



Genetic similarity among soybean (*Glycine max* (L) Merrill) cultivars released in Brazil using AFLP markers

Ana Lúcia V. Bonato¹, Eberson S. Calvo², Isaias O. Geraldi³ and Carlos Alberto A. Arias⁴

¹Centro Nacional de Pesquisa de Trigo, Empresa Brasileira de Pesquisa Agropecuária, Passo Fundo, RS, Brazil.

²Tropical Melhoramento & Genética Ltda., Cambé, PR, Brazil.

³Departamento de Genética, Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, SP, Brazil.

⁴Centro Nacional de Pesquisa de Soja, Empresa Brasileira de Pesquisa Agropecuária, Londrina, PR, Brazil.

Abstract

Genetic similarity among soybean genotypes was studied by applying the amplified fragment length polymorphism (AFLP) technique to 317 soybean cultivars released in Brazil from 1962 to 1998. Genetic similarity (GS) coefficients were estimated using the coefficient of Nei and Li (Nei and Li 1979), and the cultivars were clustered using the unweighted pair-group method with averages (UPGMA). The parentage coefficients of 100 cultivars released between 1984 and 1998 were calculated and correlated with the genetic similarity obtained by the markers. The genetic similarity coefficients varied from 0.17 to 0.97 ($\bar{x} = 0.61$), with 56.8% of the coefficients being above 0.60 and only 9.7% equal to or less than 0.50. The similarity coefficients have remained constant during the last three decades. Dendrogram interpretation was hindered by the large number of cultivars used, but it was possible to detect groups of cultivars formed as expected from their genealogy. Another dendrogram, composed of 63 cultivars, allowed a better interpretation of the groups. Parentage coefficients among the 100 cultivars varied from zero to one ($\bar{x} = 0.21$). However, no significant correlation ($r = 0.12$) was detected among the parentage coefficients and the AFLP genetic similarity. The results show the efficiency of AFLP markers in large scale studies of genetic similarity and are discussed in relation to soybean breeding in Brazil.

Key words: AFLP, genetic similarity, molecular markers, parentage coefficient, soybean.

Received: June 1, 2005; Accepted: May 18, 2006.

Introduction

Knowing the degree of genetic similarity among different genotypes is of fundamental importance for efficient plant breeding programs. Such information is useful for organizing a working collection, identifying heterotic groups, and selecting parents for crosses.

The parentage coefficient (Malécot, 1947) based on information regarding genotype genealogy has been used to estimate genetic similarity and to study the genetic structure of cultivated soybean germplasm. Pedigree analyses of American germplasm showed a high level of genetic relationship (Delannay *et al.*, 1983), but more recent studies have revealed that cultivars from the north and south of the United States have contrasting genetic bases (Gizlice *et al.*, 1996). These studies also showed that the genetic diversity

of North American soybean germplasm as a whole has been reduced over the last 50 years (Gizlice *et al.*, 1993). Pedigree analyses have also shown a narrow genetic base in Brazilian soybean germplasm (Vello *et al.*, 1988), although the use of Malécot's coefficients depends on the availability and precision of genealogical information.

Genetic diversity between individuals may be directly estimated by using biochemical and molecular markers, although the use of biochemical markers, such as isoenzymes, has been hindered in soybean by the low degree of polymorphism in this species (Cox *et al.*, 1985). This problem has been overcome by using molecular markers. Sneller *et al.* (1997) clearly separated elite American lines from the north and south using restriction fragment length polymorphism (RFLP) markers. This technique has also been used to study exotic soybean germplasm and it has allowed the identification of different gene pools (Kisha *et al.*, 1998). Similar studies have been carried out using other types of molecular markers, such as RAPD markers

(Abdelnoor *et al.*, 1995; Brown-Guedira *et al.*, 2000), simple sequence repeat (SSR) markers (microsatellites) (Dewan and Cregan, 1997) and amplified fragment length polymorphism (AFLP) markers (Zhu *et al.*, 1999; Ude *et al.*, 2003).

