Biotecnología Aplicada

Elfos scientiae
http://elfosscientiae.cigb.edu.cu/

RESEARCH

Analysis of genetic polymorphisms among glyphosate-resistant soybean [Glycine max (L.) Merr.] lines

Celia Delgado[1], Oreste Corrales[1], Yuniet Hernández[2], Natacha Soto[1], Aleines Ferreira[1], Oswalt R Jiménez[3], ✍Gil A Enríquez[1]

[1]Departamento de Biotecnología de Plantas, Dirección de investigaciones Agropecuarias, Centro de Ingeniería Genética y Biotecnología, CIGB
Ave. 31 entre 158 y 190, Cubanacán, Playa, CP 11600, La Habana, Cuba
[2]Plant RNAi Biology Group, International Center for Genetic Engineering and Biotechnology, New Delhi, India
[3]The Institute for Training, Research and Environmental Development, CIDEA, University of Central America, Managua, Nicaragua
✍gil.enriquez@cigb.edu.cu

ABSTRACT

The Cuban National Program for Soybean Genetic Breeding [*Glycine max* (L.) Merr.] has been focused on improving yield as well as adding new traits for better crop management, including the introgression of transformation event GTS 40-3-2 into advanced lines for conferring resistance to glyphosate herbicide. However, there are no reports on the use of molecular markers to evaluate genotypes obtained from this program. Therefore, in this work, a practical application of RAPD, ISSR and SSR markers is provided, to analyze the genetic diversity, estimate pedigree relationships and confirm genetic identity of nine elite soybean genotypes engaged in the current breeding program. Five glyphosate-resistant soybean lines and their respective parents were studied. Molecular markers were able to distinguish among cultivars. The UPGMA dendrogram, constructed using the three marker systems together, effectively clustered each inbred line with one of their parents. The low genetic similarity coefficient (0.4) obtained from parents IncaSoy1, IncaSoy36, CEB2 and CEB4 confirmed the observed genetic differences. Two genotypes, CEB4 and RP5 were precisely detected from a collection of 15 soybean cultivars during validation testing. RAPD amplicons were converted into two new SCAR markers that successfully detected CEB4 and contributed to build a molecular profile for the promising line IncaSoy36, something crucial for future actions based on the breeding potential of both lines. This study demonstrates that the use of this marker system could provide substantial benefits for soybean breeders and the seed industry in Cuba and other countries which breeding projects rely on local germplasm and facilities.

Keywords: soybean, breeding, genetic diversity, RAPD, ISSR, SSR, SCAR markers

RESUMEN

Análisis del polimorfismo genético en líneas de soya [*Glycine max* (L.) Merr.] resistentes a glifosato. El Programa de Mejoramiento Genético de la Soya [*Glycine max* (L.) Merr.] en Cuba se ha enfocado en mejorar el rendimiento y adicionar nuevos caracteres para un mejor manejo del cultivo. Este programa incluye la introgresión del evento de transformación GTS 40-3-2 en líneas avanzadas para conferir resistencia al herbicida glifosato. Sin embargo, no se han evaluado los genotipos obtenidos mediante marcadores moleculares. En este trabajo se usaron marcadores moleculares RAPD, ISSR y SSR para el análisis de la diversidad genética, la estimación de las relaciones de parentesco y para confirmar la identidad genética de nueve genotipos élite de soya (cinco líneas resistentes a glifosato y sus respectivos parentales). Los marcadores moleculares permitieron distinguir a los distintos cultivares. El dendrograma UPGMA construido a partir de estos agrupó a cada línea singénica con uno de sus parentales. El bajo coeficiente de similitud genética (0.4) de los parentales IncaSoy1, IncaSoy36, CEB2 y CEB4 confirmaron las diferencias genéticas observadas. Se detectaron con precisión los genotipos CEB4 y RP5 con el uso de marcadores SCAR y SSR, respectivamente, a partir de una colección de cultivares durante diferentes pruebas. Los amplicones de RAPD se convirtieron en dos nuevos marcadores SCAR, lo que permitió detectar al genotipo CEB4 y la elaboración de un perfil molecular del cultivar IncaSoy36, cruciales para futuras acciones basadas en el potencial de mejora de ambas líneas. Este sistema de marcadores podría proveer beneficios sustanciales para los mejoradores de soya y su industria semillera, y en proyectos basados en germoplasma y capacidades locales en otros países.

Palabras clave: soya, mejoramiento, diversidad genética, marcadores RAPD, marcadores ISSR, marcadores SSR, marcadores SCAR

How to cite (Vancouver style):
Delgado C, Corrales O, Hernández Y, Soto N, Ferreira A, Jiménez OR, et al. Analysis of genetic polymorphisms among glyphosate-resistant soybean [Glycine max (L.) Merr.] lines. Biotecnol Apl. 2019;36(3):3301-6.

Introduction

Soybean [*Glycine max* (L.) Merr.] production is highly important for many countries, because it provides valuable products for human and animal feeding. In Cuba, the production of this legume has become a priority due to the need for decreasing imports of soybean and all its derivatives, but yields are still too low (approx. 1.5 tons per ha) despite soybean production areas are increasing. Insufficient weed management is one of the factors influencing on the low performance of available cultivars. Thus, the utilization of herbicide-resistant genotypes seems to be an option to overcome this limitation [1].

Since 2008, a soybean breeding program has been developed, aiming to obtain herbicide-resistant

1. Delgado C, Enríquez GA, Ortiz R, Céspedes O, Soto N, Hernández Y, *et al.* Glyphosate resistance trait into soybean Cuban varieties: agronomical assessment of transgenic lines until F6 generation. Int J Agron Agric Res. 2015;7:75-85.

soybean cultivars through a combination of traditional and biotechnological methods. The project has included the introgression of GTS 40-3-2 event, which confers glyphosate resistance, into well-adapted soybean genotypes. Elite soybean cultivars were selected as parents based on their morpho-agronomical traits and yield performance. Agronomic value was the fundamental criterion for parental selection, and genetic divergence remains to be considered as well, because novel and better-quality cultivars can be generated from sources of high genetic variability [2]. However, little is known about the genetic composition of Cuban soybean germplasm. Pedigree and morpho-agronomic characters of the cultivars are just the evidences available at the Inscription Office of Vegetable Varieties belonging to Agriculture Ministry. This lacking information about genetic variability have triggered that much of the applied research have been conducted without proper genetic knowledge at DNA level.

