

## DIVERSIDAD GENÉTICA EN SOYA DEL TRÓPICO HÚMEDO DE MÉXICO DETERMINADA CON MARCADORES AFLP

### GENETIC DIVERSITY IN SOYBEAN OF THE HUMID TROPICS OF MÉXICO DETERMINED BY AFLP MARKERS

Víctor Pecina Quintero<sup>1\*</sup>, Laura Maldonado Hernández<sup>2</sup>, Nicolás Maldonado Moreno<sup>1</sup>, June Simpson<sup>3</sup>, Octavio Martínez de la Vega<sup>3</sup> y Katia del Carmen Gil Vega<sup>3</sup>

<sup>1</sup>Campo Experimental Río Bravo, Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias (INIFAP). Apdo. Postal 172, C. P. 88900. Río Bravo, Tamaulipas, México. Tel: 01 (899) 934-1046. Correo electrónico: pecina.victor@inifap.gob.mx <sup>2</sup>Facultad de Ciencias Químicas, Universidad Autónoma de Tamaulipas. Apdo. Postal 1025 C. P. 88740 Reynosa, Tamaulipas, México. <sup>3</sup>Centro de Investigaciones y de Estudios Avanzados, Instituto Politécnico Nacional. Apdo. Postal 629 C. P. 36500 Irapuato, Gto., México.

\* Autor para correspondencia

#### RESUMEN

La soya [*Glycine max* (L.) Merrill] posee bajos niveles de diversidad genética en comparación con otras especies autógamas. Por lo anterior y con el fin de determinar si existe la necesidad de introducir nuevo germoplasma para generar nuevas variedades, se evaluó la diversidad genética de 24 líneas elite de soya para el trópico húmedo de México y nueve introducciones de otros países, mediante marcadores moleculares tipo AFLP. Los resultados revelaron que las líneas e introducciones no se agruparon de acuerdo a su origen geográfico, además de indicar un alto nivel de polimorfismos (60 %). Los genotipos de México son altamente similares a los de Brasil. El índice de diversidad genética fue bajo (31 %), y semejante a los reportados en otros estudios con RFLP y RAPD. Aunque la muestra de genotipos no representó a la totalidad del germoplasma de soya cultivada en el trópico húmedo de México, se considera que existe suficiente diversidad genética para generar nuevas variedades, ya que tres líneas mexicanas fueron más divergentes a nivel molecular que los materiales considerados filogenéticamente distintos de otros países (basados en características morfológicas y fisiológicas) utilizados como testigos.

**Palabras clave:** *Glycine max*, diversidad genética, polimorfismos, AFLP.

#### SUMMARY

Soybean [*Glycine max* (L.) Merrill] has a low level of genetic diversity in comparison with other self-pollinated species. In order to determine the need to introduce new germplasm for the development of new varieties, genetic diversity of soybean from the humid tropical region of México was evaluated using AFLP markers. A total of 24 lines developed in México and nine introductions from other countries were analyzed. Results revealed that lines did not group together according to their geographical origin. A high polymorphism level (60 %) was observed and the Mexican genotypes share a high level of genetic similarity with those from Brazil. The low genetic diversity index (31 %) was similar to those reported in other studies. Although the genotypes sampled did not represent all soybean germplasm in the humid tropics, results showed that the genetic diversity is to generate new cultivars, because three of the Mexican lines were more diver-

gent at the molecular level than foreign genotypes considered phylogenetically different (based on morphological and physiological traits), which were used as controls.

**Index words:** *Glycine max*, genetic diversity, polymorphism, AFLP.

#### INTRODUCCIÓN

La soya [*Glycine max* (L.) Merrill] es una especie originaria de regiones templadas, que ha sido introducida a los trópicos del mundo con los consecuentes problemas de adaptación debido a su sensibilidad al fotoperiodo. En México este cultivo fue introducido de los Estados Unidos a principios del siglo XX, y adquirió importancia a partir de 1960 cuando se evaluaron y seleccionaron las primeras variedades. Posteriormente se formó el banco de germoplasma del trópico con nuevas introducciones. Desde entonces se han generado nuevas líneas y variedades de soya a partir de hibridaciones, algunas de las cuales han perdurado durante largo tiempo en el mercado. Todo programa de mejoramiento debe contar con amplia base genética para generar las variedades mejoradas que requiere cada región. Estudios previos han revelado que las variedades de soya tienen un bajo nivel de diversidad genética, en comparación con otras especies autógamas.

