimórficos (%P) e Índice de fijación (*F*). Para obtener esta información se hizo uso de los algoritmos incluidos en el programa GenALEX 6.1 (Peakall y Smouse, 2006). La estructura y diferenciación genética entre zonas se calculó mediante los estadísticos *F* (*F<sub>ST</sub>*, *F<sub>IS</sub>* y *F<sub>IT</sub>*) de diferenciación de Wright (1965) y el Flujo genético (*N<sub>m</sub>*) (Nei, 1987), fue estimado indirectamente a partir de *F<sub>ST</sub>* (Wright, 1965) utilizando el programa GenALEX 6.1 (Peakall y

Smouse, 2006). Para evaluar la estructura poblacional se realizó un Análisis de la Varianza Molecular (AMOVA) con fuentes de variación entre y dentro de zonas geográficas. Con esta AMOVA se calcularon los estadísticos *F*. El test de significancia estadística se realizó a través de 999 permutaciones con el programa GenALEX 6.1 (Peakall y Smouse, 2006). Para conocer el tipo de relaciones genéticas entre zonas geográficas se realizó un Análisis de Coordenadas Principales, mediante el programa GenALEX 6.1 (Peakall y Smouse, 2006), utilizando la matriz de distancia genética de Nei (1972). Adicionalmente se realizó un Análisis de Coordenadas Principales, tomando como unidades los individuos y/o accesiones para el cual se utilizó la matriz de similitud genética de Nei y Li (1979); este análisis fue realizado mediante el programa NTSYS 2.1 (Rohlf, 2000).

## Resultados y discusión

### Polimorfismo

Los 10 SSR analizados presentaron un excelente resultado en la amplificación (figura 2), hecho que facilitó la obtención de la información. Para determinar la longitud de cada alelo y evitar la lectura de alelos falsos, se tuvieron en cuenta los rangos de longitud reportados por Billotte *et al.* (2001), Billotte *et al.* (2005), Bakoume *et al.* (2007) y Singh *et al.* (2008), pudiéndose verificar que los alelos observados en este trabajo se encontraron dentro de los rangos de amplificación reportados (tabla 1). Sin embargo los alelos observados para los locus sEg00066 y mEgCIR3282, se encuentran por encima del rango reportado por los autores. Lo anterior evidencia que aunque la longitud de los alelos reportados por otros autores, sirven como una guía para realizar una correcta lectura; no es una regla general. Debido a que el número de repeticiones de cada lo-

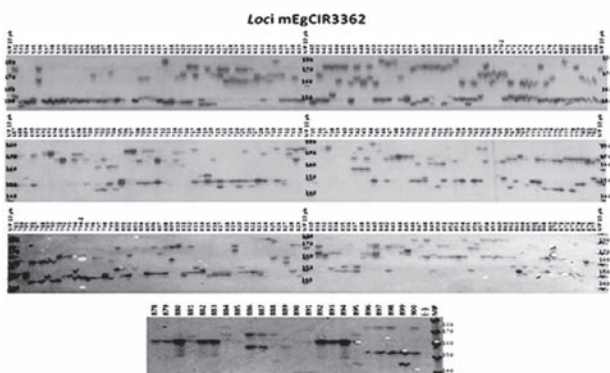
cus, así como el número de los alelos detectados depende de la naturaleza del germoplasma y por ende de su variabilidad.

**Tabla 1.** Rangos de amplificación reportados y observados para los 10 SSR evaluados.

Loci SSR	Motivo de Repetición	Rango de longitud reportado (pb)	Rango de longitud observado (pb)
mEgCIR0008	(GA) <sub>18</sub>	200-225	200-225
mEgCIR0067	(GA) <sub>17</sub>	135-187	145-175
sEg00066	(AT) <sub>8</sub>	192-215	238-260
mEgCIR0173	(GA) <sub>18</sub>	124-158	110-140
mEgCIR0802	(GA) <sub>12</sub>	236-272	218-242
mEgCIR1730	(CT) <sub>17</sub> (GT) <sub>5</sub>	248-274	248-274
mEgCIR3282	(GA) <sub>20</sub>	162-206	218-240
mEgCIR3362	(GA) <sub>19</sub>	146-178	146-178
mEgCIR3363	(GA) <sub>17</sub>	186-228	166-204
mEgCIR3546	(GA) <sub>5</sub> GT(GA) <sub>20</sub>	286-336	262-310