In this setting, molecular markers have proved to be a valuable tool for examining genetic variation in crops. Even in self-pollinated species like soybean, a precise characterization and unique identification can be performed using DNA markers [3]. Among them, Random Amplified Polymorphic DNA (RAPD) and Simple Sequence Repeats (SSR) markers have been broadly used for diversity analyses and cultivar identification in soybean [2, 4-8]. The RAPD analysis is quick and easy to execute, small amount of DNA is required and generates large number of dominant molecular markers often distributed all over the genome. But, it has a poor reproducibility [9].

Remarkably, the conversion of RAPD markers into Sequence Characterized Amplified Regions (SCAR) ones [10] greatly solves this fault. Meanwhile, SSR markers are highly polymorphic, codominant and have great ability to identify unique alleles in soybean populations [11-13]. Finally, Inter Simple Sequence Repeat (ISSR) markers are highly polymorphic and the technique is simple, quick and very reproducible. These markers have been useful in genetic diversity studies in soybean [14] and to create soybean molecular identity systems [15, 16]. Many studies about molecular characterization of soybean germplasm using DNA markers are available and the amount is still growing. Knowledge on genetic polymorphism among soybean lines could be useful to test genetic purity and to evaluate breeding potential of adapted germplasm. But there are none using molecular markers to evaluate soybean genotypes obtained from soybean breeding programs in Cuba.

Therefore, in this work, the genetic polymorphisms of five glyphosate-resistant soybean lines developed in the soybean breeding program and their respective parents were analyzed using RAPD, ISSR and SSR markers together. Furthermore, novel SCAR markers were generated, useful for cultivar identification in some of the assessed genotypes.

Materials and methods

Plant material and genomic DNA isolation

A set of nine soybean genotypes were analyzed using RAPD, ISSR and SSR markers. Five inbred lines I1B2-2 (IncaSoy1×CEB2); I1B2-3 (IncaSoy1×CEB2); I1B4 (IncaSoy1×CEB4); I36B4 (IncaSoy36×CEB4), RP5 (IncaSoy36 backcrossing) and their parents IncaSoy1, IncaSoy36, CEB2 and CEB4 were used. The inbred lines and the parents CEB2 y CEB4 bear the transformation event GTS 40-3-2 that confers resistance to glyphosate. Other twelve lines were used for SCAR primer validation: IncaSoy27, IncaSoy35, SCIGB-L1, Inifat-120, Inifat-195, Inifat-243, Inifat-304, DT84, FT2, DT-84, CubaSoy23 and Williams 82. Glyphosate-resistant genotypes were obtained from the Cuban soybean breeding program developed at the Center for Genetic Engineering and Biotechnology of Havana, Cuba.

Total genomic DNA of each soybean genotype was isolated from young leaves through the CTAB protocol described by Doyle and Doyle [17]. DNA quality was checked in a 0.8 % agarose gels, stained with ethidium bromide and its DNA concentration was assessed in a nano-spectrophotometer (NP80, IMPEM, Germany). The isolated DNA was adjusted to a final concentration of 50 ng/μL and stored at –20 °C until use.

RAPD, ISSR and SSR analyses

All PCR reaction tubes contained genomic DNA (50 ng for RAPD and SSR, 200 ng for ISSR), 1× buffer (50 mM KCl, 10 mM Tris-HCl, pH 8.3), dNTP mix (0.2 mM), $MgCl_2$ (3 mM for RAPD and ISSR, 1.5 mM for SSR), primers (0.4 μM for RAPD and ISSR, 0.2 μM of each SSR primer) and Taq DNA Polymerase (5 U) in a final volume of 25 μL. RAPD, ISSR and SSR primer sequences appear in table 1.

Table 1. List of primers/locus of RAPD, ISSR, SSR and SCAR markers used to assess the genetic polymorphism of soybean genotypes I1B2-2, I1B2-3, I1B4, I36B4, RP5, IncaSoy36, IncaSoy1, CEB2 and CEB4

Marker	Primer/ Locus	Sequence (5´-3´)
RAPD	UBC-240	ATGTTCCAGG
	UBC-256	TGCAGTCGAA
	UBC-259	GGTACGTACT
	UBC-268	AGGCCGCTTA
	OPC-14	TGCGTGCTTG
ISSR	R-5	$(AGC)_6C$
	R-7	$(TCG)_6G$
	UBC-807	$(AG)_8T$
	UBC-857	$(AC)_8CG$
SSR	Satt-005	F: GTCGATTAGGCTTGAAATA
		R: TATCCTAGAGAAGAACTAAAAAA
	Satt-009	F: CTTACTAGCGTATTAACCCTT
		R: CCAACTTGAAATTACTAGAGAAA
	Satt-141	F: CCGTCATAAAAAGTCCCTCAGAAT
		R: CGGTGGTGGTGTGCATAATAA
	Satt-146	F: GTGGTGGTGGTGAAAACTATTAGAA
		R: AAGGGATCCCTCAACTGACTG
	Satt-373	F: GGCCAGATACCCAAGTTGTACTTGT
		R: TCCGCGAGATAAATTCGTAAAAT
SCAR	$IntSC_{538}$	F: GTCGATTAGGCTTGAAATA
		R: TATCCTAGAGAAGAACTAAAAAA
	$FqSC_{591}$	F: CTTACTAGCGTATTAACCCTT
		R: CCAACTTGAAATTACTAGAGAAA
	$AdSC_{1161}$	F: CCGTCATAAAAAGTCCCTCAGAAT
		R: CGGTGGTGGTGTGCATAATAA
	ID1G-F	AAGACGATTTTTCAAAGAACTGTATTC
	ID3G-F	TTCTACTTTGTGACATAATGTTCCAGG
	ID4G-R	TCAGTAAGACTAATTTCTAATTTTTGC

* F, R: Forward and Reverse primers, respectively.