A comparative study on the performance of different types of markers in soybean genetic analysis showed that microsatellite markers have a greater degree of polymorphism and, thus, better discrimination between genotypes. However, AFLP markers have greater multiplex efficiency (*i.e.* a large number of loci can be simultaneously analyzed in a gel) and are considered an efficient tool for distinguishing highly related genotypes (Powell *et al.*, 1996). As a result of such characteristics, this technique continues to be used in current genetic diversity studies (Oliveira *et al.*, 2004).

Although genetic diversity studies using molecular markers have been carried out with various types of markers and diverse genotypes, as expected, they confirmed the presence of a larger amount of genetic diversity in exotic germplasm (Zhu *et al.*, 1999). However, genetic diversity estimates between soybean cultivars obtained using both the parentage coefficient and molecular markers, have shown variable results. The magnitude of the correlation between these two estimates was 0.54 to 0.91 in RFLP studies (Manjarrez-Sandoval *et al.*, 1997; Kisha *et al.*, 1998) but Helms *et al.* (1997) obtained no apparent relationship between the two types of estimates when using RAPD markers. Abdelnoor *et al.* (1995) reported some cases of discrepancy in the genetic distance between Brazilian cultivars analyzed by RAPD and by pedigree, in spite of overall agreement in the data. Since molecular marker measurements are a direct measure of the genetic distance it is possible that these discrepancies reflect errors related to pedigree assessments.

Analysis of Brazilian soybean germplasm by molecular markers has been reported. Abdelnoor *et al.* (1995), assessing the molecular marker approach using the RAPD technique to measure the genetic diversity of 30 Brazilian cultivars, found five different subgroups. However, the application of these results in a breeding program was hindered by the reduced number of genotypes used in the study. Recently, alleles of 12 microsatellite loci of 186 Brazilian soybean cultivars were used to morphologically distinguish similar groups and their use allowed the determination of 184 profiles for all cultivars (Priolli *et al.*, 2002). Our present study was carried out to investigate the use of AFLP markers in the genetic similarity analysis of 317 soybean cultivars released in Brazil from 1962 to 1998.

Material and Methods

Genetic material

We investigated 317 soybean cultivars released in Brazil between 1962 and 1998 (Table 1). The genetic mate-

rial was obtained from the Active Germplasm Bank of the National Soybean Research Center of the Brazilian Corporation for Agriculture Research (Embrapa Soja), Londrina, PR, Brazil.

DNA extraction and quantification

For each cultivar, 30 leaves were collected from different greenhouse-grown plants and immediately frozen in liquid nitrogen and stored at -80 °C for subsequent genomic DNA extraction according to Saghai-Marooof *et al.* (1984). The DNA concentration was estimated in 0.8% agarose gels by visual comparison of DNA band intensity with undigested lambda DNA standards and the DNA samples were then diluted to 30 ng μL^{-1} and stored at -20 °C until needed.

AFLP genotyping

All AFLP analyses were made with the AFLP Analyses Kit I (Gibco-LifeTechnologies, Rockville, MD, USA) essentially as described in the kit manual. All the amplifications were conducted in a Perkin-Elmer Gene Amp 9600 thermocycler (Perkin-Elmer Corp., Norwalk, CT, USA). The AFLP products were fractionated using 5% (w/v) polyacrylamide sequencing gels, dried and the autoradiography was performed by exposing Kodak Bio Max MR-2 film (Eastman Kodak Co., Rochester NY, USA). Six *EcoRI*/*MseI* primer combinations (E-AAC/M-CAT, E-AAC/M-CTA, E-AAC/M-CTC, E-AAC/M-CTG, E-AAG/M-CTT, and E-ACT/M-CAT) were selected based on the previously reported polymorphism rate (Bonato *et al.*, 2006).