Keim *et al.* (1989) encontraron que en Estados Unidos la soya posee un nivel de polimorfismo de 16 %, con base en un análisis del germoplasma con marcadores tipo RFLP (Restriction Fragment Length Polymorphisms). Posteriormente, Keim *et al.* (1992) observaron niveles de diversidad genética de 30 %, pero consideran que probablemente se sobreestimó la diversidad genética, ya que los genotipos

estudiados fueron seleccionados para representar la mayor diversidad genética posible. Así mismo, Gizlice *et al.* (1994) concluyeron que el número de ancestros que constituyen la base genética del programa de mejoramiento de soya de Norteamérica es muy limitada, ya que 85 % de la aportación genética a las variedades generadas entre 1947 y 1988 tuvo su origen en sólo 16 líneas.

En estudios más recientes, la utilización de nuevos sistemas de marcadores genéticos como los RAPD (Random Amplified Polymorphic DNA) y AFLP (Amplified Fragment Length Polymorphisms) han revelado un mayor número de polimorfismos que los RFLP (Powell *et al.*, 1996). Maughan *et al.* (1996), evaluaron 23 genotipos de soya que incluyen genotipos ancestrales (p.e., *G. soja* Sieb y Zucc); utilizaron 15 pares de iniciadores AFLP y encontraron 36 % de variantes polimórficas. Sin embargo, el grupo que incluyó únicamente genotipos de *G. max* presentó 17 % de polimorfismo. Thompson *et al.* (1998) analizaron 18 líneas ancestrales y 17 introducciones del banco de germoplasma de soya de los Estados Unidos de Norteamérica con 125 iniciadores RAPD y detectaron 34 % de polimorfismos, por lo que concluyeron que este sistema de marcadores puede ser útil para detectar la transferencia de alelos favorables a las líneas elite de los programas de mejoramiento genético.

También los microsatélites o SSR (Simple Sequence Repeat) han detectado alto número de polimorfismos en soya (Akkaya *et al.*, 1992; Rongwen *et al.*, 1995; Maughan *et al.*, 1996). Diwan y Cregan (1997) lograron distinguir 35 genotipos de soya que representan 95 % de los alelos presentes en las variedades de Norteamérica (10.1 alelos por locus), por medio de microsatélites. Liu *et al.* (2000) reportaron índices de diversidad genética de 43.9 a 66.8 % en China, mediante marcadores SSR. La técnica de AFLP ha sido reportada como un sistema de marcadores moleculares con alto potencial para detectar polimorfismos, debido al gran número de fragmentos que pueden visualizarse en cada reacción (Powell *et al.*, 1996; Maughan *et al.*, 1996). En este trabajo se evaluó la diversidad genética en germoplasma de soya del trópico húmedo de México, con el objetivo de determinar la diversidad y relaciones genéticas entre el germoplasma y establecer si existe la necesidad de realizar nuevas introducciones.

## MATERIALES Y MÉTODOS

### Material genético y aislamiento de ADN

Un total de 33 genotipos fueron evaluados, de los cuales 24 son líneas elite de soya de México, siete de Brasil, una de EE.UU. y una de Taiwán, pertenecientes al banco de germoplasma de soya para el trópico húmedo de Méxi-

co del Instituto Nacional de Investigaciones Forestales Agrícolas y Pecuarias (INIFAP) (Cuadro 1). La extracción de ADN se basó en el protocolo de Dellaporta *et al.* (1983), para lo cual se colectó 0.5 g de tejido joven de al menos 5 plantas de cada genotipo de soya sembrados en invernadero. El tejido se lavó con agua destilada estéril y se congeló a -80 °C, hasta su utilización. El ADN fue cuantificado en forma visual en geles de agarosa 1 %, con uso de una concentración conocida de ADN del fago Lambda.