2. Bisen A, Khare D, Nair P, Tripathi N. SSR analysis of 38 genotypes of soybean (*Glycine Max* (L.) Merr.) genetic diversity in India. Physiol Mol Biol Plants. 2015;21:109-15.

3. Chowdhury AK, Srinives P, Tongpamnak P, Saksoong P, Chatwachirawong P. Genetic relationship among exotic soybean introductions in Thailand: Consequence for varietal registration. Sci Asia. 2002;28:227-39.

4. Brown-Guedira G, Thompson J, Nelson R, Warburton M. Evaluation of genetic diversity of soybean introductions and North American ancestors using RAPD and SSR markers. Crop Sci. 2000;40:815-23.

5. Li Y, Guan R, Liu Z, Ma Y, Wang L, Li L, et al. Genetic structure and diversity of cultivated soybean (*Glycine max* (L.) Merr.) landraces in China. Theor Appl Genet. 2008;117:857-71.

6. Tantasawat P, Trongchuen J, Prajongjai T, Jenweerawat S, Chaowiset W. SSR analysis of soybean (*Glycine max* (L.) Merr.) genetic relationship and variety identification in Thailand. Austr J Crop Sci. 2011;5:283-90.

7. Khan F, Hakeem KR, Siddiqi TO, Ahmad A. RAPD markers associated with salt tolerance in soybean genotypes under salt stress. Appl Biochem Biotechnol. 2013;170:257-72.

8. Anwar-Malik MF, Tariq K, Qureshi AS, Khan MR, Ashraf M, Naz G, *et al*. Analysis of genetic diversity of soybean germplasm from five different origins using RAPD markers. Acta Agric Scand, Section B-Soil Plant Sci. 2017;67:148-54.

9. Skroch P, Nienhuis J. Impact of scoring error and reproducibility RAPD data on RAPD based estimates of genetic distance. Theor Appl Genet. 1995;91:1086-91.

10. Paran I, Michelmore R. Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce. Theor Appl Genet. 1993;85:985-93.

11. Narvel JM, Fehr WR, Chu WC, Grant D, Shoemaker RC. Simple sequence repeat diversity among soybean plant introductions and elite genotypes. Crop Sci. 2000;40:1452-8.

12. Wang L, Guan R, Zhangxiong L, Chang R, Qiu L. Genetic diversity of Chinese cultivated soybean revealed by SSR markers. Crop Sci. 2006;46:1032-8.

13. Rodrigues DH, de Alcântara Neto F, Schuster I. Identification of essentially derived soybean cultivars using microsatellite markers. Crop Breed Appl Biotechnol. 2008;8:74-8.

14. Baloch FS, Kurt C, Arioğlu HH, Özkan H. Assaying of diversity among soybean (*Glycine max* (L.) Merr.) and peanut (*Arachis hypogaea* L.) genotypes at DNA level. Turkish J Agric Forest. 2010;34:285-301.

15. Brick A, Sivolap YM. Molecular genetic identification and certification of soybean (*Glycine max* L.) cultivars. Russ J Genet. 2001;37:1061-7.

16. Meena RK, Verma AK, Thakur S. Molecular identity using Inter-simple Sequence Repeat & Random Amplified Polymorphic DNA markers in soybean (*Glycine max*) cultivars. Curr Trends Biotechnol Pharm. 2015;9:16-22.

17. Doyle J, Doyle J. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull. 1987;19:11-5.

DNA amplifications were performed in a PTC-100™ Thermal Cycler (JM Research, Inc) programmed for each marker system as follow: RAPD: pre-denaturation at 94 °C for 30 s, then 3 cycles of denaturing at 94 °C for 1 min, annealing at 38 °C for 1 min and extension at 72 °C for 2 min, followed by 35 cycles at 94 °C for 1 min, 40 °C for 1 min and 72 °C for 2 min and final extension at 72 °C for 5 min. ISSR: pre-denaturation at 94 °C for 3 min, following for 49 cycles at 94 °C for 1 min, 52 °C (UBC-807 and UBC- 857) or 58 °C (R-5 and R-7) for 1 min and 72 °C for 2 min. A final extension step was carried out at 72 °C for 10 min. SSR: pre-denaturation at 95 °C for 3 min, 40 cycles at 95 °C for 50 s, 48 °C for 40 s and 72 °C for 1 min, and final extension at 72 °C for 10 min. RAPD and ISSR amplification products were resolved by 1.5 % agarose gel electrophoresis, whereas SSR products were fractionated in a 3 % agarose gel, both in 1× TBE buffer, under 110 V and 240 mA current. DNA bands were stained with ethidium bromide (10 µg/mL) and visualized under ultraviolet light. The size of amplified products was determined by comparison with 100 bp & 1 kb DNA ladder size marker (Promega, USA) depending of band sizes. Each experiment was repeated at least three times to confirm the specific pattern and reproducibility.

Data analysis

RAPD, ISSR and SSR bands were manually scored as present (1) or absent (0). Only clear and strong bands were recorded and entered into a binary data matrix. Genetic similarities were estimated according to the method developed by Dice [18]. Similarity values from the matrix were used for cluster analyses via the unweighted pair group method with arithmetic mean (UPGMA) method. The PAST program version 1.99 [19] was used for genetic similarity computing and dendrogram construction.

The polymorphism information content (PIC) of each marker was calculated according to the formula:

$$PIC = 1 - \sum_{i=1}^{n} f_i^2$$

where: f_i^2 is the band frequency of the i^{th} allele [20].

In case of RAPD and ISSR markers, the PIC value was calculated as 1– p2 – q2, where p is the band frequency and q is the no band frequency [21]. RAPD and ISSR primer index were calculated by adding all the PIC values of all markers amplified by the same primer and divided by the total number of markers [22].

