

Desarrollo de una base de datos genéticos para la caracterización del germoplasma argentino de soja

Construction of a genetic database to characterize the argentine soybean germplasm

OLSINA, Cesar¹; CAIRO, Carlos²; PESSINO, Silvina²

¹ Bolsa de Comercio de Rosario
Rosario, Santa Fe, Argentina, colsina@bcr.com.ar

² Facultad de Ciencias Agrarias, UNR
Parque Villarino, S2125ZAA, Zavalla, Argentina,
ccairo@unr.edu.ar, pessino@arnet.com.ar

Resumen

Argentina es actualmente el primer exportador mundial de aceites y harinas de soja y el tercer productor mundial de semillas de soja. El objetivo de este trabajo fue desarrollar un banco de datos genéticos basado en marcadores moleculares de SSR (Repeticiones de Secuencia Simple, o Microsatélites) y AFLP (polimorfismos en el largo de los fragmentos amplificados) para caracterizar el germoplasma argentino de soja, facilitar la identificación de variedades y asistir al mejoramiento genético. Se estudiaron 51 loci SSR y 15 combinaciones de cebadores AFLP en los 86 cultivares de soja más utilizados en Argentina. Los SSR revelaron en promedio 5 alelos por locus, con una diversidad génica media de 0,523. Los AFLP detectaron en promedio 17 bandas polimórficas por combinación de oligonucleótidos, con una diversidad génica media de 0,1924. Los 86 cultivares estudiados mostraron coeficientes de similitud genética medios de 0,292 y 0,921 para SSR y AFLP, respectivamente. La base de datos de SSR se validó en forma exitosa a través de la identificación de muestras incógnitas remitidas al laboratorio. Además, se seleccionó un subconjunto de 20 marcadores SSR altamente informativos que permite discriminar cualquier accesión del banco de manera simple y eficiente, y que pueden ser utilizados en estudios rápidos de identificación o diferenciación de nuevos cultivares. Los marcadores AFLP produjeron dendrogramas basados en la distancia genética donde los materiales agruparon de acuerdo al obtentor. Estos resultados sugieren que los marcadores AFLP reflejan fielmente las relaciones de parentesco entre las variedades, resultando apropiados para seleccionar materiales genéticamente divergentes.

Palabras clave:

Glycine max-marcadores moleculares-huellas genéticas

Introducción

La soja (*Glycine max* L. Merrill) es una leguminosa nativa del Asia Oriental. Las evidencias arqueológicas más antiguas de su uso en China, Corea y Japón datan de circa 1000 AC, aunque se calcula que su utilización como alimento y en la composición de medicamentos comenzó hace aproximadamente 5000 años (Hymowitz & Newell, 1981; Dong et al., 2004; Zhao & Gai, 2004). En 1765 fue trasladada desde China a los Estados Unidos, observándose una notable expansión del cultivo a partir de 1840. El primer registro de su ingreso en Argentina data de 1862 (Martínez, 2010).

A principios de 1960 la producción argentina de soja comenzó a crecer sostenidamente, incrementándose desde 136.700 hasta 52.700.000 toneladas en el periodo 1971-2010 (Sistema Integrado de Información Agropecuaria, //www.sii.gov.ar/).

Abstract

Argentina is currently the leading country in terms of soybean oil and flour exports, and ranks third in soybean seed production. The aim of this work was to develop a genetic database based on SSR (Simple Sequence Repeats, or Microsatellites) and AFLP (Amplified Fragment Length Polymorphisms) markers, in order to characterize the argentine soybean germplasm, allow cultivar identification and provide assistance to breeding. Fifty-one SSR loci and 15 AFLP primer combinations were analyzed on the 86 soybean cultivars most used in Argentina. The SSR markers revealed an average of 5 alleles per locus and a genic diversity of 0.523. The AFLP markers showed an average of 17 polymorphic bands per primer combination and a genic diversity of 0.1924. The 86 cultivars under analysis showed coefficients of mean genetic similarities of 0.292 and 0.921 for SSR and AFLP, respectively. The SSR database was successfully validated for the identification of unknown samples. A subset of 20 highly-informative SSR markers was selected to achieve an effective discrimination of any databank accession by means of a simple and efficient procedure, in order to be used for rapid identification analysis and differentiation of new cultivars. The AFLP markers produced dendrograms based on genetic distance in which the plant materials grouped according to the identity of the breeder. These results suggest that AFLP markers accurately reflect the parentage relationship among cultivars, and are appropriate to the selection of genetically divergent materials.

Keywords:

Glycine max - molecular markers - genetic fingerprints

Pasó de ocupar 37.700 hectáreas en la campaña 1970/71 a 18.343.000 hectáreas en 2010 (informe GEA-BCR 2010, www.bcr.com.ar). Los principales factores que ocasionaron el asombroso aumento observado en la producción de soja en Argentina fueron el uso de la siembra directa, los altos precios internacionales del complejo soja debido a la importante demanda asiática, el desarrollo de un gran número de variedades adaptadas a las distintas zonas de producción argentinas y la introducción de cultivares transgénicos resistentes a herbicidas (Olsina, 2012). Argentina es actualmente: 1) el primer exportador mundial de aceites y harinas de soja, con un volumen de exportación de 5,4 millones de toneladas de aceite y 27,3 millones de toneladas de harina de soja en el año 2010; 2) el tercer productor mundial de soja, con 52,7 millones de toneladas, luego de Estados Unidos y Brasil, según datos de la

campana 2009/2010 (Sistema Integrado de Información Agropecuaria, //www.siaa.gov.ar/). Los granos, harinas y aceites representan el 25 % del valor total de las exportaciones argentinas (Leavy & Dewes, 2009). Considerando la producción conjunta de Argentina y Brasil, los países del MERCOSUR ocupan el primer lugar en la producción de soja a nivel mundial.

En Argentina existe un número significativo de semilleras locales que se dedican al mejoramiento de este cultivo. En el país está vigente una legislación que regula los derechos de propiedad intelectual de las obtenciones vegetales, y controla su propagación y comercialización. Se trata de la Ley de Semillas y Creaciones Fitotécnicas N° 20.247, que establece que un cultivar puede ser objeto de título de propiedad cuando reúna condiciones de distinguibilidad, homogeneidad y estabilidad que permitan la caracterización del mismo (Bonamico et al., 2004). El organismo responsable de su aplicación es el Instituto Nacional de Semillas (INASE), dependiente del Ministerio de Agricultura, Ganadería, Pesca y Alimentos de la Nación. Por su parte, el país está adherido a la Unión Internacional para la Protección de Nuevas Variedades de Plantas (UPOV/TG/1/2, 1979), lo que permite extender la confianza nacional e internacional en el ámbito de la concesión de propiedad al creador de una nueva variedad (Olsina, 2012).

Tradicionalmente la determinación de las condiciones de distinguibilidad, homogeneidad y estabilidad de un cultivar se realizaban a través de análisis morfológicos y bioquímicos. Pero en las últimas décadas, la implementación de las técnicas de marcadores moleculares basadas en el análisis de la estructura del ADN permitieron un estudio directo sobre el genotipo en lugar del fenotipo (Powell et al., 1996). Los marcadores moleculares tienen la ventaja de revelar la estructura de numerosos *loci* en la secuencia de bases del ADN, facilitando la identificación de polimorfismos genéticos, que en su mayoría no se reflejan en el fenotipo. En este contexto, la Oficina de Protección de Variedades Vegetales del Departamento de Agricultura de los Estados Unidos (USDA) acepta actualmente los perfiles alélicos de marcadores moleculares de secuencias simples repetitivas (SSR) como evidencia adicional de la singularidad de un nuevo cultivar (Diwan & Cregan, 1997; Song et al., 1999). Asimismo, diferentes organismos internacionales como la Organización Internacional para la Estandarización (ISO), la Asociación Internacional para el Análisis de Semillas (ISTA) y la Unión Internacional para la Protección de Nuevas Variedades de Plantas (UPOV) están realizando permanentes validaciones y estandarizaciones de marcadores moleculares aplicados a la identificación de variedades de los cultivos mayores, entre ellos la soja (Olsina, 2012).

Es un hecho ampliamente aceptado que *G. max* posee un bajo nivel de diversidad morfológica y genética. Por ello, la utilización de marcadores moleculares es particularmente útil en lo que se refiere a la diferenciación de cultivares de esta especie. Inicialmente se reportaron estudios basados en marcadores de RFLP (restriction fragment length polymorphisms) aplicados a su caracterización (Kein et al., 1992; Shoemaker et al., 1992). Luego se hizo más común el uso de métodos basados en la reacción de amplificación en cadena de la polimerasa (PCR) (Mullis et al., 1986), entre ellos los de RAPD (Random Amplified Polymorphic DNA) (Williams et al., 1990; Welsh & McClelland 1990), AFLP (Amplified Fragment Length Polymorphism) (Vos et al., 1995) y SSR (Simple Sequence Repeats), también conocidos como microsatélites (Weber, 1990; Akkaya et al., 1992).

El método de RAPD es simple, necesita mínimas cantidades de ADN y no requiere información previa de secuencia. Consiste en

la amplificación del ADN al azar, usando cebadores decámeros únicos de secuencia variable (Williams et al., 1990; Welsh & McClelland 1990). Sin embargo, se ha demostrado que en la soja los RAPDs tienen poca utilidad, debido al bajo nivel de polimorfismo detectado (Giancola et al., 2002; Powell et al., 1996; Brown-Guedira et al., 2000). Por otro lado, los AFLP (Vos et al., 1995) son muy efectivos porque revelan una gran cantidad de marcadores por unidad experimental, son muy reproducibles, no necesitan información de secuencia previa y requieren una baja cantidad de ADN de partida. La técnica de AFLP se basa en la digestión del ADN genómico simultáneamente con dos endonucleasas de restricción, una de ellas de corte frecuente y la otra de corte poco frecuente. Luego se ligan adaptadores específicos para cada uno de los extremos cohesivos generados por estas enzimas y se pre-amplifican los fragmentos utilizando cebadores complementarios a los adaptadores, que pueden o no presentar bases adicionales al azar en el extremo 3'. Finalmente, una porción seleccionada de los fragmentos se re-amplifican utilizando cebadores complementarios a los adaptadores, que poseen una, dos o tres bases de secuencia al azar en el extremo 3' (Vos et al., 1995). La utilización de diferentes combinaciones de cebadores con distintas bases al azar asegura la amplificación de diferentes grupos de fragmentos genómicos. Los AFLP son una buena opción cuando se quiere obtener una caracterización rápida de genotipos, ya que se puede generar una cantidad considerable de marcadores a partir de una única combinación de cebadores (Mueller & LaReesaWolfenbarger, 1999). Se han utilizado exitosamente para caracterizar el germoplasma en cultivares comerciales de soja como así también en poblaciones naturales (Maughan et al., 1996; Ude et al., 2003; Bonato et al., 2006). Por último, los SSR son marcadores con características únicas que los hacen valiosos en los experimentos de identificación, debido a su naturaleza multialélica y codominante, su abundancia relativa y su amplia cobertura genómica. Requieren solo pequeñas cantidades de ADN de partida y son muy reproducibles entre laboratorios (Powell et al., 1996; Jones et al., 1997). Sin embargo, deben ser desarrollados a partir del conocimiento de las secuencias específicas que rodean un sector genómico de repeticiones cortas, de entre dos y seis bases, localizadas en tándem, ya que para su obtención se diseñan cebadores que amplifiquen la región completa donde está incluida la repetición. Este requerimiento de conocimiento de secuencia previo puede representar un inconveniente en la caracterización de especies poco estudiadas, pero actualmente no representa un obstáculo para los investigadores que estudian las especies mayores, ya que en las bases de datos públicas suelen estar disponibles numerosas secuencias de oligonucleótidos para amplificar marcadores SSR desarrollados por otros equipos de investigación. Respecto al caso particular de la soja, el grupo del Dr. Cregan (Servicio de Investigación en Agricultura dependiente del Departamento de Agricultura de los Estados Unidos, ARS-USDA, Beltsville) llevó a cabo la identificación de numerosos *loci* de microsatélites y diseñó cebadores específicos para SSR. Finalmente se determinó la posición en el genoma de estos marcadores mediante mapeo genético (Akkaya et al. 1992; Cregan et al. 1994). Las secuencias de cada *locus* SSR, cebadores, condiciones experimentales para su amplificación y ubicación en el mapa de ligamiento de soja, se pueden encontrar en forma libre en internet en la página web: www.soybase.org. Se ha reportado un elevado nivel de polimorfismo en los *loci* de SSR de soja (Akkaya et al. 1992; Cregan et al. 1994; Rongwen et al. 1995; Diwan & Cregan 1997; Song et al. 1999; Brown-Guerida et al. 2000; Narvel et al. 2000; Meesang et al. 2001). Analizando 20 SSR, Diwan and Cregan (1997) fueron capaces de distinguir cultivares de soja comerciales que eran considerados idénticos sobre la base de análisis de RFLP, morfología y marcadores de pigmentación. Brown-Guerida et al. (2000) pudieron identificar grupos

relacionados de germoplasma de soja, observando mayor diversidad con tres SSR que con 46 RAPDs. Priolli et al. (2002) diferenciaron grupos morfológicamente similares de 186 cultivares de soja brasileña utilizando únicamente 12 marcadores SSR.

En Argentina existe un único antecedente de caracterización molecular del germoplasma de soja. Se trata del estudio publicado por Giancola et al. (2002), donde se utilizaron distintos tipos de marcadores moleculares y descriptores clásicos con el fin de caracterizar y diferenciar el conjunto de las 100 variedades de soja de uso comercial más utilizadas en ese momento. Los autores concluyeron que la variabilidad genética de los cultivares comerciales de soja argentina era baja. Aun así, se pudieron obtener huellas genéticas (fingerprints) discriminatorias para las 100 variedades analizadas. De los marcadores moleculares utilizados (AFLP, SSR y RAPD), los marcadores SSR mostraron ser los más efectivos, además de manifestar la correlación más alta con los descriptores clásicos (Giancola et al. 2002). Según estos autores, los marcadores moleculares de SSR podrían aplicarse eficazmente para la clasificación del germoplasma de soja y su protección intelectual (Giancola et al. 2002).