### Análisis AFLP

Aproximadamente 500 ng de ADN genómico se digirieron con 1.25 U  $\mu\text{L}^{-1}$  de las enzimas de restricción *MseI* y *EcoRI* del paquete "AFLP Analysis System I" (Invitrogen Life Technologies, Carlsbad, California), en un volumen total de 25  $\mu\text{L}$  durante 4 h a 37 °C; la mezcla se calentó a 70 °C por 15 min para inactivar las enzimas de restricción. A los productos de restricción, se agregó la mezcla de ligación con los adaptadores *EcoRI* (5'-CTCGTAGACTGCGTACC-3'/3' CTGACG-CATGGTTAA-5') y *MseI* (5'-GACGATGAGTCCTGAG-3'/3'-TACTCAGGACTCAT-5'). La ligación se incubó durante 12 h a 16 °C. Posteriormente, el producto de la digestión-ligación se diluyó 1:10 en TE (Tris-HCl pH 8.0 10 mM, EDTA 1 mM) y se tomaron 5  $\mu\text{L}$  para realizar la preamplificación en un volumen final de 50  $\mu\text{L}$  con el siguiente programa: 20 ciclos de 30 s a 94 °C, 60 s a 56 °C y 60 s a 72 °C. Se incluyeron dos testigos, el negativo que incluyó todos los componentes de la reacción excepto ADN, y el positivo con ADN de tomate (*Lycopersicon esculentum* L.). Los oligonucleótidos iniciadores utilizados en la preamplificación fueron *EcoRI*+A y *MseI*+C. La reacción de preamplificación se diluyó 1:10 en TE antes de la segunda amplificación. En la amplificación selectiva se marcó radiativamente el iniciador *EcoRI*+ACT con  $^{32}\text{P}$  mediante T<sub>4</sub> cinasa. Además, se evaluaron seis combinaciones diferentes del iniciador *MseI*+3, con el siguiente programa de PCR: 1 ciclo de 30 s a 94 °C, 30 s a 65 °C y 60 s a 72 °C y después la temperatura de alineamiento disminuyó 0.7 °C cada ciclo durante 11 ciclos hasta llegar a 56 °C, y continuó con 23 ciclos de 30 s a 94 °C, 30 s a 56 °C y 60 s a 72 °C. Los productos del PCR fueron desnaturalizados mediante incubación a 90 °C por 3 min y colocados en hielo antes de ser cargados en geles de acrilamida (6 %) y separados por electroforesis durante 90 min a 2000 V. Los geles fueron secados y expuestos a película de rayos X durante 12 h.

### Análisis de datos

Se contabilizaron las bandas polimórficas y monomórficas de cada gel. Se consideró que las bandas con mismo

peso molecular en diferentes individuos son idénticas. La presencia de una banda fue indicada por uno (1) y la ausencia como cero (0).

Las relaciones genéticas entre aislados se calcularon por el coeficiente de apareamiento simple (Sokal y Sneath, 1963), con el software S-Plus Version 4.0 (S-Plus 1997). La matriz de distancias generada se utilizó para producir un dendrograma por medio del método UPGMA (Unweighted Pair Group Method with Arithmetic Average). Se usó el método de Felsenstein (1985) para obtener el intervalo de confianza para los grupos formados en cada nodo del dendrograma, por lo cual se realizó un análisis de muestreo con remplazo (Bootstrap con 1000 repeticiones) de los datos AFLP originales. Los intervalos de confianza de Felsenstein se refieren al porcentaje de veces que cada nodo del dendrograma es repetido en las 1000 muestras Bootstrap, lo que a su vez indica la robustez de los grupos formados en cada nodo. Un porcentaje cercano a 100 % sugiere que el grupo refleja la estructura de la población y que el error de muestreo es mínimo, mientras que bajos porcentajes (30 % o menos) sugieren que puede existir variación entre los genotipos de algún nodo en particular, cada vez que realiza un análisis, por lo que es necesario aumentar el tamaño de muestra. El índice de diversidad (*ID*) fue calculado mediante la ecuación:  $ID = 1 - p_i^2$ ,

donde  $p_i$  es la frecuencia del alelo  $i^n$ ; en este caso, cada alelo individual se considera como un *locus* único y a su vez como un fragmento de amplificación (Powell *et al.*, 1996).