Cloning and sequencing of a RAPD amplicon

The specific RAPD amplicon obtained with UBC-240 primer from IncaSoy36 and its progeny were separated by gel electrophoresis and excised from LGT agarose gel. Amplicon purification was performed using the Wizard® Plus SV Minipreps, DNA Purification System (Promega). The eluted DNA fragments were then ligated into the pGEM-T Easy Vector (Promega) following supplier instructions and transformed into *Escherichia coli* XL-1 Blue competent cells. The recombinant plasmid DNA was isolated from the white colonies using the method proposed by Birnboim and Doyle [23] with minor modifications. Five distinct white colonies were chosen from the Luria-Bertani (LB)/ampicillin/X-gal/isopropy1-β-D-1-thiogalactopyranoside (IPTG) plate and the recombinant plasmid was isolated from each overnight colony. The inserted fragments were sequenced using M13 primers in a 96-capillary ABI 3730xl sequencer (Macrogen Inc., Korea). DNA sequence analysis was performed using the BLAST sequence analysis programs at the National Centre for Biotechnology Information (NCBI; https://blast.ncbi.nlm.nih.gov/).

Primer design and SCAR marker analysis

Three SCAR primer pairs were designed based on the sequence analysis of the cloned fragment mentioned above. $IntSC_{538}$-F/$IntSC_{538}$-R primers were designed to hybridize in the 5′ and 3′ ends of the 546 bp fragment; $FqSC_{591}$-F /$FqSC_{591}$-R in the flanking sequences of the fragment and $AdSC_{1161}$-F/$AdSC_{1161}$-R in adjacent regions surrounding the flanking sequences. The nine soybean genotypes previously analyzed were also evaluated using these SCAR primers. PCR conditions were the same as described for RAPD except that the genomic DNA concentration was 500 ng/µL. Amplification products obtained from IncaSoy36, RP5, I36B4 and CEB4 using $AdSC_{1161}$-F/$AdSC_{1161}$-R primers were purified, cloned and sequenced as described above. Then, a multiple sequence alignment was performed through the Clustal Omega program (http://www.ebi.ac.uk/Tools/msa/clustalo/). New SCAR primers were designed: ID3G-F from IncaSoy36 sequence, ID1G-F from CEB4 sequence and ID4G-R from a common sequence of both. Table 1 shows the sequences of all SCAR primers.

The SCAR primer combinations: ID3G-F/ID4G-R and ID1G-F/ID4G-R were assessed in the nine soybean genotypes used in this study. Additionally, the analyses were conducted on a set of 15 soybean cultivars: IncaSoy1, IncaSoy27, IncaSoy35, IncaSoy36, Inifat-120, Inifat-195, Inifat-243, Inifat-304, DT84, FT2, DT-84, CubaSoy23 and Williams 82, provided by the National Institute of Agricultural Sciences of Cuba; and CEB4 and SCIGB-L1, provided by the Cuban soybean breeding program. Genomic DNA of each genotype was isolated as described above. PCR reactions were performed in a 25 µL reaction tube containing 500 ng of template DNA, 1× buffer, dNTP mix (0.2 mM), $MgCl_2$ (3 mM), SCAR primers (0.2 µM) and Taq DNA polymerase (5 U). Thermal cycling conditions were optimized as: 95 °C for 5 min; 30 cycles at 95 °C for 30 s, 60 °C for 30 s and 72 °C for 1 min; and a final extension at 72 °C for 10 min. Amplification products were stained with ethidium bromide (10 µg/mL) and visualized under UV light.

Results

The DNA marker systems RAPD, ISSR and SSR were used to detect the polymorphism among nine soybean genotypes. A total of 51 DNA fragments which sizes were from 250 to 3000 bp were amplified using five RAPD primers. Out of 51 scored bands, 14 were polymorphic. Primer UBC-240 had the highest percentage of polymorphism meanwhile primer OPC-14 had the lowest (Table 2). Unique band patterns performed by three RAPD primers were observed in some genotypes such as I36B4 (UBC-240), I1B4 (UBC-268), I1B2-3 and IncaSoy36 (UBC-259). In ISSR analysis,

18. Dice LR. Measures of the amount of ecologic association between species. Ecology. 1945;26:297-302.

19. Hammer Ø, Harper DAT, Ryan P. PAST-Palaeontological statistics. 2001 [cited 2018 Oct 17]. Available from: www.uv.es/~pardomv/pe/2001_1/past/pastprog/past.pdf

20. Smith JS, Chin EC, Shu H, Smith OS, Wall SJ, Senior ML, *et al*. An evaluation of the utility of SSR loci as molecular markers in maize (*Zea mays* L.): comparisons with data from RFLPs and pedigree. Theor Appl Genet. 1997;95:163-73.

21. Ghislain M, Zhang D, Fajardo D, Huamán Z, Hijmans RJ. Marker-assisted sampling of the cultivated Andean potato *Solanum phureja* collection using RAPD markers. Genet Resour Crop Evol. 1999;46:547-55.

22. Raina S, Rani V, Kojima T, Ogihara Y, Singh K, Devarumath R. RAPD and ISSR fingerprints as useful genetic markers for analysis of genetic diversity, varietal identification, and phylogenetic relationships in peanut (*Arachis hypogaea*) cultivars and wild species. Genome. 2001;44:763-72.

23. Birnboim H, Doyle J. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. 1979;7:1513-23.

30 amplification products were scored using four ISSR primers. All primers generated polymorphic bands. The maximum percentage of polymorphism was 40 % (R-7) and the minimum 14.3 % (UBC-857). Only one distinctive pattern was produced identifying CEB4 (UBC-857). The average of the Polymorphic Index Content (PIC) values calculated for each individual markers of RAPD and ISSR ranged from 0.03 to 0.20 (Table 2). Five SSR loci were also screened across the nine soybean genotypes. The number of alleles per locus remained between two and three in all loci. The allelic sizes ranged from 150 to 320 bp (Table 2). A unique allele of Satt-005 was detected in RP5. All genotypes were heterozygous for Satt-146 and Satt-373 loci. PIC values for individual locus ranged from 0.35 (Satt-009) to 0.67 (Satt-141).