El número de alelos obtenido por locus varió en un rango de 8 a 16, con un total de 106 alelos (tabla 2). En los loci mEgCIR0008 y mEgCIR0067 evaluados en este estudio se obtuvo un número mayor de alelos al reportado por Billotte *et al.* (2001). Estos autores indican que para el locus mEgCIR0008 amplificaron 6 alelos y para el locus mEgCIR0067 8 alelos, en 18 muestras de palma de aceite *E. guineensis* de diferentes orígenes incluyendo Camerún. Para el marcador microsatélite sEg00066 se obtuvo un número menor de alelos al reportado por Singh *et al.* (2008), quienes indicaron que para este marcador amplificaron 10 alelos, en 76 muestras de *E. guineensis* de diferentes orígenes incluyendo Camerún. La cantidad de datos faltantes para las accesiones analizadas, alcanzó un porcentaje muy bajo (0.19%), en los 10 locus analizados para los 311 individuos, razón por la cual todos los marcadores fueron incluidos en el análisis de datos.

Al comparar el número de alelos obtenido en esta investigación con los reportados por Bakoumé *et al.* (2007), quienes evaluaron 494 muestras de *E. guineensis* de 10 países Africanos, los loci mEgCIR1730 y mEgCIR3546 mostraron un número mayor de alelos; mEgCIR0173 evidenció el mismo número de alelos al reportado, mientras que para los demás loci el número de alelos fue menor al reportado. Los resultados anteriores mostraron que la cantidad de alelos está directamente relacionada con el número de muestras y con los diferentes orígenes utilizados en un estudio. Es



**Figura 2.** Patrón de amplificación del Locus mEgCIR3362. En cada columna se indica el patrón de amplificación de cada una de las 311 muestras de palma de aceite de Camerún. Donde MP representa el patrón de bandas del marcador de peso molecular y (-) es la columna correspondiente al control negativo.

**Tabla 2.** Número total de alelos, frecuencia de alelos y porcentaje de datos faltantes para 10 SSR evaluados.

Locus SSR	No. de alelos reportados	No. de alelos amplificados	Alelos >5%	Alelos <5%	Datos faltantes (%)
mEgCIR0008	6	8	0	8	0.60
mEgCIR0067	8	11	0	11	0.00
sEg00066	10	8	0	8	0.30
mEgCIR0173	11	11	0	11	0.00
mEgCIR0802	12	9	0	9	0.60
mEgCIR1730	8	10	0	10	0.30
mEgCIR3282	12	8	0	8	0.00
mEgCIR3362	22	16	0	16	0.00
mEgCIR3363	17	12	0	12	0.00
mEgCIR3546	12	13	0	13	0.00
Total	118	106	0	106	0.19

importante reconocer que las 311 muestras de palma de aceite presentaron un alto número de alelos a pesar de incluirse un solo país de origen (Camerún), observación que concuerda con Bakoumé *et al.* (2007). Por otra parte las frecuencias reportadas en este estudio, mostraron que ninguno de los alelos encontrados para los loci analizados se presentó recurrentemente en porcentajes mayores al 5% (tabla 3). Al respecto, Kimura (1983) y Joyce y Tavaré (1995) consideran que si la frecuencia de un alelo, medida bajo su frecuencia relativa, resulta menor que el valor establecido para  $q$  (0.5), como en este caso, los 106 alelos encontrados serán descritos como alelos raros.

Por otra parte la prueba de chi-cuadrada aplicada para todos los loci, excepto para el locus mEgCIR3282, fue altamente significativa ( $p < 0.001$ ) (tabla 3). Por lo tanto se rechaza la hipótesis nula de que la distribución de las frecuencias es rectangular y se acepta la hipótesis alterna de que la distribución de las frecuencias no es rectangular o uniforme.

### **Diversidad genética dentro de cada zona geográfica**

Como resultados relevantes se obtuvieron los valores promedio para el número medio de alelos por locus ( $N_a=8.433$ ) y el número efectivo de alelos por locus ( $N_e=4.756$ ) (tabla 4). Laurentin (2009) indica que cuando estos valores son muy cercanos, las frecuencias alélicas son similares. Sin embargo, las amplias diferencias entre los valores de  $N_a$  y  $N_e$  permiten inferir que los 106 alelos encontrados para estas poblaciones son considerados como alelos raros, confirmando los resultados obtenidos para el porcentaje de las frecuencias alélicas. Bakoumé *et al.* (2007) reportaron valores para el número medio de alelos por locus ( $N_a$ ) con

rangos entre 4.4 a 5.5 y el número efectivo de alelos por locus ( $N_e$ ) con rangos entre 2.5 a 3.9, para 102 materiales de palma de aceite de Camerún, valores que son menores a los reportados en esta investigación. El alto porcentaje de alelos raros encontrados podría estar relacionado con el porcentaje de loci polimórfico (%P), ya que todos los loci estudiados para las 6 zonas mostraron un 100% de polimorfismo bajo el criterio del 0,99. Es importante resaltar que la selección de los marcadores SSR fue influenciada por los reportes de los autores quienes diseñaron y describieron preliminarmente estos 10 cebadores como los más polimórficos dentro de sus investigaciones.