## RESULTADOS Y DISCUSIÓN

Cada par de iniciadores AFLP amplificó entre 13 y 54 bandas, con un promedio de 38 bandas por combinación. Se detectaron 228 fragmentos amplificados de los cuales 139 (60 %) fueron polimórficos (Cuadro 2), lo que demuestra que los AFLP pueden detectar variabilidad genética aún en especies donde la diversidad es baja (Keim *et al.*, 1989; Keim *et al.*, 1992; Gizlice *et al.*, 1994). Es posible que el número de polimorfismos sea alto en comparación con otros autores, como Maughan *et al.* (1996), debido a que en este trabajo se incluyeron en el estudio variedades híbridas provenientes de diferentes líneas introducidas de otros países (Cuadro 1). Las dos líneas elite H96-1311 y H96-0092 y la variedad Huasteca 200 fueron más divergentes que los genotipos de otros países considerados filogenéticamente distintos (diferente grupo de madurez, fotoperiodo, resistencia a insectos, etc.) (Figura 2). El patrón de bandas típico de productos obtenidos por AFLP se presenta en la Figura 1.

Cuadro 1. Origen geográfico y genealogía de las líneas y variedades de soya evaluadas.

Número	Genotipo	Origen	Genealogía
1	H96-1316	México	Huasteca 200 x Santa Rosa
2	H96-1311	México	Huasteca 200 x Santa Rosa
3	Huasteca 200	México	F81-5344 x Santa Rosa
4	Huasteca 100	México	Santa Rosa x Júpiter
5	H98-1395	México	Huasteca 100 x H88-0956
6	H98-1414	México	Huasteca 100 x H88-0956
7	H98-1387	México	Huasteca 100 x H88-0956
8	H96-1153	México	Huasteca 100 x H88-0956
9	H98-2579	México	H80-1784 x Santa Rosa
10	H88-1880	México	Santa Rosa x H80-2535
11	H98-1047	México	H88-1880 x H88-3868
12	H98-1075	México	H88-1880 x H88-3868
13	H98-1028	México	H88-1880 x H88-3868
14	H98-1009	México	H88-1880 x H88-3868
15	H98-1076	México	H88-1880 x H88-3868
16	H98-1617	México	H91-0235 x H88-3868
17	H98-1092	México	H88-1880 x H91-0235
18	H96-0847	México	(GO83-22621 x Santa Rosa) x Santa Rosa
19	H96-0431	México	(GO83-22621 x Santa Rosa) x Santa Rosa
20	H96-0144	México	(H84-1617 x Santa Rosa) x Santa Rosa
21	H96-0092	México	(H84-1837 x Lamar) x Lamar
22	H94-0668	México	F85-1108 x F85-7356
23	H94-0721	México	F91-2201
24	H94-0692	México	F91-3570
25	DOKO	Brasil	
26	FT-2001	Brasil	
27	EMBRAPA-48	Brasil	
28	Carla	Brasil	
29	Tucano	Brasil	
30	Curio	Brasil	
31	Seguranza	Brasil	
32	F81-5428	USA	
33	AGS-154	Taiwán	