Materiales y Métodos

Material Vegetal

Se investigaron 86 cultivares comerciales de soja liberados para su comercialización en Argentina (Tabla 1). El material genético fue obtenido a partir de los diferentes criaderos originadores (ACA, A. L. Marchionni, Don Mario, FN Semillas, La Tijereta, Nidera, Pioneer, Relmó, SPS, Syngenta). En forma adicional, se analizaron muestras incógnitas provenientes de multiplicadores o semilleros, cuya identidad genética probable o cultivar a contrastar fue indicada por los mismos (Tabla 2).

Extracción y cuantificación de ADN genómico

Para cada cultivar o muestra, se molieron 50 semillas en un molino a cuchillas A11 Basic (IKA-Works, Inc) con el objeto de representar en forma homogénea la constitución alélica (Diwan & Cregan, 1997). Una fracción de 200 mg de semillas molidas se procesó con el equipo de extracción de ADN genómico DNeasy Plant Extraction kit (Qiagen) de acuerdo a las instrucciones del fabricante. El ADN obtenido fue cuantificado utilizando un espectrofotómetro de fluorescencia Modulus (Turner Biosystems, Sunnyvale) a una longitud de onda de absorción y emisión de 360 y 415 nm, respectivamente. A 3 mL de una solución conteniendo 10 mM Tris (Promega), 1 mM EDTA (Promega), 200 mM NaCl (Merck) pH 7.4, Hoesch 33258 0,15 µg/mL (Sigma-Aldrich) se le adicionaron 5 µL de extracto de ADN. La curva de calibración del fluorómetro se realizó con 1000; 750; 500; 250 y 100 ng de ADN del fago Lambda (Promega). Los ADN se diluyeron con agua ultrapura hasta alcanzar una concentración de 10 ng/µL para los análisis de SSR y de 150 ng/µL para los análisis de AFLP, y se conservaron a -20°C hasta su uso. Estas diluciones y la posterior preparación de los cócteles de reacción de PCR fueron llevadas a cabo en una campana de flujo laminar NuAire® (Plymouth, Minnesota), de manera de minimizar la posibilidad de contaminaciones.

Experimentos de SSR

La técnica de SSR se implementó en base al polimorfismo y la cobertura genómica reportados previamente para soja (Cregan et al., 1999; Song et al., 2004). Se seleccionaron 51 marcadores SSR obtenidos a partir de Soybase (www.soybase.org) (Tabla 3). Las reacciones de PCR se llevaron a cabo según se describe en Akkaya et al. (1995), con pequeñas modificaciones. Las mezclas

de reacción utilizadas en el estudio de Giancola et al. (2002) habían sido inscriptas antes de la aparición del evento GTS 40-3-2 de resistencia al herbicida glifosato, por lo que su participación en la siembra actual es muy baja o nula. La mayoría de estas variedades han caducado su derecho de propiedad hace ya algunos años, por lo que la base genética de los cultivares de soja argentina utilizados en la actualidad puede haber cambiado. Es por eso que resulta de interés realizar una actualización de la caracterización de los cultivares argentinos de soja, para proporcionar herramientas moleculares que puedan ser aplicadas a la protección de propiedad intelectual vegetal y la implementación de programas de mejoramiento. Este trabajo tuvo como objetivo desarrollar una metodología basada en matrices de datos genéticos de SSR y AFLP que pueda utilizarse para clasificar en forma inequívoca las variedades comerciales de soja más utilizadas actualmente en nuestro país, para contribuir a la protección de la propiedad intelectual, asegurar la distinguibilidad de nuevos materiales y aportar información aplicable a la selección eficaz de variedades para el mejoramiento.

de reacción incluyeron 60 ng de ADN genómico, solución amortiguadora de Platinum Taq 1 X, 1,5 mM MgCl₂, 100 nM de cada cebador, 0,5 U de Platinum Taq ADN Polimerasa (Invitrogen) y 0,2 µM de dNTP (Promega), en un volumen final de 25 µL. Las condiciones para la amplificación fueron: desnaturalización inicial del ADN a una temperatura de 95°C durante 3 minutos, 40 ciclos de desnaturalización a una temperatura de 95°C durante 50 segundos, hibridación a 48°C durante 40 segundos y extensión a 72°C durante 60 segundos, en un ciclador térmico programable modelo ABI 9600 (Applied Biosystems). A los productos de amplificación obtenidos se les agregaron 5 µL de solución de siembra (formamida 95% P/V, NaOH 10 mM y azul de bromofenol 2 mg/ml) y se los almacenó a -20°C hasta su uso. Luego fueron resueltos en geles de 6 % de acrilamida (BioRad), 0,3 % de metilen-bisacrilamida (BioRad) y 7,5 M de urea (Invitrogen), 360 µL de persulfato de amonio al 10 % (APS, Sigma Aldrich), 36 µL de N,N,N₂,N₁-tetrametil-etilendiamino (TEMED, Sigma Aldrich) y corridos en una cuba vertical Sequigen GT (BioRad) utilizando como tampón de corrida 0,5X TBE [50 mM de Tris, 50 mM de ácido bórico y 1 mM de EDTA (pH 8)] en el reservorio superior y 1X TBE en el inferior. Previo a la corrida electroforética, se realizó una precorrida de la matriz a una potencia constante de 90 W por aproximadamente 1 hora hasta que el gel alcanzó una temperatura homogénea de 50 °C. Se cargaron 5µL de cada producto de amplificación por calle del gel y al menos 3 calles con marcador de peso molecular de 10 pares de base de separación entre bandas (10 bp DNA Ladder, Invitrogen). La corrida electroforética se realizó a una potencia constante de 85 W por 2 horas, con una temperatura del gel de 50°C. Para visualizar los marcadores se utilizó el método de tinción con nitrato de plata descrito en el manual "DNA Silver Staining System" (Promega). Los pasos que se siguieron se detallan a continuación: fijación del gel en ácido acético al 10 % por 20 minutos con agitación suave; 3 lavados de 2 minutos con agua destilada con agitación suave; tinción por incubación de 30 minutos con agitación suave en una solución conteniendo: 1 g/L de nitrato de plata (MERCK) y 0,05% de formaldehído; lavado con agua destilada por 10 segundos; revelado con 2 L de solución reveladora conteniendo 30 g/l de carbonato de sodio (MERCK), 0,05 % de formaldehído y 2,5 mg/L de tiosulfato de sodio hasta

Tabla 1: Cultivares de soja utilizados en el presente estudio. Se indica el grupo de madurez y el obtentor.

Cultivar	Obtentor	Grupo de madurez	Cultivar	Obtentor	Grupo de madurez
ACA360	ACA	III	TJS2068	La Tijereta	VII
ACA420	ACA	IV	TJS2170	La Tijereta	VII
ACA470	ACA	IV	TJS2171IR	La Tijereta	VII
ACA480	ACA	IV	TJS2130	La Tijereta	III
ACA530	ACA	V	A3901RG	Nidera	III
ACA570	ACA	V	A6445RG	Nidera	VI
ALM3530	A. L. Marchionni	III	A5417RG	Nidera	V
ALM4200	A. L. Marchionni	IV	A5520RG	Nidera	V
ALM4500	A. L. Marchionni	IV	A4910RG	Nidera	IV
ALM4650	A. L. Marchionni	IV	A4725RG	Nidera	IV
DM2000	Don Mario	II	A5901RG	Nidera	V
DM2200	Don Mario	II	A4303RG	Nidera	IV
DM3100	Don Mario	III	A8000RG	Nidera	VIII
DM3700	Don Mario	III	A5409RG	Nidera	V
DM4200	Don Mario	IV	A6019RG	Nidera	VI
DM4400	Don Mario	IV	A4505RG	Nidera	IV
DM4600	Don Mario	IV	A5009RG	Nidera	IV
DM4670	Don Mario	IV	A7636RG	Nidera	VII
DM4800	Don Mario	IV	A4613RG	Nidera	IV
DM4870	Don Mario	IV	94B73	Pioneer	IV
DM5800	Don Mario	V	Rosario65	Relmó	VI
DM6200	Don Mario	VI	ANTA82	Relmó	VIII
DM50048	Don Mario	V	Mercedes70	Relmó	VII
DM5.1i	Don Mario	V	Ayelen22	Relmó	II
DM8002	Don Mario	VIII	Azul35	Relmó	III
DM4250	Don Mario	IV	Maravilla45	Relmó	IV
DM3500	Don Mario	III	Natalia49	Relmó	IV
DM5.2i	Don Mario	V	NvaMaría55	Relmó	V
DM4970	Don Mario	IV	Rafaela58	Relmó	V
DM7.0i	Don Mario	VII	SPS3900	SPS	III
DM5.8i	Don Mario	V	SPS4X0	SPS	IV
DM5.5i	Don Mario	V	SPS4X4	SPS	IV
FN360	FN semillas	III	SPS4500	SPS	IV
FN410	FN semillas	IV	SPS500	SPS	V
FN485	FN semillas	IV	SPS5X5	SPS	V
TJS2136	La Tijereta	III	SPS6X0	SPS	VI
TJS2138	La Tijereta	III	SPS6X2	SPS	VI
TJS2145	La Tijereta	IV	SPS8X0	SPS	VIII
TJS2049	La Tijereta	IV	SPS9800	SPS	IX
TJS2156	La Tijereta	V	NKCoker68	Syngenta	VI
TJS2164	La Tijereta	VI	NK37-00	Syngenta	III
TJS2165	La Tijereta	VI	NKCoker75	Syngenta	VII
TJS2139	La Tijereta	III	NKMireya	Syngenta	IV

la observación de las bandas; finalización de la reacción de revelado por el agregado de 1 L de ácido acético al 10 % por 2 minutos; 2 lavados de 2 minutos con agua destilada; secado en estufa a 37 °C. Los patrones de bandas se analizaron observando los vidrios sobre un transiluminador de luz blanca. Finalmente, los geles teñidos fueron registrados para su documentación utilizando un escáner DCP-8065 DN (Brother).

Experimentos de AFLP

La generación de marcadores moleculares de AFLP se realizó según lo descrito en Vos *et al.* (1995), con las modificaciones

pertinentes para la utilización de tinción con nitrato de plata en lugar del empleo de radioactivo (Olsina, 2012). Los oligonucleótidos utilizados en los experimentos de AFLP fueron diseñados por Key Gene (AFLP Protocol for Public Release, Version 2.0, Netherlands, 1994) y sintetizados por Invitrogen Life Technologies. El ADN genómico de las líneas analizadas fue digerido con dos enzimas de restricción simultáneamente: *EcoRI* (G/AATTC), de corte poco frecuente y *MseI* (T/TAA), de corte frecuente. Las mezclas de digestión enzimática incluyeron: 600 ng de ADN genómico, 1X de amortiguador NEBEcoRI 1X (50 mM NaCl 100, mM, Tris-HCl, 10 mM MgCl₂, 0.025 %

Tabla 2: Muestras incógnitas de soja utilizadas en el presente estudio

Muestra incógnita	Identidad putativa ^a	Marcadores utilizados en el análisis ^b
11209	A5009RG	Satt 324; 173; 30; 175; 294; 242; 373; 009; 147; 141; 197; 191; 458; 308; 237; 436; 453; 511; 281; 311; 181; 180; 94 y 005
11210	DM4870	Satt 324; 173; 30; 175; 294; 242; 373; 009; 147; 141; 197; 191; 458; 308; 237; 436; 453; 511; 281; 311; 181; 180; 94 y 005
11995	A4613RG	Satt 534; 324; 173; 30; 70; 231; 175; 242; 373; 09; 147; 38; 197; 243; 475; 308 y 237
1023	DM4870	Satt 534; 324; 173; 30; 42; 70; 231; 175; 242; 414 009; 147; 038; 141; 197; 253; 191; 114; 117; 577;
1262	DM4200 DM3700	Satt 242; 141; 475; 197; 458; 308; 177; 114; 577; 175; 42; 324; 173; 30; 373; 42
1355	A6019RG	Satt 534; 324; 173; 231; 294; 242; 373; 191; 030; 577; 147; 141; 453; 197; 237; 175; 009; 038; 253; 308
1357	A7636RG	Satt 534; 324; 173; 231; 294; 242; 373; 197; 191; 030; 577; 147; 141; 453; 197; 237; 175; 009; 038; 253; 308
1269	DM4600 A4303RG	Satt 534; 324; 173; 231; 294; 242; 373; 197; 191; 030; 577; 147; 141; 453; 197; 237; 175; 009; 038; 253; 308

^a identidad putativa declarada por el solicitante, quien requirió su comparación con cultivares testigo.