On the basis of combined data from RAPD, ISSR and SSR marker analysis, a UPGMA dendrogram was constructed (Figure 1). The nine genotypes were separated into four clusters (I-IV). Each genotype was grouped together with one of the parental cultivars. Cluster I included I36B4, RP5 and their parental IncaSoy36. Cluster II contained I1B2-2 and CEB2 and Cluster III comprised IncaSoy1 and I1B2-3. I1B4 and CEB4 were grouped in Cluster IV. IncaSoy1 and I1B2-3 had the maximum genetic similarity (0.87), while I36B4 had the minimum (0.62) with IncaSoy36 and RP5.

SCAR marker developed from specific RAPD marker

One out of the RAPD primers used in this study, UBC-240, detected the higher level of polymorphism and amplified a band of 546 bp only present on IncaSoy36 and its progeny profiles (Figure 2). In order to obtain a SCAR marker, the 546 bp band was purified, cloned and sequenced. A BLAST search showed a 98 % of sequence identity with a region of chromosome 14 of *G. max* cv Williams 82 which E-value was 0.0. No homology with other organism was found.

Three primer pairs were designed from the 546 bp sequence and its flanking regions. The primer pairs $IntSC_{538}$-F/$IntSC_{538}$-R and $AdSC_{1161}$-F/$AdSC_{1161}$- R produced PCR fragments in all genotypes, but the $FqSC_{591}$-F/$FqSC_{591}$-R primers did not (data not shown). Therefore, the flanking region of the 546 bp-sequence of the genotypes IncaSoy36, RP5, I36B4 and CEB4 was examined in detail to discover their differences. The amplification products of $AdSC_{1161}$- F/$AdSC_{1161}$- R primers of these genotypes were cloned and sequenced. A multiple sequence alignment showed that IncaSoy36, RP5 and I36B4 share the same 546 bp-sequence and its flanking regions. Meanwhile, CEB4 has seven inserted bases into the 546 bp-sequence and point changes on base sequences within its flanking region which prevented the RADP primer hybridization (Figure 3).

New SCAR primers were designed from the sequence information gathered. The ID4G-R primer was designed to hybridize in a common sequence of the analyzed genotypes. Likewise, ID3G-F and ID1G-F matched in a specific region of IncaSoy36 and CEB4, respectively. Different combinations of these primers discriminated IncaSoy36 from its progeny (ID3G-F/ ID4G-R combination) and CEB4

Table 2. Genetic polymorphisms and unique band patterns detected by RAPD, ISSR and ISSR markers in nine soybean genotypes

Marker	Primer/ Locus	Range of amplicons (bp)	PB/TB	Unique band patterns	Polymorphisms (%)	PIC
RAPD	UBC-240	250-2000	4/8	1	50.0	0.20
	UBC-256	700-3000	2/10	0	20.0	0.08
	UBC-259	250-2000	3/8	2	37.5	0.07
	UBC-268	600-2000	3/13	1	23.1	0.09
	OPC-14	400-1700	2/12	0	16.7	0.07
ISSR	R-5	400-1600	2/6	0	33.3	0.14
	R-7	750-2000	2/5	0	40.0	0.20
	UBC-807	300-1800	3/12	0	25.0	0.12
	UBC-857	500-2000	1/7	1	14.3	0.03
SSR	Satt-005	150-180	3/3	1	100	0.57
	Satt-009	160-220	2/2	0	100	0.35
	Satt-141	150-200	3/3	0	100	0.67
	Satt-146	250-380	3/3	0	100	0.59
	Satt-373	250-280	2/2	0	100	0.49

PB/TB: Polymorphic bands/Total bands. PIC: Polymorphic Index Content.

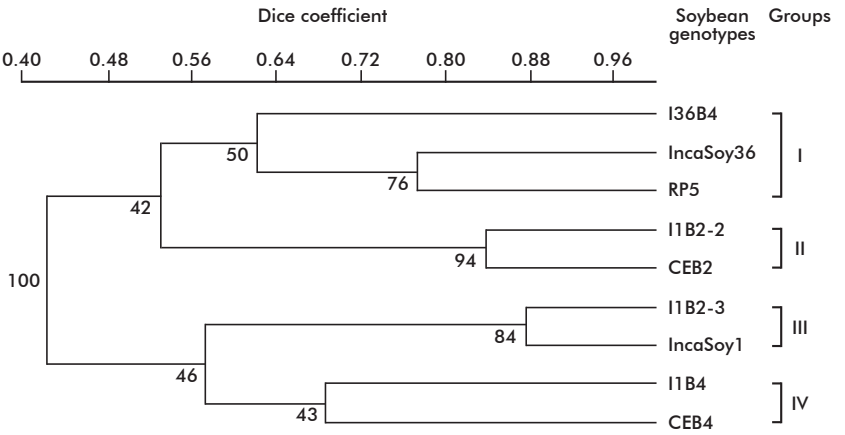

Figure 1. Genetic relationship among nine soybean genotypes revealed by UPGMA cluster analysis of Dice similarity coefficient, calculated from combined data of RAPD, SSR and ISSR markers. The numbers at the nodes indicate the bootstrapping values.