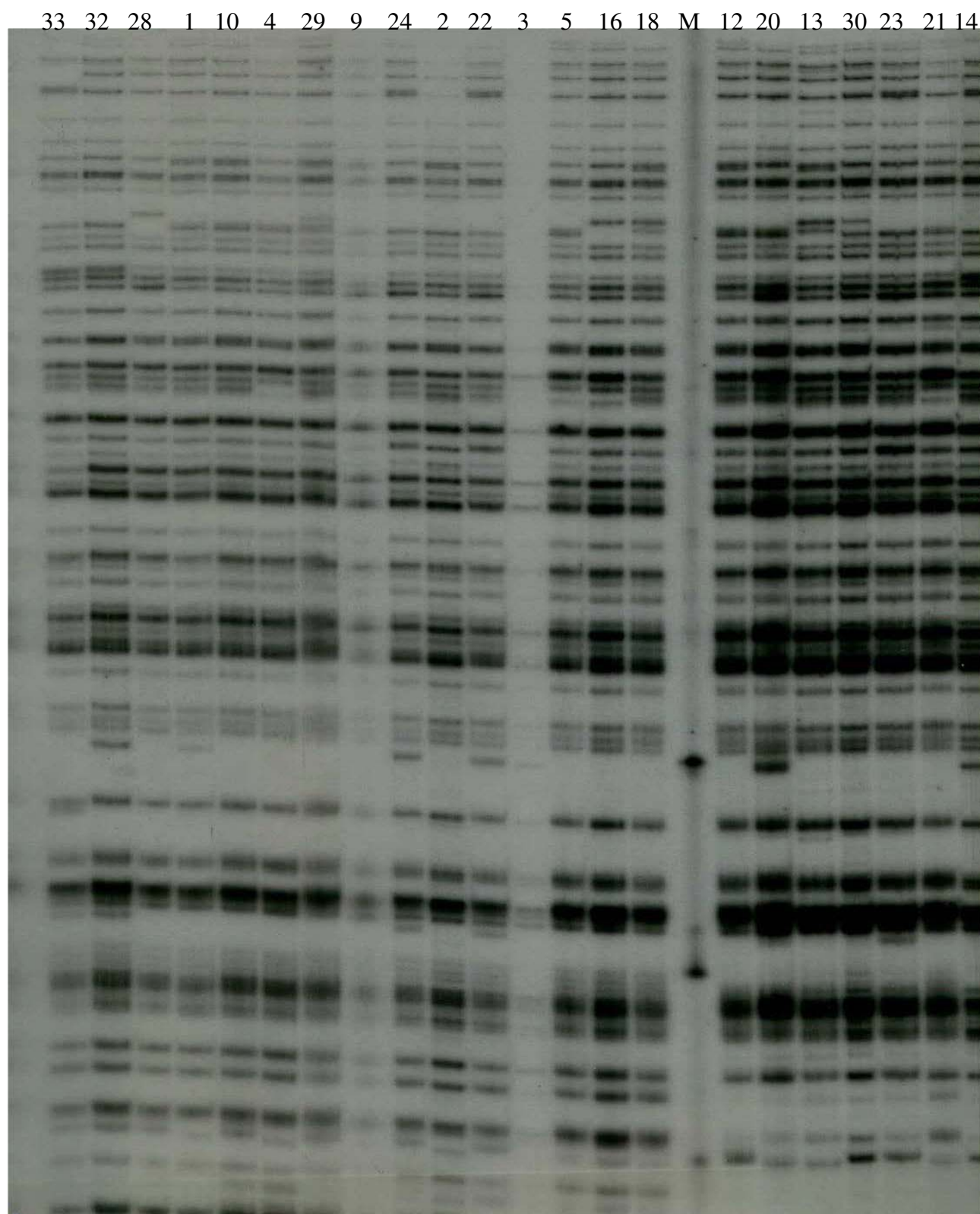


Figura 1. Productos AFLP de diferentes líneas y variedades de soya (*G. max*) del trópico húmedo de México, obtenidos con las combinaciones de oligonucleótidos ACT/CAG.

De los genotipos de soya 85 % se agruparon en dos grupos principales (Figura 2). El grupo A fue más homogéneo, con distancias genéticas que varían de 0.04 a 0.06 entre los 13 genotipos incluidos, donde la línea más divergente fue la H96-0431 con 0.08 de distancia genética. Además, en el grupo A cinco genotipos (H98-1047, H98-1009, H98-1075, H98-1076 y H98-1028) son descendientes de la misma cruce (H88-1880 x H88-3868) y la línea H98-1092 comparte un progenitor (H88-1880) con los cinco anteriores; por tanto, las pequeñas diferencias moleculares encontradas son debidas al progenitor (H91-0235),

dado que el patrón de bandas amplificadas fue muy similar. Por otra parte, aunque no se analizó por sí sola la variedad Santa Rosa, la genealogía de los demás genotipos incluidos en el grupo A (Cuadro 1), con excepción de H94-0721, muestra que todos comparten alelos de la variedad Santa Rosa por cruce directa o indirecta. También en el grupo A se observó a Curio y Seguridad, variedades de Brasil. Aunque no se cuenta con la genealogía de estas líneas, es probable que estén emparentadas, ya que la variedad Santa Rosa es originaria de este país.