Por otra parte, la media de la heterocigosidad observada ( $H_o$ ) para las 6 zonas estudiadas fue 0.644. La zona F presentó la  $H_o$  más alta (0.710) mientras que la zona C presentó la más baja (0.585). Bakoumé *et al.* (2007) reportaron un valor medio de  $H_o= 0.405$  para 102 materiales procedentes de Camerún. Singh *et al.* (2008) reportaron un valor de  $H_o= 0.338$  para 10 materiales procedentes de Camerún.

El valor de diversidad genética encontrado es similar entre las 6 zonas geográficas de Camerún, presentando valores de  $H_e$  que oscilan entre 0.746 y 0.798, con un valor medio de 0.781. Estos valores son mayores a los reportados por Billote *et al.* (2001) quienes presentaron un valor medio de  $H_e = 0.680$  para 18 muestras de *E. guineensis* de diferentes orígenes, evaluadas con 21 SSR. Montoya *et al.* (2005) reportaron un valor de  $H_e= 0.454$  para 48 muestras de palma de aceite provenientes de Angola evaluadas con 11 SSR. Bakoumé *et al.* (2007) reporta un valor de  $H_e = 0.660$ , para 102 materiales de Camerún y  $H_e = 0.644$  para 494 muestras provenientes de 10 países Africanos evaluados

**Tabla 3.** Frecuencia alélica, desviación estándar y prueba de chi-cuadrado para los loci evaluados.

Locus	Alelo	Frecuencia	Desviación Estándar	gl	Chi-cuadrado	Probabilidad	Significancia
mEgCIR0008	200	0.421	0.330	36	371.778	0.000	***
	207	0.023	-0.069				
	210	0.228	0.137				
	212	0.154	0.063				
	215	0.043	-0.048				
	218	0.026	-0.066				
	220	0.074	-0.018				
	225	0.024	-0.068				
mEgCIR0067	145	0.071	-0.021	55	264.398	0.000	***
	148	0.098	0.006				
	150	0.051	-0.040				
	155	0.180	0.088				
	157	0.241	0.150				
	158	0.235	0.143				
	160	0.103	0.011				
	162	0.013	-0.079				
	168	0.002	-0.090				
	172	0.005	-0.087				
175	0.002	-0.090					
sEg00066	238	0.180	0.088	36	447.966	0.000	***
	240	0.058	-0.034				
	246	0.106	0.014				
	249	0.178	0.087				
	250	0.206	0.114				
	252	0.154	0.063				
	258	0.103	0.011				
	260	0.011	-0.080				
mEgCIR0173	110	0.024	-0.067	55	692.959	0.000	***
	112	0.163	0.072				
	114	0.261	0.169				
	116	0.460	0.368				
	118	0.019	-0.072				
	120	0.036	-0.056				
	124	0.005	-0.087				
	130	0.005	-0.087				
	132	0.018	-0.074				
	134	0.003	-0.088				
140	0.006	-0.085					
mEgCIR0802	218	0.021	-0.071	45	618.643	0.000	***
	220	0.172	0.080				
	222	0.037	-0.055				
	226	0.116	0.024				
	228	0.232	0.140				
	230	0.201	0.109				
	238	0.084	-0.008				
	240	0.108	0.016				
242	0.024	-0.068					