^b Marcadores de SSR utilizados para el análisis particular de cada muestra

Triton X-100 pH 7.5, New England Biolab, Inc), 5 U de cada enzima de restricción, 100 µg/mL de albúmina sérica bovina (New England Biolabs, Inc) y agua tridestilada estéril hasta completar un volumen final de 50 µL. La mezcla de reacción fue incubada a 37 °C durante 2 horas. Posteriormente, se procedió a ligar de los adaptadores de AFLP a los fragmentos digeridos. La mezcla de reacción de ligación incluyó: 5 pmol del adaptador *EcoRI*, 50 pmol del adaptador *MseI*, 1X de solución tampón de T4 DNA Ligase (50 mM Tris-HCl, 10 mM MgCl₂, 1 mM ATP, 10 mM DTT pH 7.5 New England Biolabs, Inc), 1 U de enzima T4 DNA Ligasa (New England Biolabs, Inc) y agua ultrapura hasta completar un volumen de 10 µL. La mezcla de ligación se agregó directamente a los productos de la reacción de corte, se incubó a 37 °C por 1 hora, y luego se inactivó la enzima T4 DNA Ligasa, incubando a 65°C por 10 minutos. Las reacciones de amplificación se realizaron en dos pasos: (i) reacción de pre-amplificación usando cebadores de AFLP *EcoRI/MseI* con un nucleótido selectivo en el extremo 3' (+1), y (ii) reacciones de amplificación selectiva usando cebadores alternativos de AFLP *EcoRI/MseI* con tres nucleótidos selectivos en el extremo 3' (+3). Una vez realizado el ligado de los adaptadores, se llevó a cabo la primera reacción de PCR o pre-amplificación en un volumen final de 25 µL conteniendo: 2 µL de la mezcla de corte y ligación, 1X de amortiguador *Taq* (Invitrogen), 0,2 M de cada dNTP (Promega), 1,5 mM de MgCl₂, 1 U de enzima *Taq* polimerasa (Invitrogen), 1,2 ng/µL de cada cebador *EcoRI/MseI* +1 (Tabla 1) y agua ultrapura estéril hasta completar el volumen final. Las condiciones de pre-amplificación fueron las siguientes: una etapa de desnaturalización inicial de 30 segundos a 94 °C seguida de 20 ciclos de 30 segundos a 94 °C, 30 segundos a 60 °C y 1 minuto a 72 °C. Luego se realizaron reacciones de amplificación selectiva, utilizando 15 combinaciones de los cebadores de AFLP *EcoRI/MseI* (+3). El producto de pre-amplificación se diluyó 10 veces en amortiguador TE [10 mM Tris-HCl y 0,1 mM EDTA (pH 8)] y en estas condiciones se utilizó como molde en las amplificaciones subsiguientes. Las reacciones de amplificación selectiva se realizaron en un

volumen final de 10 µL conteniendo: 2 µL del producto de amplificación diluido, 1X de amortiguador *Taq* (Invitrogen), 0,2 M de cada dNTP (Promega), 1,5 mM de MgCl₂, 1 U de enzima *Taq* polimerasa (Promega), 1,2 ng/µL de cada cebador *EcoRI/MseI* +3 (Tabla 2) y agua ultrapura hasta completar el volumen final. Las condiciones de amplificación fueron las mismas que las descritas en Vos *et al.* (1995): un primer ciclo de 30 segundos a 94 °C, 30 segundos a 65 °C y 1 minuto a 72 °C, seguido de 12 ciclos donde la temperatura de hibridación (65 °C) se disminuyó en 0,7 °C por ciclo, y finalmente 23 ciclos de 30 segundos a 94 °C, 30 segundos a 56 °C y 1 minutos a 72 °C. Los productos de amplificación fueron suplementados con 5 µL de solución tampón de siembra (formamida 95% p/v, NaOH 10 mM y azul de bromofenol 2 mg/mL), denaturalizados a 96 °C por 5 min, y resueltos en geles de poliacrilamida, en una cuba vertical Sequigen GT (BioRad). La matriz del gel se preparó usando 4,75 % de acrilamida (BioRad), 0,25 % de metilen-bisacrilamida (BioRad) y 7,5 M de urea (Invitrogen) en presencia de 0,5X de amortiguador TBE [50 mM de Tris, 50 mM de ácido bórico y 1 mM de EDTA (pH 8)]. A 80 mL de esta solución se agregaron 360 µL de persulfato de amonio al 10 % (APS, Sigma Aldrich) y 36 µL de N,N,N₂,N₁-tetrametil-etilendiamino (TEMED, Sigma Aldrich) para permitir la polimerización. Se realizó una recorrida de la matriz a una potencia constante de 60 W por aproximadamente 1 hora, hasta que el gel alcanzó una temperatura homogénea de 50 °C. Se utilizó amortiguador 0,5X TBE en el reservorio superior de la cuba y 1X TBE en el reservorio inferior. Se sembraron 5 µL de cada muestra en los correspondientes pocillos y se realizó la corrida electroforética a una potencia constante de 60 W por 1 hora 30 minutos. Los marcadores se revelaron utilizando el procedimiento de tinción con plata descrito anteriormente. Los marcadores moleculares fueron identificados de acuerdo a la combinación de cebadores de AFLP utilizada, empleando la denominación alfa-numérica establecida por los proveedores seguido por un número en orden de mayor a menor peso molecular en forma sucesiva (Olsina, 2012).

Tabla 3: Marcadores de SSR seleccionados a partir de la base de datos Soybase (soybase.org)

Locus	Grupo de Ligamiento	Posición cM	Nº de acceso GenBank	Cebador directo 5'-3'	Cebador reverso 5'-3'
Satt 511	A1	94,2	BH126682	GCGACTTTACTGAAAACCTGGAAA	GCTTCAAACCAACAAACAACTTA
Satt 449	A1	27,784	CC453967	GCGTGCTTCTTATATTAGGTGTTAGT	GCGCATTGGAGTTTTTGCTTTT
Satt 042	A1	27,664	BH126311	GACTTAATTGCTTGCTATGA	GTGGTGCACACTCACTT
Satt 409	A2	145,565	BH126589	CCTTAGACCATGAATGTCTCGAAGATA	CTTAAGGACACGTGGAAGATGACTAC
Satt 177	A2	36,77	BH126385	CGTTTCATCCCATGCCAATA	CCCGCATCTTTTCAACCAC
Satt 233	A2	100,086	BH126434	AAGCATACTCGTCGTAAC	GCGGTGCAAAGATATTAGAAA
Satt 453	B1	123,962	BH126627	GCGGAAAAAAAAACAATAACAACA	TAGTGGGGAAGGGAAGTTACC
Satt 197	B1	46,393	BH126404	CAGTCTTTTCCCTCTCT	AAGATACCCCAACATTATTGTAA
Satt 556	B2	73,213	BH126723	GCGATAAACCCGATAAATAA	GCGTTGTGCACCTTGTTTTCT
Satt 534	B2	87,592	BH126703	CTCCTCCTGCGCAACAACAATA	GGGGGATCTAGCCATGAC
Satt 577	B2	6,047	BH126743	CAAGCTTAAGTCTTGGTCTTCTCT	GGCCTGACCCAAAATAAGGGAAGTG
Satt 070	B2	72,808	BH126318	TAAAAATAAAATACTAGAAGACAAC	TGGCATTAGAAAATGATATG
Satt 294	C1	78,645	BH126487	GCGGGTCAAATGCAAATATTTTT	GCGCTCAGTGTGAAAGTTGTTTCTAT
Satt 180	C1	127,774	BH126388	TCGCGTTTGTGACG	TTGATTGAAACCCAATA
Satt 307	C2	121,271	BH126498	GCGCTGGCCTTTAGAAC	GCGTTGTAGGAAATTTGAGTAGTAAG
Satt 281	C2	40,298	BH126476	AAGCTCCACATGCAGTTCAAAAC	TGCATGGCACGAGAAAGAAGTA
Satt 436	D1a	70,691	BH126614	GCGTATAAAGAAAACGAGCATATCAT	GCGCTATAAAGGCTTGTAAGGACACT
Satt 147	D1a	108,89	BH126359	CCATCCCTTCTCCAAATAGAT	CTTCCACCCCTAGTTGTAGTGACAA
Satt 141	D1b	72,887	BH126353	CGTGGTGGTGTGCATAATAA	CCGTCATAAAAAGTCCCTCAGAAT
Satt 005	D1b	75,292	BH146210	TATCCTAGAGAAGAACTAAAAAA	GTCGATTAGGCTTGAATAA
Satt 703	D1b	98,745	CC454052	GCGATGGGTGGATTCTAGAT	GCGGTGCTGTAGGCTTTCCAAAAT
Satt 095	D1b	25,6	BH126329	TATTTGTTATTGGTGAATTAAGA	ATTCTGTAGAAATGTGATCTG
Satt 458	D2	24,517	BH126631	TTGGGTTGACCGTGAGAGGGAGAA	GCGAACCACAAACAACAATCTTCA
Satt 226	D2	85,148	BH126427	GCGAAACAACCTCACTTAAGCAATACAT	GCGTCTCCTACCTTTCTTATC
Satt 311	D2	84,62	BH126502	GCGTCTCCTACCTTTCTTATC	GTTGAAGCTCAGGCTGTGATGAAT
Satt 231	E	70,227	BH126432	GCGTGTGCAAAATGTTTCATCATCT	GGCACGAATCAACATCAAAACTTC
Satt 146	F	1,918	BH126358	AAGGGATCCCTCAACTGACTG	GTGGTGGTGGTGAACACTATTAGAA
Satt 030	F	3,952	BH126307	AAAAAGTGAACCAAGCC	TCTTAAATCTTATGTTGATGC
Satt 114	F	63,689	BH126332	GGGTATCTCCCAATA	ATATGGGATGATAAGGTGAAA
Satt 324	G	33,263	BH126515	GTTCCAGGTCCCACCATCTATG	GCGTTTCTTTTATACCTTCAAG
Satt 191	G	96,572	BH126398	CGGATCATGTCTCTG	GGGAGTTGGTGTTCCTTGTG
Satt 038	G	1,84	CC453951	GGGAATCTTTTTTCTTCTATTAAGTT	GGGCATTGAAATGGTTTATGTA
Satt 181	H	91,116	BH126389	TGGTAGCAGATTGACA	GGAGCATAGCTGTTAGGA
Satt 253	H	67,168	BH126451	GCGCCTAATAAAGATAAGACAAG	GCGTGGCTTTTCCCAATTA
Satt 292	I	82,777	BH126485	GCGGAATTAGAAGCTCCAGTAAAGA	GCGAGGCCAACATTGAAAAGT
Satt 354	I	46,22	BH126543	GCGAAATGGACACCAAAAGTAGTTA	GCGATGCACATCAATTAGAATATACAA
Satt 414	J	37,042	CC453965	GCGTATTCTAGTCACATGCTATTCA	GCGTCATAATAATGCCTAGAACATAAA
Satt 596	J	39,643	CC453972	TCCCTTCGTCCACCAAAT	CCGTCGATTCCGTACAA
Satt 242	K	14,354	BH126443	GCGTTGATCAGGTCGATTTTTATTGT	GCGAGTGCCAACCTACTCTTTATGA
Satt 417	K	46,204	BH126596	TCTTGCTAATTGCTTCAATTCAT	AATTGCTTGGGATTTTCATTT
Satt 475	K	78,683	BH126647	AAATGCAACTTACACAACCTTAT	TGCTTGCTTCAATTTGGATGAGTG
Satt 373	L	107,24	BH126560	TCCGCGAGATAAATTCGTAATAAAT	GGCCAGATACCAAGTTGACTTGT
Satt 156	L	56,142	BH126368	CGCACCCCTCATCTATGTA	CCAATAATCCAGGGACTTACTT
Satt 308	M	130,756	BH126499	GCGTTAAGGTTGGCAGGGTGAAGTG	GCGCAGCTTTATACAAAATCAACAA
Satt 175	M	61,932	BH126384	GACCTCGCTCTGTTTCTCAT	GGTGACCACCCCTATTCTTAT
Satt 150	M	18,582	BH126362	AAGCTGAGGTTATTCGAAAATGAC	TGCCATCAGGTTGTGTAAGTGT
Satt 237	N	74,987	BH126438	GCGTGAATTTCAATCCTTTTTTTC	GCGGTTGCTCCTGTTAGAACCT
Satt 009	N	28,52	BH146212	CCAAGTTGAAATTAAGTAGAGAAA	CTTACTAGCGTATTAACCCCTT
Satt 173	O	58,398	BH126382	TGCGCAATTTATTCTTCA	AAGCGAAATCACCTCTCTCT
Satt 94	O	56,579	BH126328	CCAAGTGCCAAATGAAG	ATCCATGGTTTTTTGATG
Satt 243	O	119,504	BH126444	GCGCATTGCACATTAGGTTTTCTGTT	GCGGTAAGATCACGCCATTATTAAGA

Análisis de perfiles genéticos y establecimiento de matrices binarias

La lectura de cada gel se llevó a cabo en forma manual de manera independiente por dos operadores. Las bandas intensas, claramente diferenciables de las espurias, y confirmadas por ambos operadores, fueron utilizadas en el análisis. Las bandas que mostraron igual movilidad electroforética fueron consideradas como idénticas, y se asumió que representaban el mismo alelo. El tamaño en pares de bases (pb) de los alelos de cada uno de los marcadores SSR fue estimado mediante la aplicación de una curva de logaritmo del peso molecular vs distancia de migración desde el pocillo hasta el centro de la banda, utilizando como referencia un marcador de peso molecular de 10 pb de separación, y procesando los datos con el sistema Excel®. El cálculo del tamaño de los alelos SSR se realizó por duplicado. Las bandas de ADN obtenidas del análisis de los marcadores de SSR y AFLP fueron registradas como 1 (presencia) o 0 (ausencia) y los resultados fueron almacenados en una hoja de cálculo de Excel® en forma de matriz numérica binaria.

Análisis Estadísticos

La comparación de la eficiencia discriminatoria de cada marcador se realizó en base a los siguientes parámetros calculados por unidad de ensayo (U), según Powell et al. (1996): 1) Número de bandas polimórficas (np); 2) Número de bandas monomórficas (nnp); 3) Promedio de bandas polimórficas por unidad de ensayo (np/U); 4) Número de loci (L); 5) Número de loci por unidad de ensayo: $nu = L/U$; 6) Índice de Diversidad Genética (ID) de los loci polimórficos para un marcador genético: $ID = 1 - \sum p_i^2$, donde p_i es la frecuencia alélica para el i -ésimo alelo (Botstein et al, 1980); 7) Media aritmética del Índice de Diversidad Genética de los loci polimórficos: $ICp = \sum IC/np$; 8) Fracción de loci polimórficos: $\beta = np/(np + nnp)$; 9) Relación múltiple efectiva: $E = nu \times \beta$; 10) Índice de marcador: $MI = E \times Hep$. Para el cálculo de las matrices de similitud genética se utilizaron los coeficientes recomendados por Kosman & Leonard (2005), de acuerdo a la ploidía del organismo en estudio, tipo de marcador molecular y tipo de dominancia. En el caso de los SSR, para calcular la similitud entre pares de cultivares i y j se utilizó el coeficiente de similitud de Jaccard (1908), mientras que en el caso de marcadores AFLP se empleó el coeficiente de emparejamiento simple descripto en Sneath &

Sokal (1973). Con los datos se realizaron análisis de agrupamientos utilizando el método de apareamiento de grupos no ponderado mediante media aritmética (UPGMA) y se representaron como dendrogramas. Para determinar con cuánta exactitud los agrupamientos representaban las estimaciones de similitud genética entre los genotipos, se generó una matriz cofenética a partir de cada uno de los agrupamientos, la cual se comparó con la matriz de similitud correspondiente por la prueba de correspondencia de matrices de Mantel (Mantel, 1967). La misma estadística de Mantel fue utilizada para comparar las matrices de similitud, así como las matrices cofenéticas producidas por las técnicas de AFLP y SSR. Adicionalmente, para ambos marcadores moleculares se llevó a cabo un análisis de ordenamiento mediante la técnica de Análisis de Coordenadas Principales (Gower, 1966) y se construyó un árbol de mínima distancia no sesgada (Nei, 1978). Todos estos procedimientos se realizaron por rutinas adecuadas en el programa NTSYSpc versión 2.17 (Exeter Software, Setauket, NY). Para comparar la media y varianza de un subconjunto seleccionado de 20 marcadores SSR en relación al conjunto de 51 marcadores iniciales, se realizó un test-t de comparación de medias y un test F para la comparación de varianzas. Además se realizaron histogramas para analizar las distribuciones de frecuencias de los coeficientes de similitud genética calculados para el subgrupo de 20 loci SSR versus el conjunto de 51. Todos estos cálculos, gráficos y test de hipótesis fueron realizados utilizando Excel® (Olsina, 2012).