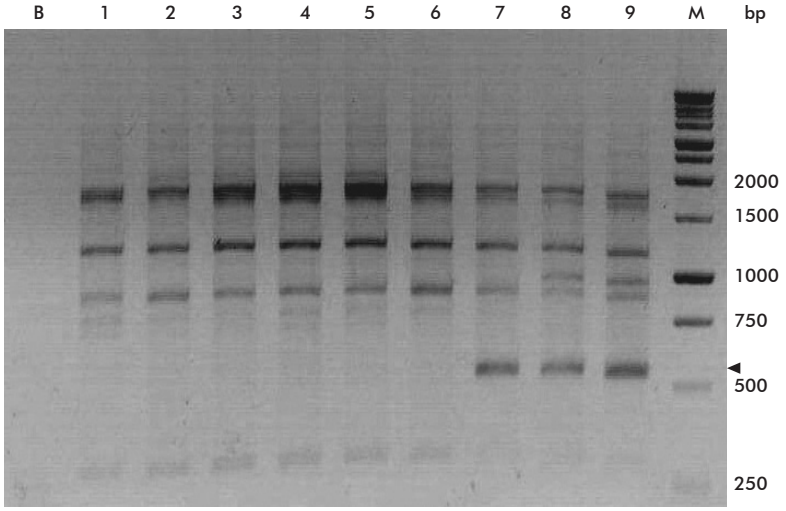

Figure 2. Amplification pattern of nine soybean genotypes assessed using the RAPD primer-UBC240. Lanes: B: Blank; 1: CEB2; 2: I1B2-2; 3: I1B2-3; 4: IncaSoy1; 5: I1B4; 6: CEB4; 7: I36B4; 8: IncaSoy36; 9: RP5. M: Molecular weight marker 1 kb DNA ladder (Promega). B: Blank, reaction mix without amplification template. The arrowhead indicates the DNA fragment converted into SCAR markers (546 bp).

(ID1G-F/ID4G-R combination) among the nine soybean genotypes evaluated.

To corroborate the SCAR primer combinations' specificity for discriminating between IncaSoy36 and CEB4 genotypes, primer sets were tested in 15 soybean cultivars. A genotype derived from IncaSoy36 (SCIGB-L1) was present in this group as sequence homology control. Ultimately, the ID3G-F/ID4G-R combination produced the expected amplicon of 648 bp in IncaSoy36 and SCIGB-L1 (Figure 4A). Meanwhile, the ID1G-F/ID4G-R pair only amplified the DNA fragment of 602 bp in CEB4 as predicted (Figure 4B).

Discussion

The information about the existing genetic variability in the elite soybean genotypes engaged in breeding programs could be useful to test genetic purity in cultivars and decide better options for future crossings. In this study, nine soybean genotypes from the Cuban soybean breeding program, seven of them expressing the event GTS 40-3-2, were analyzed using RAPD, ISSR and SSR marker systems.

The percentage of polymorphism detected by RAPD and ISSR markers ranged from 14.3 to 50.0 % (Table 2). These results were expected, considering the origin and nature of the analyzed material, because evaluated lines were derived from intraspecific crossings and selected based on the agronomic similarity with their parents. Other authors reporting higher levels of DNA polymorphism in soybean using these techniques often had studied wild populations [24, 25], cultivars with different ecological and geographic origin [8, 15, 26] or mutant lines [27, 28]. Otherwise, Baloch *et al.* [14] and Al-Saghir and Abdel-Salam [29] reported low polymorphism percentages in soybean loci by ISSR and RAPD, respectively. Contradictory reports may arise from the use of different plant material and primers, making the results not fully comparable [14]. In contrast, a 100 % of polymorphic SSR loci was detected, consistent with the previous detection of high percentage polymorphism in soybean genotypes using SSR markers [2, 11, 12, 30, 31].

Importantly, PIC values of the five assessed SSR markers (Satt-005, Satt-009, Satt-141, Satt-146 and Satt-373; PIC values greater than 0.3) were superior to those calculated from RAPD and ISSR markers. Since PIC provides an estimate of the discriminating power of a locus among different genotypes, the checked SSR loci were more informative than the RAPD and ISSR loci for distinguishing among the analyzed soybean genotypes. Similar PIC values were obtained by other researchers using these SSR markers to evaluate genetic diversity of soybean genotypes from Thailand [6], Argentina [32], India [33, 34] and Brazil [2].

On the other hand, the genetic relationship among the assessed genotypes was determined using the RAPD, ISSR and SSR data (Figure 1). In soybean, both combined and separated data from different molecular techniques used to build UPGMA dendrograms had been reported. Brick and Sivolap [15] used the information generated by AP PCR, SSR and ISSR together to show the phylogenetic relationship between soybean cultivars with different geographical and ecological origins. Likewise, based on the combined ISSR and SRAP data, Baloch *et al.* [14] clustered together

```
                    ↓  ↓
CEB4        CATAATGTTTCAGACAATTATCTTAAAAAGTCTACTATCTAAGACGATTTTTCAAAGAAC
I36B4       CATAATGTTCCAGGCAATCATCTTAAAAAGTCTACTATCTAAGACGATTTTTCAAAGAAC
RP5         CATAATGTTCCAGGCAATCATCTTAAAAAGTCTACTATCTAAGACGATTTTTCAAAGAAC
seq546bp    --------------CAATCATCTTAAAAAGTCTACTATCTAAGACGATTTTTCAAAGAAC
IncaSoy36   CATAATGTTCCAGGCAATCATCTTAAAAAGTCTACTATCTAAGACGATTTTTCAAAGAAC
                  ********** ****************************************

CEB4        TGTATTCAATCTTCAATCTTCAATGTTTAATCATGAGTTTTTATTTTGATTTCTTTCACT
I36B4       TATCTTCAA-------TCTTCAATGTTTAATCATGAGTTTTTATTTTGATTTCTTTCACT
RP5         TATCTTCAA-------TCTTCAATGTTTAATCATGAGTTTTTATTTTGATTTCTTTCACT
seq546bp    TATCTTCAA-------TCTTCAATGTTTAATCATGAGTTTTTATTTTGATTTCTTTCACT
IncaSoy36   TATCTTCAA-------TCTTCAATGTTTAATCATGAGTTTTTATTTTGATTTCTTTCACT
            * * *****       ****************************************
```

Figure 3. Fragment of multiple sequence alignment (Clustal Omega program) of the 546 bp sequence and the amplified regions of CEB4, I36B4, RP5 and IncaSoy36 genotypes using $AdSC_{1161}$-F/$AdSC_{1161}$-R primers. Underlined bases indicate the sequence of UBC240 RAPD primer used. Arrows point to the different bases present in CEB4 sequence which prevented the RADP primer hybridization.