gl: grados de libertad; ns: no significativo; \*\*\*  $p < 0.001$

Continuación **Tabla 3.**

Locus	Alelos	Frecuencia	Desviación Estándar	GL	Chi-cuadrado	Probabilidad	Significancia
mEgCIR1730	248	0.109	0.015	55	458.346	0.000	***
	250	0.076	-0.019				
	252	0.031	-0.064				
	256	0.137	0.043				
	258	0.233	0.139				
	260	0.270	0.176				
	262	0.087	-0.007				
	264	0.043	-0.051				
	270	0.005	-0.089				
274	0.006	-0.088					
mEgCIR3282	218	0.016	-0.078	28	27.659	0.483	ns
	220	0.421	0.327				
	224	0.063	-0.031				
	228	0.230	0.136				
	230	0.180	0.086				
	234	0.039	-0.056				
	236	0.042	-0.052				
240	0.010	-0.084					
mEgCIR3362	146	0.090	-0.004	120	297.717	0.000	***
	150	0.291	0.197				
	152	0.034	-0.060				
	154	0.035	-0.059				
	156	0.013	-0.081				
	158	0.021	-0.073				
	160	0.084	-0.011				
	162	0.047	-0.047				
	164	0.072	-0.022				
	166	0.064	-0.030				
	168	0.035	-0.059				
	170	0.027	-0.067				
	172	0.064	-0.030				
	174	0.085	-0.009				
176	0.013	-0.081					
178	0.024	-0.070					
mEgCIR1730	180	0.090	-0.002	55	458.346	0.000	***
	182	0.177	0.085				
	184	0.262	0.170				
	186	0.206	0.114				
	188	0.071	-0.021				
	190	0.061	-0.031				
	192	0.079	-0.013				
	194	0.039	-0.053				
	196	0.003	-0.088				
	198	0.002	-0.090				
	200	0.005	-0.087				
204	0.002	-0.090					
mEgCIR3546	228	0.002	-0.092	78	220.665	0.000	***
	262	0.013	-0.081				
	272	0.023	-0.072				
	274	0.105	0.010				
	280	0.016	-0.078				
	282	0.016	-0.078				
	284	0.064	-0.030				
	288	0.262	0.168				
	290	0.220	0.126				
	292	0.121	0.026				
	294	0.095	0.001				
	300	0.024	-0.070				
310	0.040	-0.054					

gl: grados de libertad; ns: no significativo; \*\*\* p < 0.001

con 16 SSR. Por otra parte Singh *et al.* (2008) reportaron para 10 muestras de palma de aceite provenientes de Camerún un valor de  $H_e=0.328$  y una media de  $H_e= 0.356$ , para 76 muestras provenientes de países africanos evaluados con 10 EST (*Expressed Sequence Tags*)-SSR. Sin embargo no es posible realizar una comparación equitativa con nuestros resultados, puesto que por su naturaleza los marcadores EST-SSR, tienden a ser menos polimórficos, debido a que estos son obtenidos a partir de genes transcritos con una función putativa. Actualmente estos marcadores están siendo utilizados para la selección asistida por marcadores, especialmente cuando residen en los genes responsables de un rasgo fenotípico (Varshney *et al.*, 2005). Caso contrario ocurre con los marcadores SSR obtenidos a partir del ADN genómico, que debido a esta característica son altamente polimórficos.

La diversidad genética en palma de aceite ha sido evaluada mediante otro tipo de marcadores moleculares como isoenzimas (Hayati *et al.*, 2004 y Purba *et al.*, 2000), RFLP (Maizura *et al.*, 2006) y AFLP (Barcelos *et al.*, 2002), que muestran valores de heterocigosidad esperada ( $H_e$ ) para estas técnicas, generalmente menores a los valores de  $H_e$  obtenidos mediante marcadores tipo microsatélite. Por esto los marcadores SSR son reconocidos como una excelente y poderosa herramienta para estudios de diversidad genética (Weising *et al.*, 2005 y Datta *et al.*, 2009). El índice de fijación ( $F$ ) obtenido corresponde a 0.165 para todas las poblaciones. Siendo que estos valores se miden en un rango que varía entre -1 que indica un alto grado de heterocigosidad y 1 que indica un alto grado de homocigosidad, se infiere el entrecruzamiento entre individuos, provocado un exceso de homocigotos. Estas desviaciones positivas pueden ser provocadas por tres factores esenciales: (1) altos niveles de endogamia, (2) cualidades de la estructura o subestructura poblacional, (3) el número limitado de individuos que conforman la población versus la cantidad de alelos obtenidos (Romero *et al.*, 2003). Para este estudio el exceso de homocigotos podría estar asociado a las cualidades de la estructura o subestructura poblacional. Lo anterior ha sido sustentado por los valores de  $F$  ( $F_{IT}$ ,  $F_{ST}$  y  $F_{IS}$ ), los cuales revelaron un déficit de heterocigotos entre las zonas geográficas evaluadas (tabla 5).