Identificación de muestras incógnitas

Para cada comparación se eligió un número particular de marcadores SSR procesables en un único gel, y que además presentasen alelos que permitiesen la diferenciación de los cultivares de referencia indicados como identidad putativa. Cada muestra fue comparada contra la base de datos para determinar la capacidad de los marcadores seleccionados para diferenciar correctamente los cultivares evaluados. Para ello se confeccionó una matriz de datos con los 86 cultivares previamente caracterizados, teniendo en cuenta los marcadores efectivamente evaluados en cada muestra y agregando la/s muestras reales analizadas. A cada una de estas matrices se le calculó la similitud genética y se obtuvo el correspondiente dendrograma.

Resultados

Caracterización genética con marcadores SSR

La caracterización de 86 cultivares seleccionados de soja fue realizada mediante la selección de 51 marcadores SSR descriptos para esta especie. Se seleccionaron entre 2 y 5 marcadores por cada uno de los grupos de ligamiento, con el objeto de lograr una buena cobertura genómica. Los marcadores mostraron perfiles electroforéticos claros y con buena resolución, amplificando casi siempre una sola banda por locus SSR. Esta observación está en total acuerdo con la estructura homocigota esperada para la mayoría de los loci en este cultivo (Diwan y Cregan, 1997), aunque algunos resultaron heterocigotas. Para lograr la caracterización de los cultivares analizados en este estudio con los 51 loci SSR, se realizaron más de 100 geles de poliacrilamida. La Figura 1, muestra un ejemplo correspondiente a una de esas corridas electroforéticas, donde se observa el perfil polimórfico de las bandas obtenidas.

En la Tabla 4 se muestran los resultados de la caracterización de los cultivares mediante marcadores SSR. Se obtuvieron un total de 242 bandas polimórficas, con un promedio de 5 bandas por locus, variando de 2 a 9 bandas. La frecuencia de ocurrencia observada de los 242 alelos varió de 0,01 a 0,95 con un promedio

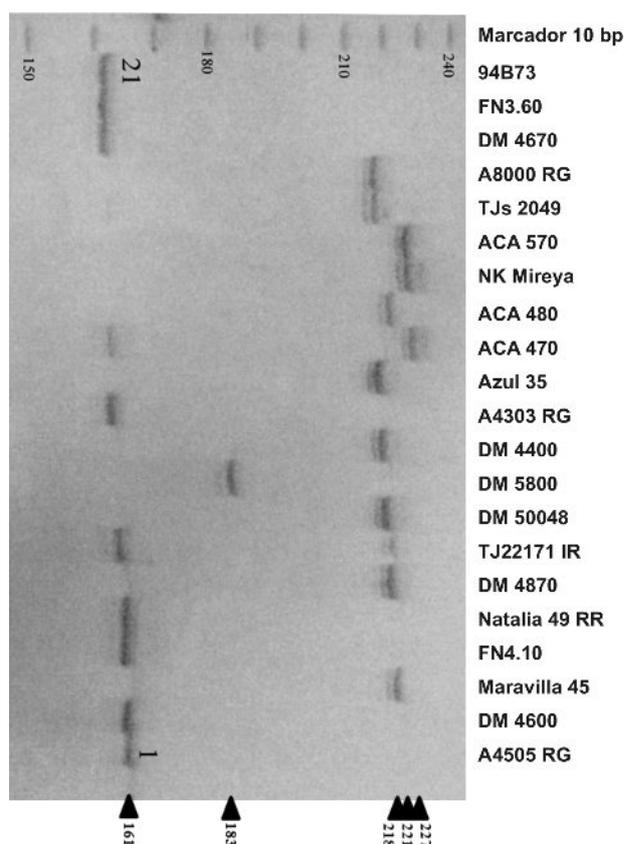
de 0,22. Cuatro alelos mostraron una frecuencia de ocurrencia de 0,8 o mayor en los cultivares analizados y 27 alelos una frecuencia mayor a 0,5, mientras que 32 alelos mostraron frecuencias menores a 0,02. Estos alelos que muestran una muy baja frecuencia pueden ser útiles como marcadores diagnósticos para la identificación inequívoca de algunos de los cultivares de soja analizados, ya que poseen un alto poder de discriminación. Los valores de índice de diversidad genética (ID), que indican la probabilidad de que al tomar dos genotipos al azar a partir de una población de individuos ensayados estos sean polimórficos, variaron de 0,081 a 0,753 con un promedio de 0,523. Los marcadores SSR Satt 141 y 373 mostraron el ID más alto, con valores de 0,753 y 0,730 respectivamente, reflejando el alto número de alelos observados y una distribución de frecuencias alélicas uniformes (Figura 2). Por otro lado, los marcadores Satt 150 y Satt 475 presentaron los ID más bajos con valores de 0,111 y 0,081, respectivamente, mostrando un bajo número de alelos y frecuencias alélicas poco uniformes (Figura 2). Como es de esperarse, se observa que la variación del ID está significativamente asociada con el número de alelos detectados por locus (Figura 3).

Tabla 4: Características de los 51 marcadores SSR utilizados sobre los 86 cultivares argentinos de soja.

Locus	Grupo de Ligamiento	Rango de tamaños de alelos (pb)	Nº de alelos	Ida
Satt 141	D1b	148 - 204	6	0.753
Satt 373	L	212 - 283	9	0.730
Satt 458	D2	158 - 194	9	0.727
Satt 146	F	278 - 317	7	0.715
Satt 005	D1b	142 - 178	6	0.705
Satt 181	H	176 - 217	5	0.702
Satt 030	F	152 - 167	5	0.695
Satt 307	C2	162 - 192	5	0.687
Satt 414	J	168 - 315	6	0.687
Satt 453	B1	221 - 269	8	0.684
Satt 556	B2	164 - 211	4	0.658
Satt 114	F	82 - 122	4	0.654
Satt 242	K	187 - 199	4	0.640
Satt 703	D1b	189 - 239	5	0.638
Satt 009	N	161 - 227	5	0.624
Satt 534	B2	232 - 261	7	0.606
Satt 577	B2	113 - 122	4	0.591
Satt 292	I	223 - 254	6	0.591
Satt 596	J	236 - 259	6	0.589
Satt 237	N	237 - 281	8	0.579
Satt 324	G	229 - 241	3	0.577
Satt 417	K	280 - 321	3	0.576
Satt 354	I	233 - 250	4	0.573
Satt 197	B1	175 - 189	5	0.569
Satt 038	G	157 - 177	5	0.567
Satt 308	M	149 - 175	4	0.566
Satt 226	D2	321 - 338	3	0.548
Satt 231	E	215 - 242	5	0.546
Satt 191	G	185 - 229	6	0.545
Satt 436	D1a	182 - 246	7	0.544
Satt 095	D1b	160 - 175	4	0.510
Satt 511	A1	244 - 256	3	0.499
Satt 070	B2	147 - 174	4	0.495
Satt 156	L	205 - 220	3	0.483
Satt 173	O	199 - 263	4	0.478
Satt 180	C1	244 - 268	3	0.469
Satt 294	C1	248 - 293	5	0.465
Satt 409	A2	165 - 195	6	0.441
Satt 094	O	148 - 163	2	0.428
Satt 243	O	202 - 211	2	0.426
Satt 449	A1	222 - 266	5	0.412
Satt 177	A2	113 - 122	2	0.400
Satt 311	D2	180 - 191	2	0.397
Satt 147	D1a	171 - 203	4	0.387
Satt 042	A1	159 - 174	4	0.387
Satt 175	M	160 - 193	6	0.289
Satt 281	C2	167 - 236	4	0.263
Satt 253	H	138 - 162	6	0.259
Satt 233	A2	188 - 209	3	0.144
Satt 150	M	189 - 200	2	0.111
Satt 475	K	231 - 252	3	0.081

a Índice del diversidad génica: probabilidad de que al tomar dos genotipos al azar a partir de una población de individuos ensayados estos sean polimórficos en ese marcador.

Figura 1. Alelos detectados para el locus SSR Satt009 en 21 de los 86 cultivares de soja analizados en el presente estudio. A la izquierda se sembró una calle con un marcador de peso molecular de ADN de 10 pares de base de separación. Al tope de la figura se indican los nombres de los cultivares analizados. Las bandas se observan como bandas dobles porque la tinción con plata permite detectar la hebra complementaria, que migra a una distancia ligeramente diferente por poseer una secuencia y composición de bases diferente. Las flechas de la derecha indican los tamaños de los alelos en pares de bases (pb). La posición de las muestras 1 y 21 se indica sobre el gel a la altura del alelo de 161 pb.



Caracterización genética con marcadores AFLP

Los cebadores de AFLP se seleccionaron a partir de estudios realizados por otros grupos, ya que habían revelado tasas altas de polimorfismo (Ude *et al.* 2003; Bonato *et al.* 2006). Se utilizaron 15 combinaciones de cebadores, para generar 636 bandas, de las cuales 336 (52,8%) fueron polimórficas. En la Figura 4 se muestra un ejemplo representativo de un gel de AFLP. Se obtuvieron entre 20 y 63 bandas totales, con una media de 41 bandas, por combinación de cebadores. El número promedio de bandas polimórficas por combinación de cebadores fue de 17, con un mínimo de 5 y un máximo de 29 bandas.

En la Tabla 5 se resumen los resultados obtenidos con los AFLPs. Las combinaciones *EcoRI*-AAG/*MseI*-CTA y *EcoRI*-ACT/*MseI*-CAT generaron los niveles más altos de bandas polimórficas, 29 (56%) y 27 (52%), respectivamente. Los valores del índice de diversidad génica variaron de 0,0604 a 0,3650 con un promedio de 0,1924. Se observó una asociación positiva entre el número de bandas polimórficas y el índice de diversidad génica calculado para cada combinación de cebadores.

Figura 2. Distribución de frecuencia alélica de marcadores SSR Satt 373; Satt 141; Satt 475 y Satt 150 en los cultivares de soja utilizados.

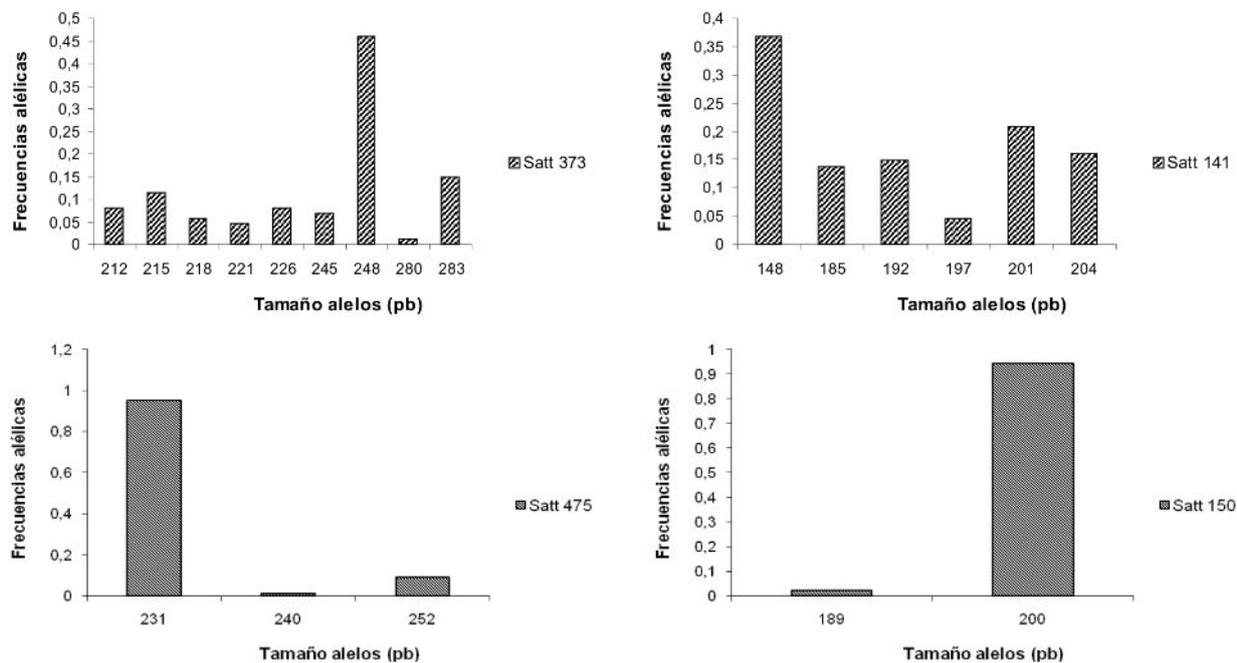
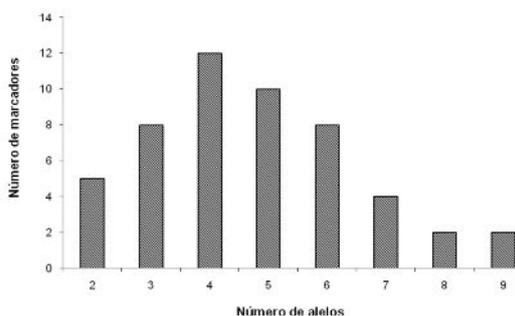
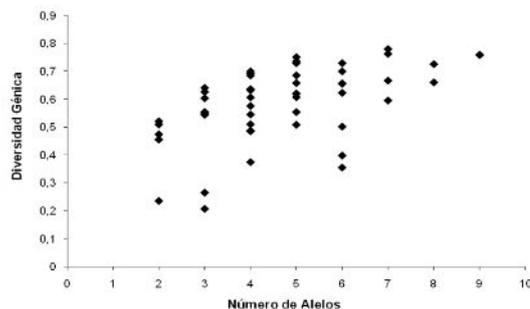


Figura 3. A) Frecuencia de los marcadores SSR en función del número de alelos detectados. B) Diversidad genética en función del número de alelos detectados, calculado en los 86 cultivares de soja utilizados en este estudio. Se observa que a medida que aumenta el número de alelos, los marcadores SSR son más informativos.