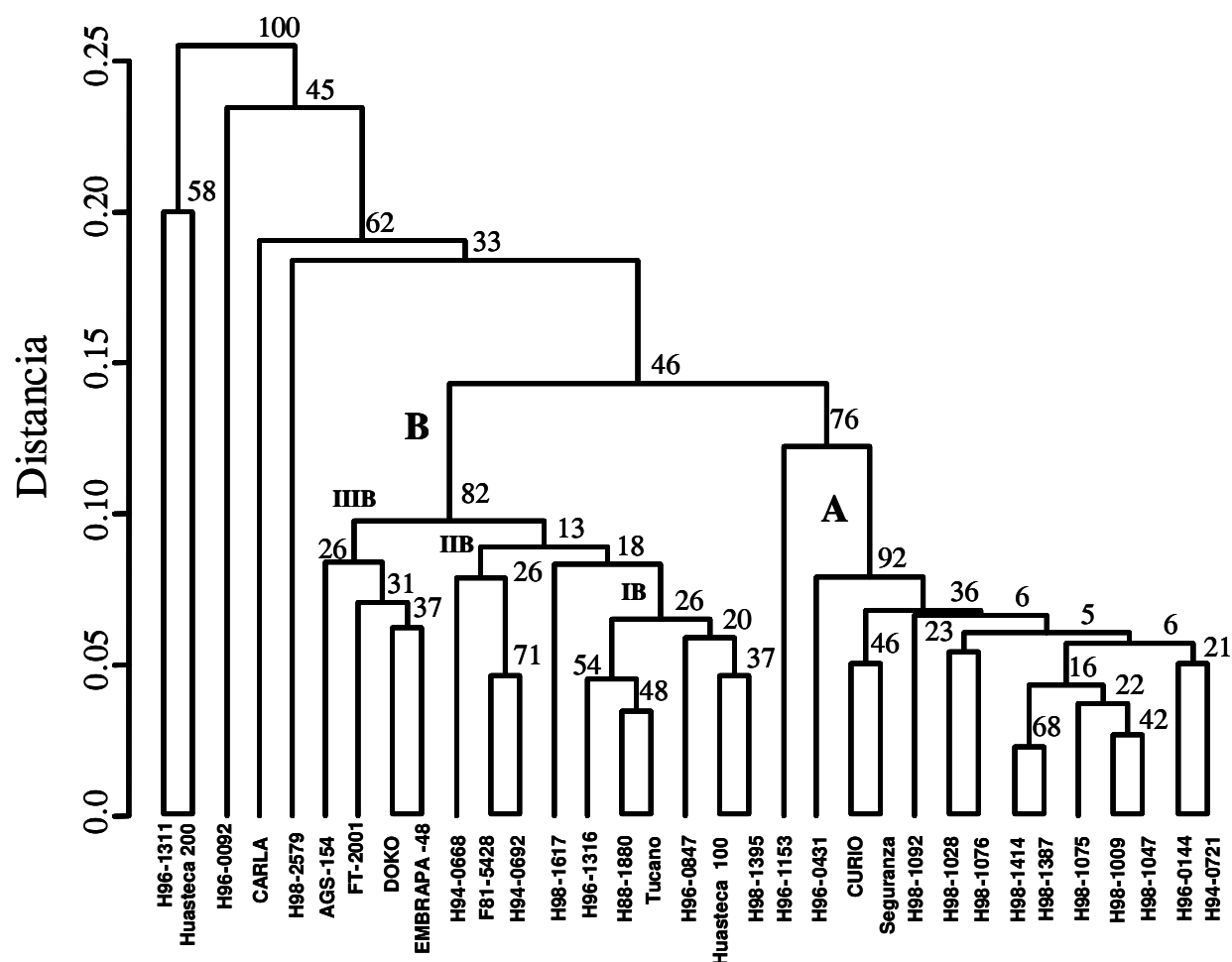


Figura 2. Dendrograma de las relaciones genéticas entre 33 líneas de soya mediante el uso de UPGMA.