### Estructura Genética entre zonas geográficas

Los rangos de interpretación propuestos por Wright (1965), indicaron un bajo grado de diferenciación genética entre las 6 zonas geográficas, pues los valores obtenidos para el índice  $F_{ST}$  se encuentran distribuidos en un rango que oscila entre 0.013 a 0.049, los cuales son estadísticamente significativos. Los resultados del análisis de varianza molecular (AMOVA) muestran que el mayor porcentaje de la varianza (80%), se encuentra dentro de las zonas geográficas, mientras que el porcentaje restante se encuentra distribuido entre

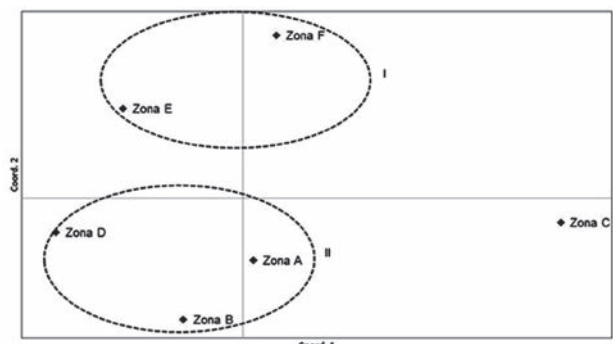
zonas geográficas (3%) y entre individuos (17%). Estos resultados son similares a los encontrados por Cochard *et al.* (2009) obteniendo como resultado que el 90.16 % de la varianza se encontraba distribuida dentro de orígenes. El término "origen" fue designado por estos autores, para referirse al sitio o zona geográfica en donde realizaron la colecta de las 318 accesiones analizadas con 14 marcadores SSR.

Los rangos de interpretación propuestos por Wright (1965) indicaron que el grado de diferenciación genética entre las 6 zonas es bajo,  $F_{ST}= 0.030$  ( $p > 0.001$ ) (tabla 5), lo cual podría estar asociado con el alto número de migrantes entre poblaciones  $N_m = 8.093$ , evidenciando que existe recombinación genética entre las poblaciones evaluadas, sobrepasando los efectos de deriva e impidiendo diferenciación local (Slatkin, 1994). Por su parte los valores de  $F_{IT}=0.204$  ( $p > 0.001$ ) y  $F_{IS}=0.180$  ( $p > 0.001$ ) fueron positivos y significativos, indicando un déficit de heterocigosidad en la población como en cada sub población respectivamente, o la presencia de alelos nulos en la población que no pudieron ser detectados. De Vicente *et al.* (2004) explica que estos valores ( $F_{ST}$ ,  $F_{IS}$ ,  $F_{IT}$ ) tienden a modificarse en función de la cantidad de individuos estudiados, influyendo en la lectura real del flujo de genes que tiene lugar en la población.

### Relaciones genéticas entre zonas geográficas

En el análisis de coordenadas principales realizado con la matriz de distancia genética de Nei (1972), mostró que las dos primeras coordenadas explican el 64.9% del total de la variación, la primera componente explica 34.47% y la segunda 30.45%. Dos grupos fueron visibles en las dos componentes (figura 3).

El grupo I está conformado por las zonas E y F. El grupo II esta con formado por tres zonas A, B y D. La zona C es la más distante a nivel genético. La matriz de la distancia genética de Nei (1972) fue interpretada con valores entre 0 (máxima igualdad genética) y 1 (máxima distancia genética). Según los valores obtenidos



**Figura 3.** Dispersión de las seis zonas geográficas revelado a través del análisis de coordenadas principales.



**Tabla 4.** Media y desviación estándar para los parámetros de diversidad genética por zona geográfica.

Zona geográfica	N	$N_a$	$N_e$	%P	$H_o$	$H_e$	F
Zona A	81	9.400 <sup>a</sup>	4.682	100	0.605	0.779	0.221
		0.819 <sup>b</sup>	0.348	-	0.054	0.019	0.066
Zona B	44	8.200	5.049	100	0.636	0.794	0.185
		0.593	0.422	-	0.047	0.023	0.061
Zona C	20	6.500	3.970	100	0.585	0.746	0.192
		0.428	0.326	-	0.069	0.029	0.103
Zona D	70	9.400	4.963	100	0.679	0.789	0.132
		0.748	0.433	-	0.025	0.019	0.031
Zona E	60	8.700	5.169	100	0.647	0.798	0.182
		0.539	0.482	-	0.034	0.019	0.037
Zona F	36	8.400	4.705	100	0.710	0.779	0.076
		0.636	0.443	-	0.035	0.024	0.036
Total	311	8.433	4.756	100	0.644	0.781	0.165
		0.281	0.169	-	0.019	0.009	0.025