A



B



Comparación de los resultados obtenidos con los marcadores SSR y AFLP

En la Tabla 6 se muestra una comparación del comportamiento de los marcadores moleculares AFLP y SSR. Puede observarse que los AFLP produjeron una mayor cantidad de datos por experimento (unidad de ensayo), ya que estos marcadores permiten detectar la presencia de múltiples *loci* por corrida. El número promedio de bandas polimórficas para los marcadores AFLP es de 22,4 mientras que para SSR es 4,75. Sin embargo, los marcadores SSR poseen una fracción mayor de *loci* polimórficos (1 contra 0,53) y un Índice de Diversidad promedio 3 veces más alto que los AFLP (0,19 AFLP vs. 0,59 SSR). Esto se debe a que muchos de los marcadores generados por AFLP están dirigidos a regiones muy conservadas del genoma, y resultan monomórficos entre las líneas. El alto valor observado del Índice del marcador para AFLP (7,19) en comparación con los SSR (0,59), es el resultado del alto componente de relación de multiplex del primero por unidad de ensayo.

Estimación de la similitud genética en cultivares de soja mediante marcadores AFLP y SSR

Se construyó una matriz de similitud genética a partir de los

cálculos obtenidos para cada par de los 86 cultivares comparados mediante ambos marcadores. La Figura 5, muestra la distribución de frecuencias de los coeficientes de similitud obtenidos para cada marcador molecular. Se pudo apreciar un mayor poder de discriminación de los marcadores SSR con respecto a los AFLP. El valor promedio de similitud fue de 0,292 para los marcadores SSR, variando desde 0,084 entre los cultivares ACA 530 y Nueva María 55 a 0,925 en los cultivares DM5.8i y DM5.5i (Tabla 7). Estos dos últimos cultivares difirieron en sólo dos alelos revelados por marcadores SSR (Satt 436 y Satt 156). Los valores relativamente altos de similitud genética estimados mediante marcadores SSR son indicativos de la uniformidad existente entre los cultivares de soja comerciales. No obstante, utilizando este conjunto de marcadores SSR se pudo diferenciar adecuadamente los 86 materiales evaluados. Para el caso de marcadores AFLP, el promedio de similitud fue de 0,921, cuando se consideró para el cálculo el total de las bandas obtenidas (polimórficas y monomórficas) (Tabla 7). Utilizando sólo las bandas polimórficas para el cálculo del coeficiente de similitud genética el valor promedio descendió a 0,848 (Tabla 7). Los cultivares A6019RG y DM3100 presentaron los valores de similitud más bajos (0,751), mientras

Tabla 5. Resultados del análisis de AFLP obtenidos a partir de combinaciones de cebadores EcoRI y MseI.

Cebadores		Número de bandas		ID ^a
Eco RI	Mse I	Totales	Polimórficas	
E-AAC	M-CAT	48	19	0.2251
E-AAC	M-CTA	49	20	0.2497
E-AAC	M-CTC	40	19	0.2227
E-AAC	M-CTG	34	13	0.0813
E-AAC	M-CTT	39	24	0.3650
E-AAG	M-CAT	63	18	0.0604
E-AAG	M-CTA	58	29	0.2039
E-AAG	M-CTC	34	13	0.2012
E-AAG	M-CTG	23	8	0.1431
E-AAG	M-CTT	61	13	0.0888
E-ACT	M-CAT	57	27	0.1812
E-ACT	M-CTA	41	11	0.2025
E-ACT	M-CTC	20	5	0.0647
E-ACT	M-CTG	35	24	0.3147
E-ACT	M-CTT	36	16	0.2816

^a ID: Índice de diversidad génica

Tabla 6: Comparación de los niveles de polimorfismo y la capacidad discriminante de los marcadores AFLP y SSR.

Índices y sus abreviaturas	Marcador		
	AFLP	SSR	
Número de unidades de ensayo	U	15	51
Número de bandas polimórficas	n _p	336	242
Número de bandas monomórficas	n _{mp}	300	0
Promedio de bandas polimórficas por unidad de ensayo	n _p /U	22,4	4,75
Numero de loci	L	636	51
Numero de loci por unidad de ensayo	n _u	42,4	1
Índice de Diversidad promedio (ID) de los loci polimórficos	ID _p	0,19	0,59
Fracción de loci polimórficos	β	0,53	1
Relación multiplex efectiva	E	22,4	1,00
Índice del Marcador	MI	7,19	0,59

Tabla 7: Análisis de las propiedades del coeficiente de similitud genética según marcador estudiado.

Coef. de similitud/ Marcador	SSR	AFLP Totales	AFLP Polimórficas	SSR y AFLP
Promedio	0,292	0,921	0,848	0,687
Min	0,0842	0,751	0,498	0,365
Max	0,925	0,991	0,982	0,949
Varianza	0,0073	0,0007	0,00306	0,0034

que TJS 2136 y TJS 2138 mostraron los más altos (0,991), difiriendo en sólo 5 bandas polimórficas. El 87% de los coeficientes estimados fueron mayores a 0,90, lo que refleja el alto grado de similitud genética existente entre los cultivares utilizados en este estudio. Sin embargo, el 1,4% de los coeficientes fueron menores o iguales a 0,80.

Mediante un test de correlación de matrices (Mantel, 1967) se determinó la correspondencia entre las matrices de similitud genética de los marcadores SSR y AFLP. Los resultados indicaron que hay una correlación muy baja entre la información genética obtenida con ambos marcadores para los cultivares

Figura 4. Ejemplo de perfilado obtenido con la combinación de cebadores AFLP, que muestra los alelos detectados en 27 de los 86 cultivares de soja analizados en el presente estudio. En paralelo se sembró una calle con un marcador de peso molecular de ADN de 10 pares de base de separación.

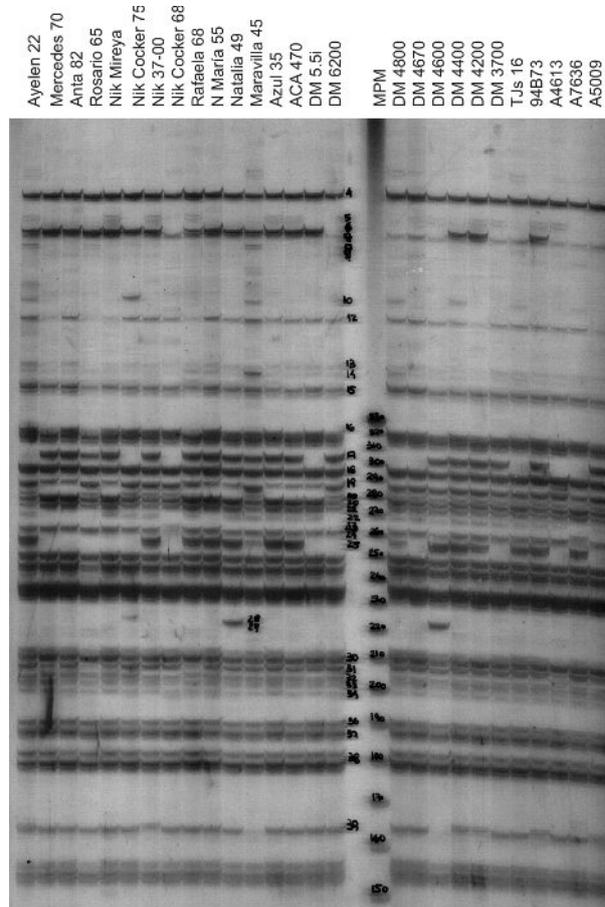
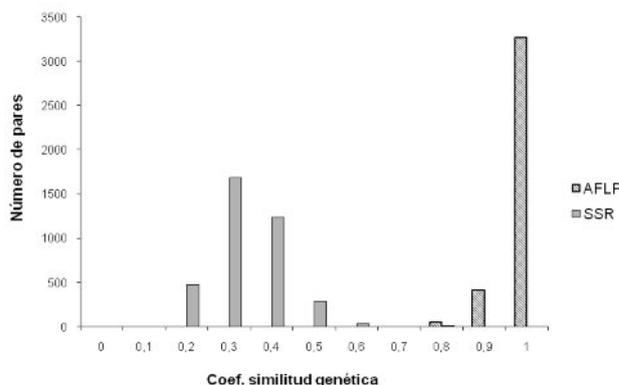
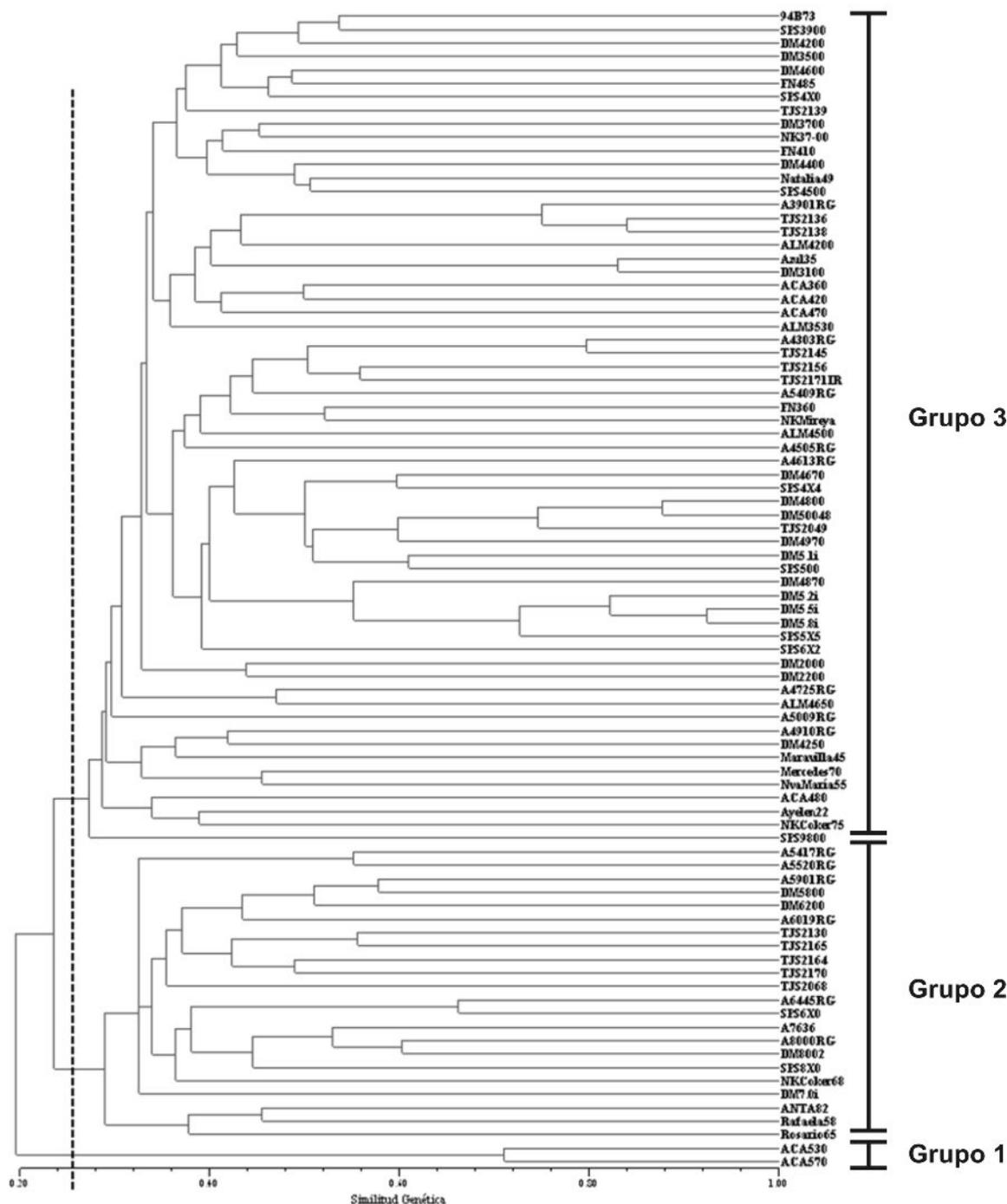


Figura 5. Distribución de frecuencias de las estimaciones de coeficientes de similitud genética para los 86 cultivares de soja con los marcadores SSR y AFLP.



evaluados ($r = 0,166$; Probabilidad Z al azar < Z observado: $p = 0,9997$). Este resultado está de acuerdo con la diferente naturaleza de las secuencias genómicas blanco de cada marcador: mientras los SSR están dirigidos específicamente a las regiones microsatélites, los AFLP cubren el genoma al azar, abarcando secuencias tanto muy conservadas, como medianamente polimórficas o muy polimórficas.

Figura 6. Dendrograma obtenido utilizando marcadores SSR en los cultivares de soja seleccionados en este estudio. La línea trazada a una similitud genética de 0,24 permite diferenciar tres grupos principales.