Cuadro 2. Combinaciones de oligonucleótidos AFLP utilizados, número de productos amplificados y número de productos polimórficos obtenidos en cada ensayo.

Iniciador	Total de fragmentos	Fragmentos polimórficos	Polimorfismos (%)	Índice de diversidad
ACT/CAA	54	39	72	0.29
ACT/CAC	38	23	60	0.37
ACT/CAT	53	30	56	0.22
ACT/CAG	37	15	40	0.28
ACT/CTA	33	25	75	0.42
ACT/CTG	13	7	53	0.30
Total	228	139	60	0.31

El grupo B se dividió en tres subgrupos, con distancias genéticas que variaron de 0.03 a 0.19 entre genotipos. En el subgrupo 1B sólo los genotipos Tucano y H98-1617 no comparten germoplasma de Santa Rosa. Además se observó estrecha relación entre las líneas F81-5428 originaria de Florida (EE.UU.) y las líneas mexicanas H94-0692 y H94-0668 que fueron derivadas de otra línea introducida de dicho estado (subgrupo IIB). En el subgrupo IIIB se ubican tres genotipos de Brasil (EMBRAPA-48, Doko y FT-2001) y uno de Taiwán (AGS-154), con distancias genéticas de 0.06 a 0.08.

Es notable que los genotipos más divergentes con respecto al total del germoplasma que se incluyó, fueron cuatro líneas de México y una de Brasil (H98-2579, H96-0092, Huasteca 200, H96-1311 y Carla), mismos que presentan grandes distancias genéticas entre líneas (0.17, 0.19, 0.22, 0.25 y 0.28 en promedio, respectivamente). Parte de estos materiales se desarrollaron de cruza en las que intervinieron la variedad Lamar de Mississippi, EE.UU. (Hartwig *et al.*, 1990) (resistencia a insectos defoliadores) y la línea F81-5344 (periodo juvenil largo) de Florida, con materiales de Brasil (mayor sensibilidad al fotoperiodo). Huasteca 200 y H96-1311 se agruparon en un nodo, lo que coincide con la genealogía (Cuadro 1), ya que Huasteca 200 es progenitor de H96-1311, mientras que la línea hermana H96-1316 que también fue derivada de la misma cruce, resultó más divergente debido a las recombinaciones que probablemente ocurrieron durante su desarrollo. Algo similar ocurrió con los genotipos H96-1311 y H96-0144 que en el dendrograma aparecen opuestos, no obstante que tienen como progenitor común a Santa Rosa. Por tanto la divergencia entre los dos genotipos puede ser explicada por efecto del aporte genético del padre que es distinto, además del efecto de selección que realiza el investigador, al conservar sólo genotipos con características agronómicas deseables (rendimiento, precocidad, sensibilidad al fotoperiodo, resistencia a plagas y enfermedades, etc.).

Tales resultados evidencian que las líneas de soya no quedaron agrupadas de acuerdo con su origen geográfico, ya que los materiales de Brasil se entremezclan con las lí-

neas desarrolladas en México. Esto se debe a que dichas líneas son descendientes de cruza con progenitores introducidos de Brasil tales como la variedad Santa Rosa. Lo anterior también se pudo deber al pequeño número de muestras y a la necesidad de emplear un mayor número de pares de iniciadores para generar mayor información. Ambos factores afectaron también los intervalos de confianza de Felsenstein (1985), ya que se observaron bajos valores (>30 %) para algunos nodos en particular, y reflejan que las relaciones genéticas entre genotipos no son tan robustas (Figura 2).