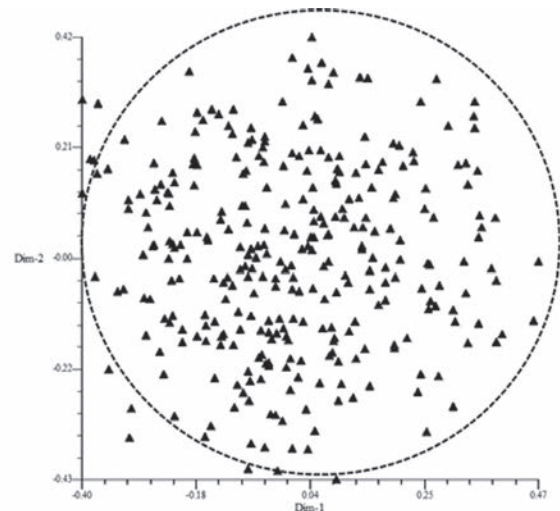
muestran que las zonas geográficas B y D son las más cercanas (0.0934), mientras que las zonas C y D fueron las más distantes (0.2441). De alguna forma lo anterior se ve reflejado en el mapa de localización de las zonas geográficas (figura 1), en donde las poblaciones B y D presentan la menor distancia geográfica entre ellas. La cercanía geográfica entre las zonas, facilita de cierta manera la dispersión de semilla, lo cual podría estar asociado con el alto número de migrantes ( $N_m = 8.149$ ) encontrado en este estudio, que actúa como un efecto en la homogenización de la variación genética entre los grupos y sobrepasa los efectos de la deriva e impide la diferenciación local (Hedrick, 2005).

**Tabla 5.** Valores totales del estadístico F y Flujo genético ( $N_m$ ) entre zonas geográficas.

Estadísticos F	Valor	P(rand >= data)
$F_{ST}$	0.030	0.001
$F_{IS}$	0.180	0.001
$F_{IT}$	0.204	0.001
$N_m$	8.149	

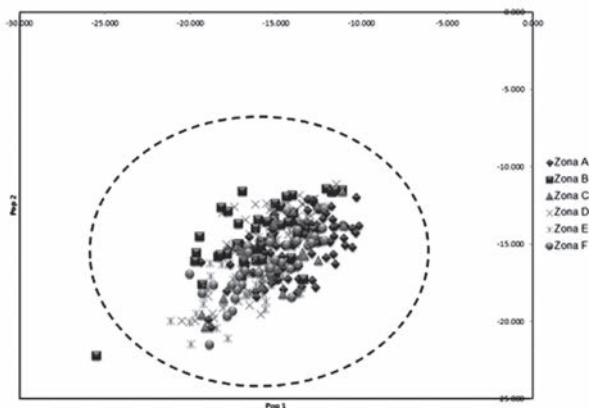
Por su parte los análisis de coordenadas principales tomando como unidades las 311 accesiones de palma de aceite, utilizando el coeficiente de similitud genética de Nei & Li (1979) (figura 4) y el análisis de

asignación de poblaciones (figura 5), mostraron a nivel espacial la conformación de un solo grupo; en donde no presentó ninguna asociación con los orígenes geográficos en donde fueron colectadas las palmas.



**Figura 4.** Análisis de coordenadas principales para 311 individuos de palma de aceite de Camerún.

Según estos resultados, se evidencia la ausencia de estructura poblacional definida, permitiendo inferir que existe un alto flujo genético entre los individuos, po-



**Figura 5.** Análisis de asignación de poblaciones para 311 individuos de palma de aceite de Camerún.

siblemente sustentado por los altos valores de diversidad genética obtenidos. Epperson (2003) corrobora esta información revelando que la estructura genética puede ser difícil de identificar cuando aumenta la diversidad genética, o ser concreta cuando ésta disminuye. Además Slatkin (1994) y Neigel (1997), muestran que el flujo genético puede influenciar en la modificación de la distribución espacial de los genes, e intervenir en procesos como selección, recombinación y la mutación. Particularmente en la palma de aceite, el flujo genético podría estar directamente relacionado con el nivel de migración del polen (debida principalmente a dispersores aéreos) y a una activa dispersión de semillas, evitando la estructuración de las poblaciones y favoreciendo la diversidad genética de la especie (Ennos, 1994 y Corley y Tinker, 2003). Los resultados de esta investigación coinciden con Billote *et al.* (2001) y Montoya *et al.* (2005), quienes reportaron que las muestras analizadas no presentaron ningún tipo de estructura genética. Estos resultados han sido confirmados por Barcelos *et al.* (2002) quienes estudiando 38 accesiones de *E. guineensis* analizados a través de RFLP y 22 accesiones de palma de aceite analizados a través de AFLP, reportan un bajo nivel de estructuración genética dentro de todos los orígenes evaluados, lo que fue atribuido a la dispersión de material sin barreras geográficas en el continente africano. Maizura *et al.* (2006) a través de RFLP analizando 359 accesiones de palma de aceite provenientes de 11 países africanos encontró que las poblaciones incluidas en el estudio pudieron haber experimentado el flujo de genes entre las poblaciones dando lugar al alto nivel de heterocigosidad ( $H_e$ ), atribuido a la dispersión de semillas a través de los ríos. En relación con estructura genética, los resultados de Hayati *et al.* (2004) contradicen lo reportado, indicando que el flujo genético obtenido (0.56) es bajo, denotando que de los 26 orígenes en los 10 países muestreados, la diferenciación genética entre las poblaciones de palma de aceite disminuyó a lo largo de una línea hacia el este de África, lo que