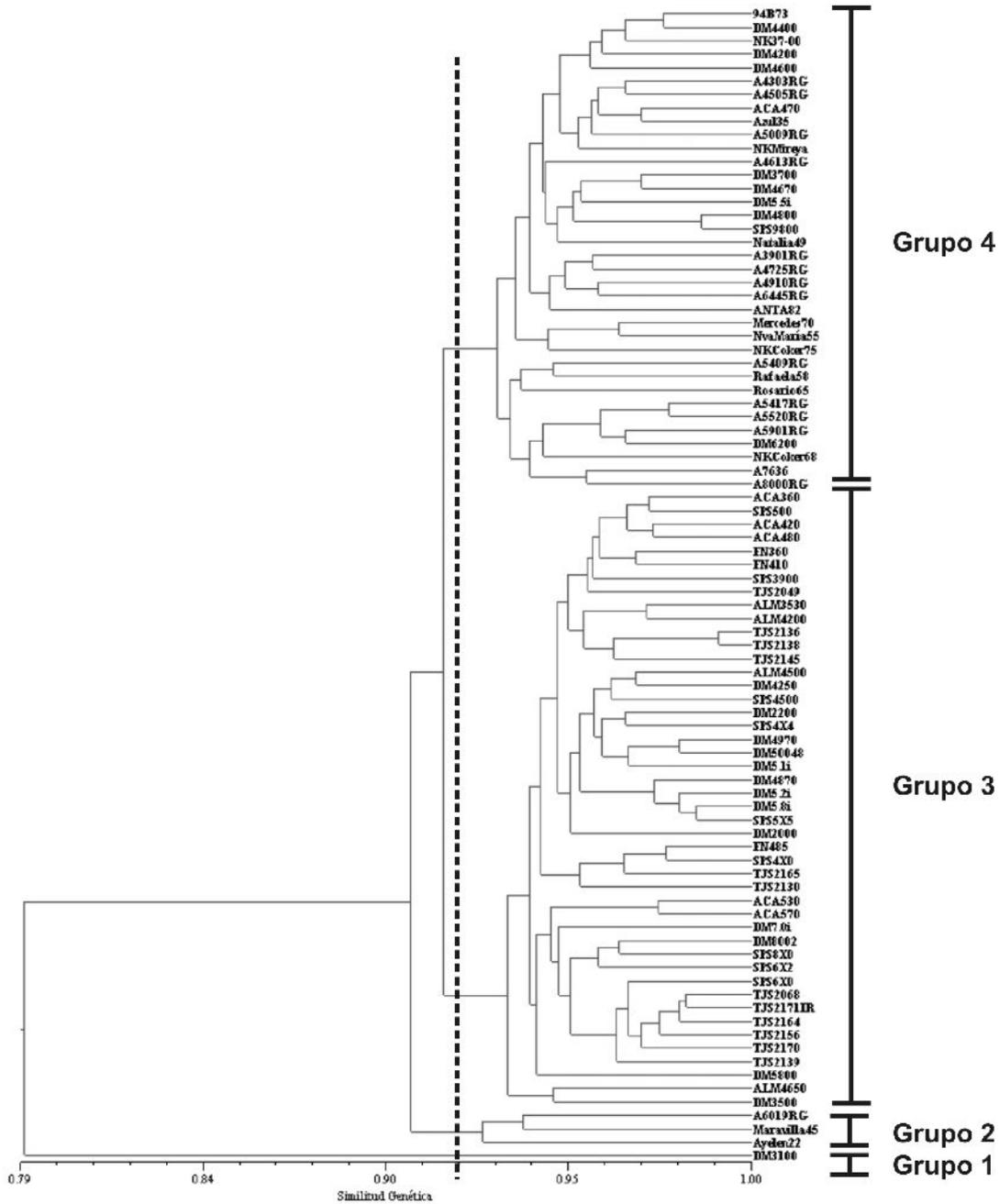
Análisis de agrupamiento

Con las matrices de similitud genética obtenidas, se realizó un análisis de agrupamiento de pares no ponderado utilizando la media aritmética (UPGMA). Este procedimiento reveló que la similitud genética promedio entre genotipos era dependiente del tipo de marcador molecular empleado. En las Figuras 6 y 7, se muestran los resultados obtenidos para cada tipo de marcador molecular (SSR o AFLP).

Los dendrogramas permitieron la diferenciación de los cultivares en estudio, sin embargo las asociaciones observadas en cada uno son características del sistema de marcadores

empleados. En el de SSRs (Figura 6) se pueden distinguir tres grupos principales a una similitud genética de 0,26. El primero de ellos (Grupo 1) está constituido por los cultivares ACA 530 y ACA 570, que se muestran muy diferentes al resto de los materiales en estudio y similares entre sí (coeficiente ~ 0,70). El coeficiente de similitud genético de estos cultivares respecto al resto es de 0,191. El resto de los genotipos se distribuyen en los dos grupos restantes (Grupos 2 y 3). No se observa una asociación entre los agrupamientos formados y los grupos de madurez de los cultivares evaluados, ni tampoco con la identidad del obtentor de las variedades. Los cultivares DM5.5i y DM5.8i, pertenecientes al grupo 3, son los más semejantes entre sí,

Figura 7. Dendrograma obtenido utilizando marcadores AFLP en los cultivares de soja seleccionados en este estudio. La línea trazada a una similitud genética de 0,92 permite diferenciar cuatro grupos principales.



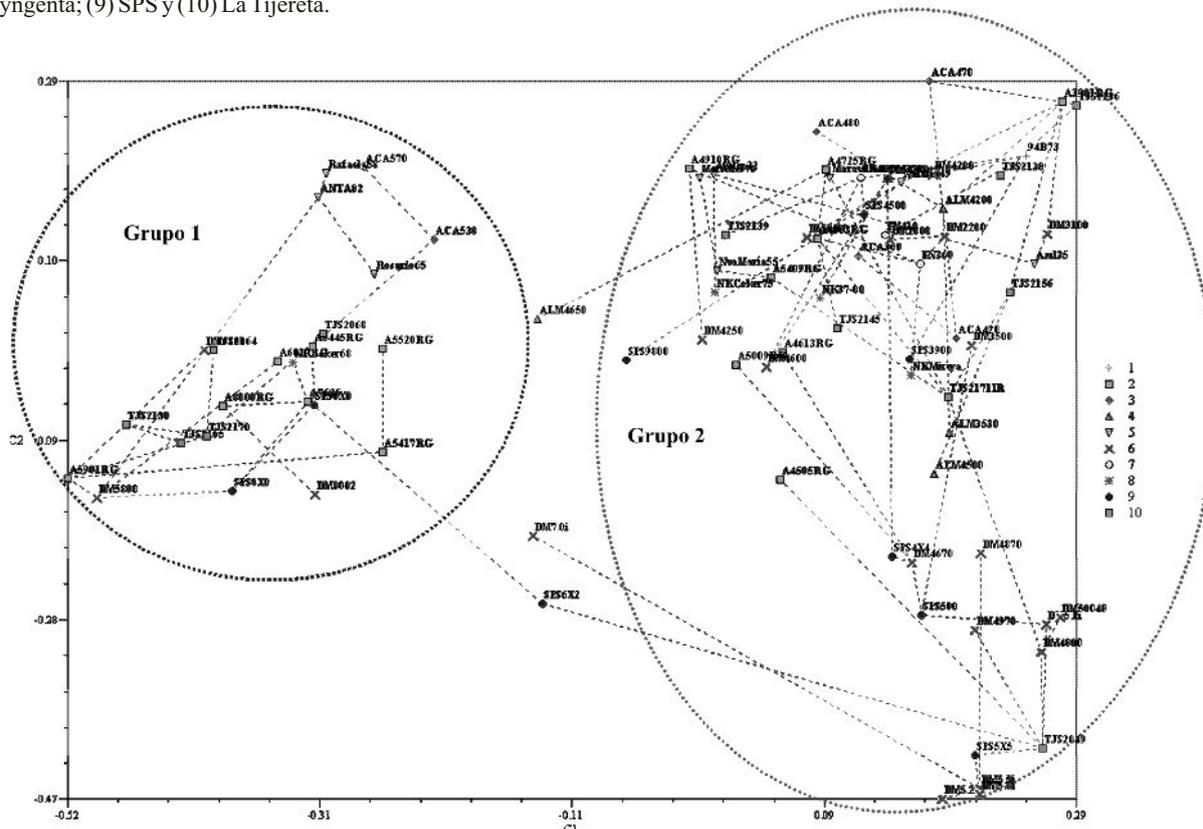
mostrando un coeficiente de similitud genética de 0,925. En el dendrograma de AFLP (Figura 7) se pueden describir cuatro grupos principales, a un valor de similitud genética de 0,92. El cultivar DM 3100 (Don Mario) constituye por si solo uno de los grupos (Grupo 1), presentando un perfil de AFLP completamente diferente a los restantes cultivares en estudio para todas las combinaciones de cebadores utilizadas, resultado que se ve expresado en la baja similitud promedio de este cultivar con el resto (0,765) y en su consecuente aislamiento en el árbol. El segundo grupo se observan los cultivares A6019RG (Nidera), Maravilla 45 (Relmó) y Ayelén 22 (Relmó). En los dos grupos restantes se observa un buen agrupamiento de los cultivares de acuerdo al obtentor. En el grupo 3 se ubicaron los cultivares de las compañías FN Semillas, A. L. Marchionni, La Tijereta, SPS y ACA (con la excepción de SPS9800 y ACA 470, que se ubican en

el grupo 2). En el grupo 4 se agruparon cultivares de Nidera, Relmó y Syngenta. Finalmente, los cultivares de Don Mario, se distribuyeron en ambos grupos (Grupo 3 y Grupo 4) (ver Discusión).

Análisis de ordenamiento

Para el estudio del ordenamiento se llevó a cabo un análisis de coordenadas principales. Los tres primeros factores del análisis de coordenadas principales a partir de la matriz de similitud de Jaccard obtenida mediante la caracterización por SSR explican el 7,82%; 4,76% y 3,8% de la variación genética respectivamente, sumando un 16,4% de la variación genética total. La Figura 8, muestra el plano definido por las dos primeras componentes. Se observan dos grupos principales, que tienen correspondencia con los agrupamientos formados a partir del

Figura 8. Ordenamiento en dos dimensiones de las 86 cultivares de soja luego de un Análisis de Coordenadas Principales y la construcción de un árbol de mínima distancia (MST) para marcadores SSR. Se muestran las dimensiones 1 y 2. Los símbolos indican los criaderos obtentores: (1) Pioneer; (2) Nidera; (3) ACA; (4) A. L. Marchionni; (5) Relmó; (6) Don Mario; (7) FN Semillas; (8) Syngenta; (9) SPS y (10) La Tijereta.



análisis de UPGMA, excepto por el cultivar DM 7.0i. El análisis mediante un árbol de mínima distancia muestra que hay una buena concordancia entre los agrupamientos propuestos en el espacio bidimensional y las distancias entre los cultivares en el espacio de N-Dimensiones original.

El análisis de coordenadas principales obtenido a partir de la caracterización por AFLP mostró una variación de 17,85%; 6,80% y 5,54% en las tres primeras dimensiones respectivamente, sumando un 30,21% de la variación genética total. En la Figura 9 se observa la proyección de los cultivares analizados en el plano definido por las dos primeras componentes. Los cultivares tienden a agruparse en dos grupos principales siguiendo un gradiente a través del primer eje. Estos agrupamientos tienen una muy buena concordancia con los agrupamientos obtenidos a partir del análisis de UPGMA. El árbol de mínima distancia superpuesto, muestra una buena asociación entre las proyecciones en las dos dimensiones del plano y las distancias obtenidas de la matriz de similitud original. Al igual que con los dendrogramas, se observa un buen agrupamiento de los cultivares de acuerdo al obtentor a través del gradiente formado en la componente 1. En el grupo 1 se encuentran ordenados los cultivares de las compañías FN semillas; A. L. Marchionni, La Tijereta, SPS y ACA (menos SPS9800 y ACA 470 respectivamente, que se ubican en el grupo 2). En el grupo 2 por otra parte se encuentran ordenados los cultivares de Nidera, Relmó y Syngenta. Finalmente, los cultivares de Don Mario, se ordenaron entre ambos grupos con los cultivares DM4670, DM4400, DM3700, DM5.5i, DM4800 y DM6200 en el grupo 2 y los 16 cultivares restantes en el grupo 1.

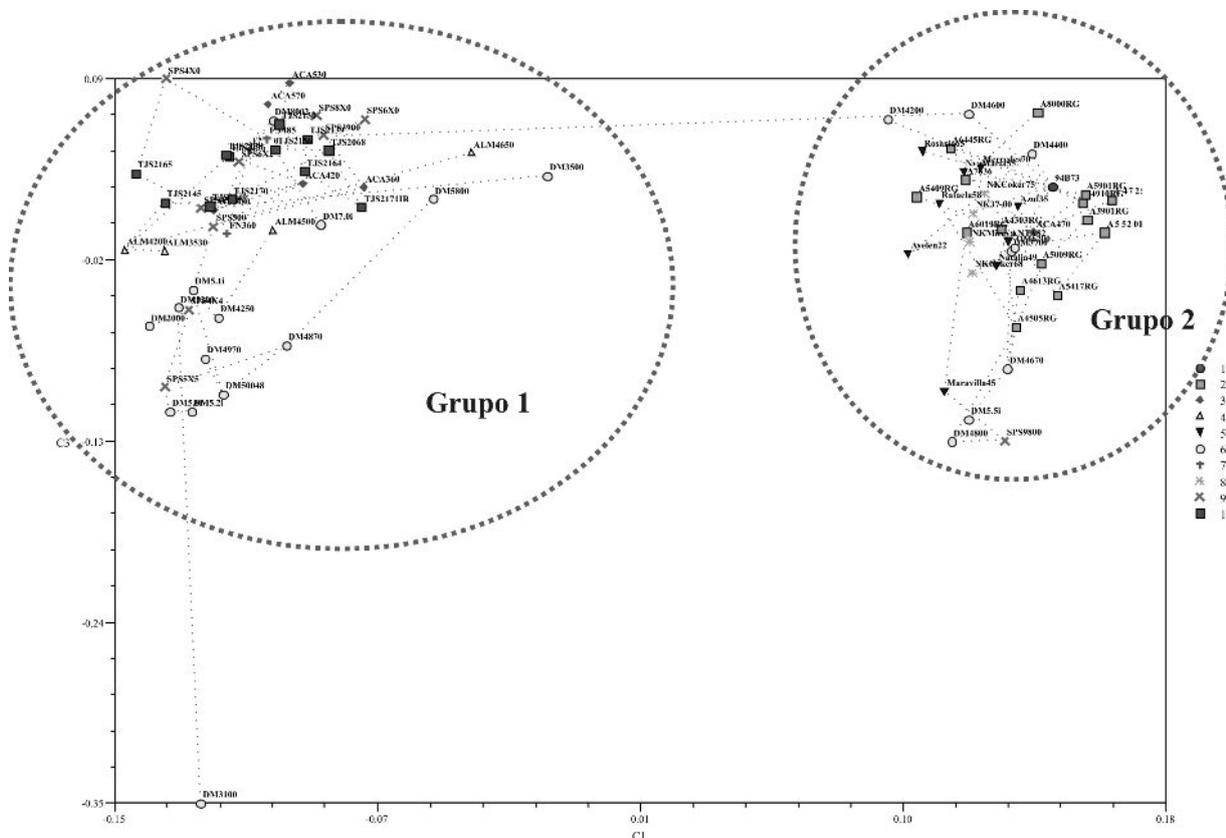
Uso de la base de datos de SSR para la identificación de muestras incógnitas. Prueba de concepto

Con el objeto de poner a prueba la base de datos de SSR de cultivares de soja se analizaron un total de 12 muestras enviadas al laboratorio por diferentes semilleros o multiplicadores, los cuales a su vez designaron una identidad putativa del cultivar que debería ser contrastado con el banco. Los resultados obtenidos se muestran en la Tabla 8. En 10 de los 12 casos los marcadores SSR seleccionados permitieron concluir eficazmente sobre la identidad las muestras ensayadas y diferenciarlas de los restantes cultivares del banco. Esto demuestra el poder discriminatorio de los marcadores SSR como herramienta para la diferenciación de cultivares de soja argentina. En los restantes dos casos no se detectó similitud genética con el cultivar propuesto como identidad putativa, pero además no se encontró similitud con ningún otro cultivar del banco.