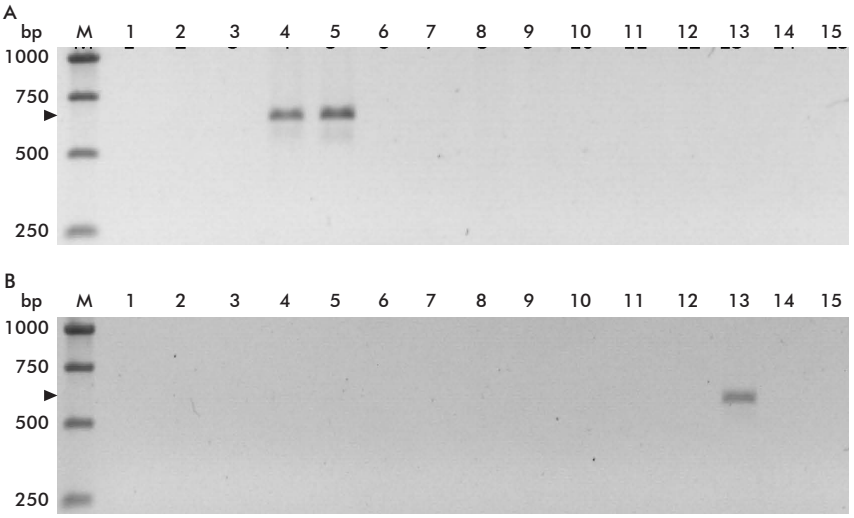

Figure 4. PCR amplification of 15 soybean genotypes using two SCAR primer combinations. A) SCAR primers ID3G-F/ID4G-R. B) SCAR primers ID1G-F/ID4G-R. Images stand for 1.5 % agarose gel electrophoresis. Lanes: M: Molecular weight 1 kb DNA ladder (Promega); 1: IncaSoy1; 2: IncaSoy27; 3: IncaSoy35; 4: IncaSoy36; 5: SCIGBL1; 6: Inifat-120; 7: Inifat-195; 8: Inifat-243; 9: Inifat-304; 10: DT84; 11: FT2; 12: DT-84; 13: CEB4; 14: CubaSoy23; and 15: Williams 82. Arrowheads indicate the respective amplicons of 648 bp in IncaSoy36 and SCIGBL1 (A) and 602 bp in CEB4 (B) genotypes.

their assessed soybean genotypes, which was in accordance with the pedigree information. On the contrary, Olsina *et al.* [32] and Hosamani *et al.* [35] found that despite they assessed two marker systems for molecular characterization of soybean genotypes, only the information of one of them was useful. A dendrogram representation from AFLP data clustered the studied soybean genotypes by owner, meanwhile no association among cultivars was observed on dendrogram from ISSR data [32]. On the other hand, the grouping of soybean genotypes based on SSR markers was found to be more informative and useful than RAPD [35]. In our case, the combined data was necessary because no informative dendrograms were obtained when trying to use them separately.

In spite of the closeness of the evaluated genotypes, it was possible to cluster each line with one of the parental using data generated by few primers of RAPD, SSR and ISSR. In fact, Groups II and III were unexpected because I1B2-2 and I1B2-3 have more

24. Jin Y, He T, Lu BR. Fine scale genetic structure in a wild soybean (*Glycine soja*) population and the implications for conservation. New Phytologist. 2003;159:513-9.

25. Zhao R, Cheng Z, Lu W, Lu B. Estimating genetic diversity and sampling strategy for a wild soybean (*Glycine soja*) population based on different molecular markers. Chinese Sci Bull. 2006;51:1219-27.

26. Panjoo M, Nazarian-Firouzabadi F, Ismaili A, Ahmadi H. Evaluation of genetic diversity among soybean (*Glycine max*) genotypes, using ISJ and RAPD molecular markers. J Plant Physiol Breed. 2014;4:55-65.

27. Hamzekhanlu MY, Izadi-Darb A, Pirvali-Beiranv N, Taher-Hallajian M, Majdabadi A. Phenotypic and molecular analysis of M7 generation of soybean mutant lines through random amplified polymorphic DNA (RAPD) marker and some morphological traits. Afr J Agric Res. 2011;6:1779-85.

morphoagronomic traits in common with the parental CEB2 than IncaSoy1 [36]. Notably, this result demonstrates the usefulness of these markers to estimate pedigree relationship among lines, because it was able to relate I1B2-3 with parent IncaSoy1 despite of phenotypic differences. In this sense, more RAPD, SSR and ISSR primers could be evaluated to cover other genome regions of these genotypes and corroborate their genetic relationship.

One out of the SSR markers analyzed in this study was able to discriminate RP5 (an inbred line derived from IncaSoy36 backcrossing) among soybean genotypes. In fact, a specific allele of the Satt-005 locus could be used as marker to distinguish RP5 (data not shown). In this sense, Rodrigues *et al.* [13] were able to discriminate two soybean varieties, one derived from the other, by means of the SSR marker Satt-115, concluding that even essentially derived varieties can be differentiated by molecular markers. Moreover, Liubov *et al.* [36] reported that Satt-005 was one of the most polymorphic locus tested in their study, and they were successful for distinguishing soybean cultivars resistant to *Fusarium* species. According to our results, Satt-005 also seems to be capable to discriminate among closely related genotypes.

Besides, markers were identified for fingerprinting these genotypes among elite soybean genotypes from the Cuban soybean breeding program, in addition to genetic diversity and relatedness determinations.