indica que las poblaciones de África central y occidental son más similares entre sí que con las poblaciones de África oriental. Sin embargo, es importante anotar que se incluyeron materiales de palma de aceite provenientes de la isla de Madagascar, lo cual pudo haber afectado la estimación del flujo genético y por ende la diferenciación genética.

## Conclusiones

Los diferentes análisis realizados en esta investigación evidenciaron la ausencia de estructura poblacional en las 311 muestras de palma de aceite *E. guineensis* provenientes de Camerún, infiriendo un alto flujo genético entre los individuos, apoyado por los altos valores de diversidad genética encontrados. Estos resultados permiten referirse a los genotipos evaluados como una población natural heterogénea heterocigota, apta para favorecer el aumento de la base genética de las poblaciones cultivadas de palma de aceite.

## Agradecimientos

Los autores expresan sus agradecimientos al Departamento Administrativo de Ciencia Tecnología e Innovación Colciencias, por su apoyo financiero para la realización de este trabajo bajo el Contrato No. 626-2009. A Myriam Cristina Duque, Investigadora del Centro Internacional de Agricultura Tropical (CIAT) y a Eloina Mesa, por la asesoría en el análisis de datos. Los autores también agradecemos a los doctores: Leonardo Rey y Claude Bakoumé, quienes realizaron la colecta de las accesiones de palma de aceite en Camerún. La investigación de Cenipalma es financiada por el Fondo de Fomento Palmero (FFP) administrado por Fedepalma.

## Referencias bibliográficas

- Arias, D.; Daza, E.; Montoya, C.; Romero, H. 2011. Colección genética de materiales de palma de aceite provenientes de Camerún. *Rev. Palmas.* 32(3): 27-36.
- 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. Implication in conservation. *Rev. Palmas.* 28(1): 149-158.
- Barcelos, E.; Amblard, P.; Berthaud, J.; Seguin, M. 2002. Genetic diversity and relationship in American and African oil palm as revealed by RFLP and AFLP molecular markers. *Pesq. Agropec. Bras.* 37(8):1105-1114.
- Billotte, N.; Rusterucci, A.; Barcelos, E.; Noyer, J.; Amblard, P.; and Baurens, F. 2001. Development, characterization, and across-taxa utility of oil palm (*Elaeis guineensis* Jacq.) microsatellite markers. *Genome.* 44 (3): 413-425.
- Billotte, N.; Marseillac, N.; Rusterucci, A. M.; Adon, B.; Brottier, P.; Baurens, F. C.; ... & Charrier, A. 2005. Microsatellite-based high density linkage map in oil palm (*Elaeis guineensis* Jacq.). *Theoretical and Applied Genetics.* 110(4): 754-765.
- Cochard, B.; Adon, B.; Rekima, S.; Billotte, N.; Desmier de Chenon, R.; Koutou, A.; Nouy, B.; Omoré, A.; Purba, A.; Glazsmann, J.; Noyer, J. 2009. Geographic and genetic structure of African oil palm diversity suggests new approaches to breeding. *Tree genetics & genomes.* 5:493-504