Diseño de una metodología eficiente para ensayos de diferenciación de cultivares

La caracterización de muestras incógnitas comentada en la sección anterior se basó en una selección específica de marcadores de acuerdo a la identidad putativa (propuesta) de la muestra, que había sido proporcionada por el solicitante de la identificación. Sin embargo, pensamos que resultaría conveniente contar con un subconjunto selecto y fijo de marcadores que permitiera diferenciar cualquier variedad caracterizada en el banco. Así, nos propusimos seleccionar el mínimo subconjunto de marcadores cuya aplicación, independientemente de la identidad de la muestra, nos permitiera diferenciarla inequívocamente.

Figura 9. Ordenamiento en dos dimensiones de las 86 cultivares de soja luego de un Análisis de Coordenadas Principales y la construcción de un árbol de mínima distancia (MST) para marcadores AFLP. Se muestran las dimensiones 1 y 2. Los símbolos indican los criaderos obtentores: (1) Pioneer; (2) Nidera; (3) ACA; (4) A. L. Marchionni; (5) Relmó; (6) Don Mario; (7) FN Semillas; (8) Syngenta; (9) SPS y (10) La Tijereta.



A partir de la caracterización realizada sobre 86 cultivares de soja se seleccionaron 20 *loci* SSR localizados en los diferentes grupos de ligamiento y que además presentaban los valores de índice de diversidad génica más altos. La utilización de este número reducido de marcadores comunes permitiría el análisis simultáneo de hasta tres muestras cualesquiera en un único gel. Para determinar la efectividad del conjunto de 20 marcadores SSR seleccionados versus el conjunto completo de 51 *loci*, se compararon los dos conjuntos en términos de su habilidad para distinguir los 86 cultivares de soja analizadas en este estudio. La similitud media calculada sobre los 20 *loci* fue numéricamente más baja, que la obtenida utilizando los 51 *loci* SSR (0,214 con 20 *loci* versus 0,293 para el caso de los 51 *loci*). La diferencia entre las medias fue estadísticamente significativa ($P < 0.001$). La varianza obtenida fue de 0,0093 mediante el análisis de los 20

loci, y de 0,0073 basándose en los 51 *loci*. Estas estimaciones de la varianza no difieren significativamente en su magnitud.

Estos resultados indican que el procedimiento utilizado para seleccionar los *loci* fue efectivo para identificar un subconjunto más informativo. Este subconjunto se comportó de manera similar al conjunto completo de 51 *loci* en términos de habilidad para distinguir cultivares de soja comerciales argentinas. El valor máximo de similitud genética obtenido para el subconjunto de 20 SSR corresponde a los cultivares DM 4800 y DM 50048, con un valor obtenido de 0,909. En el caso de los valores mínimos, para el subconjunto de 20 *loci* corresponden a los cultivares ACA470 y A5417 RG con un valor de similitud genética de 0,00, indicando que en los 20 *loci* SSR seleccionados estos cultivares no tienen ningún alelo en común.

Discusión

La caracterización mediante marcadores moleculares de los cultivares de soja que se utilizan actualmente en la producción argentina es necesaria, debido a la importancia que tiene la diferenciación e identificación de los mismos para el mejoramiento, así como también para el control del comercio y de la calidad de lotes de semilla. Esta caracterización permite contar con una base de datos de huellas genéticas de las variedades locales, determinar la variabilidad genética existente y seleccionar parentales divergentes para la obtención de nuevas variedades mediante cruzamiento.

En el presente trabajo se realizó una caracterización de 86 cultivares de soja argentina resistentes a glifosato, utilizadas en la actualidad y provenientes de las compañías semilleras más importantes del país, por medio de dos tipos de marcadores moleculares, SSR y AFLP. Se construyó una base de datos que fue luego validada mediante el ensayo de muestras incógnitas enviadas por semilleros o multiplicadores de semilla de soja, invirtiendo un tiempo de ensayo corto y utilizando un método robusto y reproducible. Se diseñó además un procedimiento simplificado para una clasificación eficaz y estandarizada de cualquier muestra incógnita bajo estudio.

El nivel de polimorfismos obtenidos para los 51 *loci* SSR analizados arrojó una media de 4,7 alelos por *locus* variando entre 2 y 9 alelos, y fue más bajo que el informado en otros trabajos como los de Rongwen et al. (1995) y Diwan & Cregan (1997). En el estudio de la diversidad genética en soja de Rongwen et al. (1995), realizado sobre 96 genotipos, se amplificaron entre 11 y 26 alelos por cada *locus* de microsatélites, aunque este número se vio reducido entre 5 y 10 alelos por *locus* en los 26 cultivares provenientes de programas de mejoramiento de soja norteamericano. Una posible razón para esta diferencia es que los materiales utilizados en el presente estudio provienen justamente de cultivares comerciales inscriptos, los cuales poseen una base genética estrecha. Sin embargo, los resultados de nuestro trabajo están en consonancia con los obtenidos por Priolli et al. (2002), quienes estudiaron 186 cultivares de soja brasileños por medio de 12 *loci* SSR y amplificaron entre 4 y 8 alelos por *locus*, con un promedio de 5,3 alelos. También concuerda con el trabajo de Giancola et al. (2002), quienes analizaron 30 *loci* SSR en 100 cultivares argentinos, y observaron entre 2 y 8 alelos por *loci* con un promedio de 4,4 alelos por *locus*.

El índice de diversidad génica (ID), que describe el potencial del *locus* para la discriminación de dos variedades tomadas al azar, tiene en cuenta para su cálculo tanto el número como la frecuencia de aparición de alelos en los cultivares analizados. En este trabajo el valor promedio de ID obtenido para marcadores de SSR fue de 0,523, mostrando el alto grado de polimorfismo de estos marcadores. Valores de diversidad génica similares fueron obtenidos por Maughan et al. (1995) en 62 cultivares de soja con un valor medio de ID de 0,55 o Narvel et al. (2000) con un valor medio de ID de 0,50 en 39 cultivares elite norteamericanos. Sin embargo, en los trabajos de Priolli et al. (2002) y Giancola et al. (2002), los valores medios de diversidad génica fueron levemente mayores que el obtenido en el presente trabajo (0,64 y 0,60, respectivamente). Esto puede deberse tanto a una disminución de la variabilidad genética, como así también a que en nuestro estudio se incluyó un número mayor de *loci* SSR que poseen alelos de muy baja frecuencia, conduciendo a valores de ID muy bajos (por ejemplo Satt 475 con un ID de 0,082). Estos alelos al ser poco frecuentes son útiles para la identificación de algún cultivar en particular, pero son poco polimórficos y por lo tanto hacen que el valor promedio obtenido en el cálculo de la diversidad genética sea menor.

La presencia de estos alelos de baja frecuencia en los marcadores SSR, como los encontrados en los *loci* Satt 534; 70; 231; 175; 294; 373; 414; 038; 197; 253; 191; 475; 458; 596; 292; 436; 409; 449; 703; 146; 556 y 095, podría reflejar la tasa de mutación de los microsatélites de soja, estimada en 10^{-5} a 10^{-4} por generación (Diwan & Cregan, 1997). Esta tasa es similar a la encontrada en humanos, y no implicaría una restricción para el uso de los microsatélites en la identificación de cultivares (Diwan & Cregan, 1997). Se ha determinado que para la identificación de cultivares de soja se debe utilizar un conjunto de entre 30 y 50 individuos, ya que los alelos mutantes que pudiesen surgir no serían detectados y de esta manera, las mutaciones en una planta aislada no alterarían la constitución alélica del cultivar (Diwan & Cregan, 1997). Por ejemplo, Song et al. (1999) encontraron 10 alelos nuevos en 66 cultivares de soja, los cuales no estuvieron presentes en 35 líneas ancestrales, mientras que Narvel et al. (2000) registraron 32 alelos específicos de cultivares elite, dentro de un total de 397 alelos que han sido detectados en 40 líneas y 39 cultivares de soja. Dado que en este trabajo no se analizaron las líneas progenitoras o ancestrales que dieron lugar a los cultivares actuales, no se puede concluir que pudiesen haber surgido alelos de baja frecuencia por este mecanismo.

En la caracterización con marcadores AFLP se utilizaron 15 combinaciones de cebadores que produjeron un total de 636 bandas de las cuales un 52% fueron polimórficas. Este valor alto de bandas polimórficas se debió principalmente a la presencia del cultivar DM3100, que presenta la mayoría de dichas bandas. El número promedio de bandas polimórficas por marcador fue de 17, resultado que está en concordancia con estudios realizados en cultivares de soja brasileña (Bonato et al., 2006) y cultivares chinos y norteamericanos (Ude et al., 2003) con 13 y 18 respectivamente. Las combinaciones de cebadores más informativos fueron EcoRIAAC/MseI-CTT y EcoRI-ACT/MseI-CTG, con 24 bandas polimórficas y los valores de diversidad génica más altos. Estos resultados contrastan con los obtenidos por otros grupos en donde las combinaciones más polimórficas fueron otras. Por ejemplo, Bonato (2006) y Zhu et al. (1999) encontraron que la combinación EcoRI-AAC/MseI-CTC fue la más polimórfica, mientras que Ude et al. (2003) describieron que los marcadores EcoRI-ACC/MseI-CAA producían la mayor cantidad de bandas polimórficas. Esta diferencia entre los resultados de varios autores podría estar indicando variaciones genéticas subyacentes en los germoplasmas caracterizados.

Los resultados obtenidos en Giancola et al. (2002) en cultivares de soja argentino mediante la utilización de marcadores AFLP no son comparables en forma directa con este trabajo, dado que estos autores utilizaron cebadores selectivos +4 (con cuatro bases de secuencia al azar en el extremo 3' del cebador) mientras que aquí utilizamos +3 (con sólo 3 bases de secuencia al azar), generando diferentes números y tipos de bandas. Los resultados obtenidos en este trabajo confirman resultados previos en los que se demuestra que los marcadores AFLP son una técnica molecular que detecta polimorfismos en múltiples *loci* simultáneamente, generando un gran número de marcadores reproducibles en un periodo corto de tiempo, siendo además una herramienta poderosa para el análisis de genotipos altamente relacionados (Powell et al., 1996).

El presente estudio demuestra que la diversidad génica esperada para los *loci* polimórficos utilizando marcadores SSR es mayor que para AFLP. Esto se relaciona con la variación del número medio de bandas polimórficas por unidad de ensayo, que varía de un 21% a 61% para AFLP a un máximo de 100% para SSR. El nivel de polimorfismo más alto detectado de los marcadores de SSR con respecto a AFLP destaca la capacidad discriminatoria de los primeros. Sin embargo, la comparación de la eficacia total de los dos sistemas de marcadores proporcionada por el índice del marcador (MI), indica que los marcadores AFLP poseen un MI 12 veces más alto con respecto a *loci* SSR. Esto confirma la naturaleza altamente discriminatoria del análisis de AFLP, que está en concordancia con informes anteriores en varias especies de plantas (Powell et al., 1996; Belaj et al., 2003; Saini et al., 2004; Medini et al., 2005). El valor distintivo del MI para los marcadores AFLP radica en el cociente multiplex eficaz. Es decir, en el alto número de bandas polimórficas obtenidas por experimento.

Los resultados comentados nos permiten inferir que los marcadores SSR permiten detectar un alto grado de polimorfismos, siendo muy útiles para estudios de diferenciación de cultivares (Diwan & Cregan, 1997). Pero, por otro lado, los marcadores AFLP permiten realizar estudios de comparación de cultivares muy rápidos, dado el gran número de bandas obtenidas por unidad de ensayo, de las cuales un número importante son polimórficas (Powell et al., 1996). El direccionamiento de los marcadores AFLP hacia cualquier región del genoma al azar garantiza una cobertura genómica completa, lo que potencialmente permitiría realizar análisis de

parentesco más precisos. En los dendrogramas y análisis de ordenamientos generados a partir de la matriz de AFLP se observó que los genotipos agruparon según el obtentor. Sin embargo, dado que carecíamos de datos de pedigrí, no pudimos analizar la asociación entre los perfiles de marcadores y el parentesco entre variedades.