El índice de diversidad genética observado fue bajo (31 %), y es similar a los reportados en otros estudios con diferentes sistemas de marcadores (30 % por Keim *et al.* 1992; 31 % por Powell *et al.*, 1996; 43 % por Liu *et al.* 2000). Además, si se considera que la mayor parte de los genotipos evaluados son líneas élite del programa de soya del INIFAP para el trópico húmedo de México y que no representan en su totalidad la variación genética del país, puede inferirse que existe suficiente diversidad genética posible de explotarse, ya que 60 % de los productos amplificados fueron polimórficos.

## CONCLUSIONES

El uso de los marcadores moleculares AFLP permitió detectar alto nivel de polimorfismos genéticos en las líneas y progenitores de soya del Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias, para el trópico húmedo de México. Las relaciones genéticas entre genotipos indicaron que los materiales de México comparten un alto nivel de similitud genética con los de Brasil. Aunque la muestra de genotipos no representó a la totalidad del germoplasma de soya del trópico húmedo, se considera que existe suficiente variabilidad genética para generar nuevas variedades, por lo que no es necesario realizar nuevas introducciones de germoplasma en forma inmediata, a menos que se desee incorporar alguna característica en particular, ausente en el germoplasma actual.

## AGRADECIMIENTOS

Investigación parcialmente financiada por el proyecto CONACYT-SIREYES, Clave:20000606009.

## BIBLIOGRAFÍA

- Akkaya M S, A A Bhagwat, P B Cregan (1992) Length polymorphisms of simple sequence repeat DNA in soybean. *Genetics* 132:1131-1139.
- Dellaporta S L, J Wood, J B Hicks (1983) A plant DNA miniprep-  
aration. Version II. *Plant Molec. Biol. Report* 1:19-21.

- Diwan N, P B Cregan (1997)** Automated sizing of fluorescent-labeled simple sequence repeat (SSR) markers to assay genetic variation in soybean. *Theor. Appl. Gen.* 95:723-733.
- Felsenstein J (1985)** Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39: 783-791.
- Gizlice Z, T E Carter Jr, J W Burton (1994)** Genetic base for north American public soybean cultivars released between 1947 and 1988. *Crop Sci.* 34:1143-1151.
- Hartwig E E, L Lambert, T C Kilen (1990)** Registration of Lamar soybean. *Crop. Sci.* 30:231.
- Keim P, R C Shoemaker, R G Palmer (1989)** Restriction fragment length polymorphism diversity in soybean. *Theor. Appl. Gen.* 77:786-792.
- Keim P, W Beavis, J Schupp, R Freestone (1992)** Evaluation of soybean RFLP marker diversity in adapted germplasm. *Theor. Appl. Gen.* 85:205-212.
- Liu F, F Y Dong, J J Zou, S Y Chen, B C Zhuang (2000)** Soybean germplasm diversity and genetic variance detected by microsatellite markers. *Yi Chuan Xue Bao* 27:628-633.
- Maughan P J, M A Saghai Maroof, G R Buss, G M Huestis (1996)** Amplified fragment length polymorphism (AFLP) in soybean: species diversity, inheritance, and near-isogenic line analysis. *Theor. Appl. Gen.* 93:392-401.
- Powell W, M Morgante, C Andre, M Hanafey, J Vogel, S Tingey, A Rafalski (1996)** The comparisons of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Molec. Breed.* 2:225-238.
- Rongwen J, M S Akkaya, A A Bhagwat, U Lavi, P B Cregan (1995)** The use of microsatellite DNA markers for soybean genotype identification. *Theor. Appl. Gen.* 90:43-48.
- Sokal R R, A P H Sneath (1963)** *Principles of Numerical Taxonomy*. W.H. Freeman and Co. Francisco, USA. 359 p.
- [S-Plus] S-Plus User's Guide Versión 4.0 (1997)** Seattle: Math Soft. 525 p.
- Thompson J A, R L Nelson, L O Vodkin (1998)** Identification of diverse soybean germplasm using RAPD markers. *Crop Sci.* 38:1348-1355.