- Corley, R.H.V.; and Tinker, P.B. 2003. The oil palm. 4th Ed. World Agricultural Series. Blackwell Publishers Ltd, Oxford, UK.
- Datta, S.; Kaashyap, M.; Kumar, S. 2009. Amplification of chickpea-specific SSR primers in *Cajanus* species and their validity in diversity analysis. *Plant Breeding*. 129(3): 334-340. doi:10.1111/j.1439-0523.2009.01678.x.
- De Vicente, M.; López, C.; Fulton, T. 2004. Genetic Diversity Analysis with Molecular Marker Data: Learning Module. International Plant Genetic Resources Institute (IPGRI).
- Ennos, R. 1994. Estimating the relative rates of pollen and seed migration among plant populations. *Heredity*. 72: 250-259.
- Epperson, B. 2003. Geographical Genetics. Princeton University Press: Englewood Cliffs, NJ.
- Fedepalma. 2009. La Agroindustria de la palma de aceite en Colombia y el mundo. ISBN: 9789589834169.
- Hayati, A.; Wickneswari, R.; Maizura, I.; Rajanaidu, N. 2004. Genetic diversity of oil palm (*Elaeis guineensis* Jacq) germplasm collections from Africa: implications from improvement and conservation of genetic resources. *Theor. Appl. Genet.* 108: 1274-1284.
- Hedrick, P.W. 2005. Genetics of Populations (Third Edition). Jones and Bartlett, Boston.
- Joyce, P.; Tavaré, S. 1995. The distribution of rare alleles. *J. Math. Biol.* 33: 602-618
- Kimura, M. 1983. Rare variant alleles in the light of the neutral theory. *Mol. Biol. Evol.* 1: 84-93.
- Laurentin, H. 2009. Data analysis for molecular characterization of plant genetic resources. *Genetic Resources and Crop Evolution*. 56(2): 277-292.
- Maizura, I.; Rajanaidu, N.; Zakri, A.; Cheah, S. 2006. Assessment of Genetic Diversity in Oil Palm (*Elaeis guineensis* Jacq) using restriction fragment length polymorphism (RFLP). *Genet. Resour. Crop. Evol.* 53(1):187-195(9).
- Mayes, S.; Jack, P.L.; Corley, R.H.V. 2000. The use molecular markers to investigate the genetic structure of an oil palm breeding programme. *Heredity*. 85:288-293.
- Montoya, C.; Arias, D.; Rey, L.; Rocha, P. 2005. Caracterización molecular de materiales de *E. guineensis* Jacq. Procedentes de Angola. *Fitotecnia Colombiana*. 5(2):1-10.
- Nei, M. 1972. Genetic distance between populations. *American Naturalist*. 106: 283-292.
- Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individual. *Genetics*. 89:583-590.
- Nei, M and Li, W.H. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Science*. 76: 5269-5273
- Nei, M. 1987. Molecular evolutionary genetics. Columbia University Press.
- Neigel, J. 1997. A comparison of alternative strategies for estimating gene flow from genetic markers. *Annu. Rev. Ecol. Syst.* 28:105-128.
- Peakall, R.; Smouse, P. 2006. GenALEX version 6.0: Genetic Analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology*. 6: 288-295.
- Purba, A.; Glazmann, J.; Noye, J. 2000. Geographic and genetic structure of African oil palm diversity suggests new approaches to breeding. Doi: 10.1007/s11295-009-0203-3.
- Rocha, P.; Gomez, P. 2003. Marcadores moleculares, una herramienta útil para la selección genética de palma de aceite. *Palmas*. 24(2):9-25.
- Rohlf, F. 2000. NTSYS pc: Numerical taxonomy and multivariate system, ver. 2.1, computer program. Setakuet, New York: Exeter Publishing.
- Singh, R.; Zaki, N.; Ting, N.; Rosli, R.; Tan, S.; Low, E.; Maizura, I.; Cheah, S. 2008. Exploiting an oil palm EST database for the development of gene-derived SSR markers and their exploitation for assessment of genetic diversity Biologic. *Biol. Cell. Mol.* 63(2): 227-235.
- Slatkin, M. 1994. Gene flow and population structure. *Ecological Genetics*, Editado por L. Real. pp. 3-17. Princeton University Press, Princeton, NJ.
- Romero, C.; Pedryc, A.; Muñoz, V.; Llácer, G.; Badenes, M. 2003. Genetic diversity of different apricot geographical groups determined by SSR markers. *Genome*. 46: 244-252.
- Varshney, K.R.; Graner, A.; Sorrells, M.E. 2005. Genomic microsatellite markers in plants: features and applications. *Trends in Biotechnology*. 23(1): 48-55.
- Weising, K.; Nybom, H.; Wolff, K.; Kahl, G. 2005. DNA Fingerprinting in Plants: Principles, Methods, and Applications, Second Edition. Taylor & Francis Group. United States.
- Wright, S. 1965. The interpretation of population structure by F-statistics with special regard to system of mating. *Evolution*. 19: 395-420.