En este trabajo los marcadores SSR arrojaron estimaciones de similitud genética más bajas con respecto a los marcadores de AFLP, obteniéndose valores medios de 0,292 y 0,921 para estas técnicas, respectivamente. Estos resultados eran esperables, ya que los SSR están dirigidos a regiones altamente variables, mientras los AFLPs detectan *loci* que presentan un amplio rango en su nivel de polimorfismo, desde los que son completamente conservados hasta los hipervariables, y la posibilidad de detectar este tipo de *loci* está relacionada con la frecuencia de su ocurrencia en el genoma. Nuestros resultados están en completo acuerdo con otros estudios que comparan el nivel del polimorfismo detectado con los marcadores de AFLP y SSR en la soja (Powell et al., 1996), el maíz (Pejic et al., 1998), el sorgo (Menz et al., 2004), y el triticale (Tams et al., 2005).

La correlación existente entre las estimaciones de la similitud genética entre marcadores AFLP y SSR para los cultivares de

soja evaluados en este trabajo es muy baja ($r = 0,166$), aunque significativa. Valores comparables fueron encontrados en soja por Giancola et al. (2002), que observaron una correlación entre las estimaciones de similitud genética de marcadores AFLP y SSR de 0,210. Asimismo, Powell et al. (1996) encontraron valores de correlación bajos entre estos marcadores. Las estimaciones de la similitud genética se ven influenciadas por una variedad de factores entre los que se pueden contar la distribución de los marcadores en el genoma, la naturaleza de los mecanismos que generan la variación medida y errores en la clasificación de bandas no homólogas de *loci* AFLP de movilidad similar. Por una parte, los marcadores AFLP se originan al azar, a partir de regiones que pueden resultar hipervariables, medianamente variables o muy conservadas. Sin embargo, dado que la mayor parte de las regiones del genoma de soja están conservadas entre las variedades, un mayor número de marcadores presenta perfiles monomórficos, por lo que las tasas de variación son más bajas que las observadas en los *loci* SSR, los cuales están dirigidos específicamente a regiones altamente polimórficas (Powell et al., 1996). Es de interés mencionar que aunque en general se considera que la distribución de los marcadores de AFLP en el genoma es azarosa, se han notificado casos de sobrerepresentación de ciertas secuencias genómicas, como por ejemplo regiones centroméricas (Menz et al. 2004).

Conclusiones

En este trabajo se informa el desarrollo de una base de datos genéticos de SSR y AFLP que permitió la caracterización de las 86 variedades de soja más cultivadas actualmente en Argentina. Mientras los marcadores de SSR resultaron útiles para la rápida identificación de variedades, los de AFLP fueron más apropiados para la caracterización global del genoma y la

selección de genotipos divergentes. Estas bases de datos están siendo utilizadas en la Bolsa de Comercio de Rosario para prestar un servicio de clasificación y caracterización de genotipos. Su utilización puede contribuir al mejoramiento del germoplasma argentino de soja y a la protección de la propiedad intelectual.

Agradecimientos

Agradecemos a los semilleros ACA, AL Marchionni, Don Mario, FN Semillas, La Tijereta, Nidera, Pioneer, Relmó, SPS y Syngenta, por proveernos gentilmente el material vegetal utilizado en este análisis. Este estudio fue financiado por la Bolsa de Comercio de Rosario. Silvina C. Pessino es miembro de la Carrera de Investigador Científico del Consejo Nacional de Investigaciones Científicas y Tecnológicas (CONICET).

Bibliografía

1. AKKAYA MS, BHAGWAT AA, CREGAN PB (1992) The use of microsatellite DNA markers for soybean genotype identification. *Genetics*, 132: 1131—1139.
2. AKKAYA MS, SHOEMAKER RC, SPECHT JE, BHAGWAT AA, CREGAN PB (1995) Integration of simple sequence repeat DNA markers into a soybean linkage map. *Crop Sci*, 35: 1439-1445.
3. BELAJ A, SATOVIC Z, CIPRIANI G, BALDONI L, TESTOLIN R, RALLO L, TRUJILLO I (2003) Comparative study of the discriminating capacity of RAPD, AFLP and SSR markers and of their effectiveness in establishing genetic relationships in olive. *Theor. Appl. Genet*, 107: 736-744.
4. BONAMICO N, AIASSA J, IBAÑEZ M, DI RENZO M, DÍAZ D, SALERNO J (2004) Caracterización y clasificación de híbridos simples de maíz con marcadores SSR. *Revista de Investigaciones Agropecuarias, INTA, Argentina*, 33: 129-144.
5. BONATO A, CALVO E, GERALDI I, ARIAS C (2006) Genetic similarity among soybean (*Glycine max L. Merrill*) cultivars released in Brazil using AFLP markers. *Genet Mol Biol*, 29: 692-704.
6. BOTSTEIN D, WHITE R, SKOLNICK M, DAVIS RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet*, 32: 314-331.
7. BROWN-GUEDIRA G, THOMPSON J, NELSON R, WARBURTON M (2000) Evaluation of genetic diversity of soybean introductions and North American ancestors using RAPD and SSR markers. *Crop Sci*, 40: 815-823.
8. CREGAN PB, BHAGWAT AA, AKKAYA MS, RONGWEN J (1994) Microsatellite fingerprinting and mapping of soybean. *Methods Mol Cell Biol*, 5: 49-61.
9. CREGAN PB, JARVIK T, BUSH AL, SHOEMAKER RC, LARK KG, KAHLER AL, KAYA N, VAN TOAITT, LOHNES DG, CHUNG J, SPECHT JE (1999). An integrated genetic linkage map of the soybean genome. *Crop Sci*, 39: 1464-1490.
10. DIWAN N, CREGAN PB (1997) Automated sizing of fluorescent-labeled simple sequence repeat (SSR) markers to assay genetic variation in soybean. *Theor Appl Genet*, 95: 723-733.
11. DONG YS, ZHAO LM, LIU B, WANG ZW, JIN ZQ, SUN H (2004) The genetic diversity of cultivated soybean grown in China. *Theor Appl Genet*, 108: 931-936.
12. GIANCOLA S, MARCUCCI POLTRI S, LACAZE P, HOPP HE (2002) Feasibility of integration of molecular markers and morphological descriptors in a real case study of a plant variety protection system for soybean. *Euphytica*, 127: 95-113.

13. **GOWER JC** (1966) Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika*, 53:325-338.
14. **HYMOWITZ T, NEWELL CA** (1981) Taxonomy of the genus *Glycine*, domestication and uses of soybeans. *Econ Bot*, 35:272-288.
15. **JACCARD P** (1908) Nouvelles recherches sur la distribution florale. *Bull. Soc. Vaud. Sci. Nat.*, 44:223-270.
16. **JONES CJ, EDWARDS KJ, CASTAGLIONE S, WINFIELD MO, SALA F, VAN DEWIEL C, BREDEMEIJER G, VOSMAN B, MATTHES M, DALY A, BRETTSCHEIDER R, BETTINI P, BUIATTI M, MAESTRI E, MALCEVSCI A, MARMIROLI N, AERT R, VOLCKAERT G, RUEDA J, LINACERO R, VAZQUEZ A, KARP A** (1997) Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. *Mol Breed*, 3: 381-390.
17. **KEIM P, BEAVIS WD, SCHUPP J, FREESTONE R** (1992) Evaluation of soybean RFLP marker diversity in adapted germplasm. *Theor Appl Genet*, 85:205-212.
18. **KOSMAN E, LEONARD KJ** (2005) Similarity coefficients for molecular markers in studies of genetic relationships between individuals for haploid, diploid, and polyploid species. *Molecular Ecology* 14: 415-424.
19. **LEAVY S, DEWES H** (2009) Drivers y tendencias en los escenarios futuros de la cadena de la soja argentina. Anais do Congresso da SOBER (Sociedade Brasileira de Economia, Administração e Sociologia Rural), Porto Alegre, pp. 21.
20. **MANTEL NA** (1967) The detection of disease clustering and a generalized regression approach. *Cancer Res*, 27: 209-220.
21. **MARTINEZ F** (2010) Crónica de la soja en la región pampeana argentina. Para mejorar la producción, INTA EEA OLIVEROS, pp 141.
22. **MAUGHAN PJ, SAGHAI MAROOF MA, BUSS GR, HUESTIS GM** (1996) Amplified fragment length polymorphism (AFLP) in soybean: Species diversity, inheritance, and near-isogenic line analysis. *Theor Appl Genet* 93:392-401.
23. **MEDINI M, HAMZA S, REBAI A, BAUM M** (2005) Analysis of genetic diversity in tunisian durum wheat cultivars and related wild species by SSR and AFLP markers. *Genetic Resources and Crop Evolution*, 52: 21-31.
24. **MEESANG N, RANAMUKHAARACHCHI SL, PETERSEN MJ, ANDERSEN SB** (2001) Soybean cultivar identification and genetic purity analysis using microsatellite DNA markers. *Seed Science and Technology*, 29: 637-645.
25. **MENZ MA, KLEIN RR, UNRUH NC, ROONEY WL, KLEIN PE, MULLET JE** (2004) Genetic diversity of public inbreds of *Sorghum* determined by mapped AFLP and SSR markers. *Crop Sci* 44: 1236-1244.
26. **MUELLER UG, LAREESA WOLFENBARGER L** (1999) AFLP genotyping and fingerprinting. *Trends Ecol Evol*, 14: 389-394.
27. **MULLIS K, FALOONA F, SCHARF S, SAIKI R, HORN G, ERLICH H** (1986) Specific Enzymatic Amplification of DNA *in vitro*: The Polymerase Chain Reaction. *Proceedings of the Cold Spring Harbor Symposia on Quantitative Biology*, pp. 263-273.
28. **NARVEL JM, FEHR WR, CHU WS, GRANT D, SHOEMAKER RC** (2000) Simple sequence repeat diversity among soybean plant introductions and elite genotypes. *Crop Sci* 40: 1452-1458.
29. **NEI M** (1978) Estimation of average heterozygosity and genetic distances from a small number of individuals. *Genetics*, 89: 583-590.
30. **OLSINA C** (2012) Desarrollo de un banco de datos genéticos del germoplasma de soja argentino, Tesis de Maestría, Maestría en Genética, Facultad de Ciencias Agrarias, Universidad Nacional de Rosario, Rosario, Argentina.
31. **PEJIC I, AJMONE-MARSAN P, MORGANTE M, KOZUMPLICK V, CASTIGLIONI P, TARAMINO G, MOTTO M** (1998) Comparative analysis of genetic similarity among maize inbred lines detected by RFLPs, RAPDs, SSRs, and AFLPs. *Theor. Appl. Genet.* 97: 1248-1255.
32. **POWELL W, MORGANTE M, ANDRE C, HANAFEY M, VOGEL J, TINGEY S, RAFALSKY A** (1996) The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germoplasm analysis. *Mol Breed*, 2: 255-238.
33. **PRIOLLI RHG, MENDES-JUNIOR CT, ARANTES NE, CONTEL EPB** (2002) Characterization of brazilian soybean cultivars using microsatellite markers. *Genet Mol Biol*, 25: 185-193.
34. **RONGWEN J, AKKAYA MS, BHAGWAT AA, LAVI U, CREGAN PB** (1995) The use of microsatellite DNA markers for soybean genotype identification. *Theor. Appl. Genet.* 90: 43-48.
35. **SAINI N, JAIN N, JAIN S, JAIN RK** (2004) Assessment of genetic diversity within and among Basmati and non-Basmati rice varieties using AFLP, ISSR and SSR markers. *Euphytica* 140: 133-146.
36. **SHOEMAKER RC, GUFFY RD, LORENZEN LL, SPECHT JE** (1992) Molecular mapping of soybean: Map utilization. *Crop Sci*, 32: 1091-1098.
37. **SNEATH PHA, SOKAL RR** (1973) En: *Numerical Taxonomy*, W.H. Freeman and Company, San Francisco, California, USA, pp. 230-234.
38. **SONG QJ, QUIGLEY CV, NELSON RL, CARTER TE, BOERMA HR, STRACHAN JL, CREGAN PB** (1999) A selected set of trinucleotide simple sequence repeats markers for soybean cultivar identification. *Plant Var and Seed*, 12: 207-220.
39. **SONG QJ, MAREK LF, SHOEMAKER RC, LARK KG, CONCIBIDO VC, DELANNAY X, SPECHT JE, CREGAN PB** (2004) A new integrated genetic linkage map of the soybean. *Theor Appl Genet*, 109:122-128.
40. **TAMS SH, MELCHINGER AE, BAUER E** (2005) Genetic similarity among European winter triticale elite germplasms assessed with AFLP and comparisons with SSR and pedigree data. *Plant Breeding*, 124: 154-160.
41. **UDE GN, KENWORTHY WJ, COSTA JM, CREGAN PB, ALVERNANAZ J** (2003) Genetic diversity of soybean cultivars from China, Japan, North America and North American ancestral lines determined by amplified fragment length polymorphism. *Crop Sci*, 43: 1858-1867.
42. **VOS P, HOGERS R, BLEEKER M, REIJANS M, VAN DE LEE T, HORNES M, FRIJTERS A, POT J, KUIPER M, ZABEAU M** (1995) AFLP: A new technique for DNA fingerprinting. *Nucleic Acids Res*, 23: 4407-4414.
43. **WEBER JL** (1990) Informativeness of human (dC-dA)n(dG-dT)n polymorphisms. *Genomics*, 7: 524-530.
44. **WELSH J, MCCLELLAND M** (1990) Fingerprinting genomes using de PCR with arbitrary primers. *Nucl Acid Res*, 18: 7213-7218.
45. **WILLIAMS JGK, KUBELIK AR, LIVAK KJ, RAFALSKI JA, TINGEY SV** (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucl Acid Res*, 18: 6531-6535.
46. **ZHAO TJ, GAI JY** (2004) The origin and evolution of cultivated soybean [*Glycine max (L) Merr.*]. *SciAgricult Sin*, 37: 954-962.
47. **ZHU SL, MONTI LM, AVITABILE A, RAO R** (1999) Evaluation of genetic diversity in Chinese soybean germplasm by AFLP. *Plant Genet Res Newsletter*, 119: 10-14.