Despite few RAPD, SSR and ISSR markers were used, unique band patterns of the inbred lines I1B2-3, I1B4, I36B4, RP5 and the parents IncaSoy36 and CEB4 were obtained. The highest number of unique profiles was generated by RAPD markers (Table 2). Kumar [33] had indicated that an accurate selection of polymorphic RAPD primers enhances the efficiency of the RAPD technique to generate molecular profiles for genetic purity studies in soybean cultivars. To avoid the inconvenience of the reproducibility of the RAPD technique, a polymorphic band (546 bp) amplified by UBC-240 primer in I36B4, RP5 and IncaSoy36 was selected to convert it into a SCAR marker (Figure 2). But, the size of this fragment was too small and had high sequence homology (98 %) with a region of the sequenced soybean genome (cv Williams 82). Therefore, the SCAR primers designed from this fragment and its adjacent regions were useless because they amplified a band in all genotypes.

Likewise, the SCAR primers $FqSC_{591}$-F/$FqSC_{591}$- R designed from the flanking region of the 546-bp fragment were also inadequate due to the absence of amplified products. A new sequence analysis of the DNA region adjacent to that fragment in CEB4, RP5, I36B4 and IncaSoy36 lines showed single point changes in the RAPD primer hybridization sequence that could prevent amplification. Moreover, sequence analysis revealed seven nucleotides inserted into the 546 bp-fragment of the CEB4 (Figure 3). This information was valuable to design three new SCAR primers (ID3G-F, ID1G-F, ID4G-R) of higher specificity. The ID3G-F/ID4G-R combination was able to discriminate IncaSoy36 and SCIGB-L1 (cultivar derived from IncaSoy36) among 15 soybean cultivars (Figure 4A) and CEB4 was specifically identified with the ID1G-F/ID4G-R combination (Figure 4B). These SCAR markers would be useful to generate molecular profiles to IncaSoy36 and CEB4, respectively.

Additionally, the two SCAR markers generated here could be useful to confirm the identity of CEB4 and contribute to build a molecular profile for IncaSoy36, while requiring further research to identify new markers to accomplish fingerprinting of genotypes assessed. This last cultivar has been used successfully in experiments of tissue culture [37], genetic transformation [38] and artificial crossing [35]. Hence, a deep genetic knowledge of the IncaSoy36 genome could be valuable for soybean breeders.

In summary, the three molecular marker systems (RAPD, SSR and ISSR) employed in this study confirmed previous reports on their efficacy for conducting diversity analysis, determining pedigree relationships and genotyping in soybean [1, 5, 13, 33]. By means of only a few polymorphic markers, soybean cultivars could be effectively distinguished as it was attained for genotypes CEB4 and RP5, which were precisely identified. Data generated by these molecular markers clustered together each inbred line with one of their parental lines and the maximum coefficient of genetic similarity was nearly 0.9. Parental lines IncaSoy1, IncaSoy36, CEB2 and CEB4 showed low genetic similarity coefficient (0.4), indicative of good parental selection for planning crossings. These soybean cultivars would be employed in other crossing combinations to generate new and variable genetic material. Finally, the present study demonstrates that these markers systems could be applied by soybean breeders and the seed industry in Cuba and in other countries exhibiting breeding projects relying on local germplasm and facilities.

Acknowledgements

The authors thank Dr. Rodobaldo Ortíz from the National Institute of Agricultural Sciences of Cuba, who provided the seeds of the soybean cultivars used in this study, technical knowledge for working with them and their pedigree information.

Conflicts of interest statement

The authors declare that there are no conflicts of interest.

Received in March, 2019.
Accepted in June, 2019.

28. Mudibu J, Nkongolo K, Mehes-Smith M, Kalonji-Mbuyi A. Genetic analysis of a soybean genetic pool using ISSR marker: effect of gamma radiation on genetic variability. Int J Plant Breed Genet. 2011;5:235-45.

29. Al-Saghir MG, Abdel-Salam ASG. Genetic diversity of North American soybean (*Glycine max* L.) cultivars as revealed by RAPD markers. J Plant Sci. 2011;6:36-42.

30. Kuroda Y, Tomooka N, Kaga A, Wanigadeva S, Vaughan, DA. Genetic diversity of wild soybean (*Glycine soja* Sieb. et Zucc.) and Japanese cultivated soybeans [*G. max* (L.) Merr.] based on microsatellite (SSR) analysis and the selection of a core collection. Genet Resour Crop Evol. 2009;56:1045-55.

31. Priolli RHG, Pinheiro JB, Zucchi MI, Bajay MM, Vello NA. Genetic diversity among Brazilian soybean cultivars based on SSR loci and pedigree data. Braz Arch Biol Technol. 2010;53:519-31.

32. Olsina C, Cairo C, Pessino S. Desarrollo de una base de datos genéticos para la caracterización del germoplasma argentino de soja. Cienc Agronóm. 2012;12:23-39.

33. Bizari EH, Unêda-Trevisoli SH, Vianna VF, Meyer AS, Di Mauro AO. Genetic diversity in early-maturing soybean genotypes based on biometric and molecular parameters. J Food Agric Environ. 2014;12:259-65.

34. Kumar S. Studies on efficiency of RAPD primers in developing molecular profiles for genetic purity studies in soybean (*Glycine max* L.) cultivars. Genetika. 2014;46:681-92.

35. Hosamani J, Kumar M, Talukdar A, Lal S, Dadlani M. Molecular characterization and identification of candidate markers for seed longevity in soybean [*Glycine max* (L.) Merill]. Ind J Genet Plant Breed. 2013;73:64-71.

36. Liubov C, Ecaterina B, Nicolae B, Galina L. Genetic variation within *Glycine max* (L.) Merrill genotypes resistant to Fusariose as revealed by microsatellite markers. Romanian Biotechnol Lett. 2012;17:7270-8.

37. Soto N, Ferreira A, Delgado C, Enríquez GA. In vitro regeneration of soybean plants of the Cuban Incasoy-36 variety. Biotecnol Apl. 2013;30:34-8.

38. Soto N, Delgado C, Hernández Y, Rosabal Y, Ferreira A, Pujol M, *et al*. Efficient particle bombardment-mediated transformation of Cuban soybean (INCASoy-36) using glyphosate as a selective agent. Plant Cell Tissue Organ Cult. 2017;128:187-96.