Data analysis

The DNA bands were scored as 1 (presence) and 0 (absence) based on visual observation and the results entered into an Excel[®]. Genetic similarity (GS) was estimated for all genotype pairs using the equation $GS_{i,j} = 2N_{i,j}/(2N_{i,j} + N_i + N_j)$ (Nei and Li, 1979), where $GS_{i,j}$ represents the similarity estimate between the genotypes *i* and *j*, based on the AFLP data, $N_{i,j}$ is the total number of bands common to *i* and *j*, and N_i and N_j correspond to the number of bands found in genotypes *i* and *j*. The matrix generated with the GS estimates was used to cluster the genotypes in a dendrogram obtained by the unweighted pair group method using arithmetic averages (UPGMA) (Rohlf, 1997). Cophenetic correlation between GS-matrix and dendrogram cophenetic values was estimated to validate the dendrogram in relation to the original similarity estimates and the binary data matrix analyzed using the NTSYS 2.0 software (Rohlf, 1997). Bootstrap analysis (Tivang *et al.*, 1994) was used to verify if the number of markers was sufficient to characterize the cultivars for genetic similarity. The procedures for this re-sampling have been described by Barroso *et al.* (2003). Cophenetic correlation was obtained by Bionumeric Analyses (Rohlf, 1997) to express the consistency of a cluster.

Table 1 - Soybean cultivars released in Brazil from 1962 to 1998 and used for the AFLP analysis.

N.	Cultivar	N.	Cultivar	N.	Cultivar
1	BR-36*	55	FT-11 (Alvorada)*	109	CAC-1
2	BR-37	56	Cep 12 (Cambará)*	110	IAC-13
3	GO/BR-26 (Tocantins)	57	FT-6 (Veneza)	111	RS 6 (Guassupi)*
4	MS/BR-19 (Pequi)	58	FT-7 (Tarobá)	112	Coker 136
5	BR-30*	59	FT-8 (Araucária)*	113	UFV-2
6	MS/BR-34 (EMPAER 10)*	60	FT-2	114	IAC-3
7	BR-23*	61	FT-3	115	São Carlos*
8	BR-38	62	FT-1	116	Ivorá
9	BR-28 (Seridó)	63	Emgopa-308*	117	IAC-7
10	MS/BR-20 (Ipê)	64	BR/Emgopa-312 (Potiguar)*	118	IPAGRO-21
11	Emgopa-301	65	Emgopa-311	119	IAC-12
12	MS/BR-21 (Buriti)*	66	Emgopa-313 (Anhanguera)	120	Embrapa 20 (Doko RC)*
13	MS/BR-17 (São Gabriel)	67	Emgopa-309	121	IAS 4
14	BA/BR-31	68	Emgopa-310	122	UFV-4
15	MS/BR-18 (Guavira)*	69	FT-15	123	Ocepar 2 (Iapó)
16	MG/BR-22 (Garimpo)*	70	FT-Manaca	124	Ocepar 7 (Brilhante)*
17	BR-5	71	FT-16	125	Paraná
18	BR-14 (Modelo)*	72	FT-10 (Princesa)	126	Ocepar 14*
19	BR-10 (Teresina)	73	FT-Eureka	127	Vila Rica
20	BR-4	74	FT-Guaíra*	128	Ocepar 11*
21	BR-12	75	FT-14 (Piracema)	129	UFV-9 (Sucupira)
22	BR-24*	76	FT-Jatoba	130	Ocepar 10*
23	BR-15 (Mato Grosso)*	77	FT-Canarana*	131	IAS 2
24	BR-1	78	FT-Estrela	132	Embrapa 3*
25	BR-13 (Maravilha)*	79	FT-9 (Inae)	133	MG/BR-42 (Kage)*
26	BR-2 (Vagem clara)	80	FT-18 (Xavante)	134	Ocepar 5 (Piquiri)
27	BR-11 (Carajás)	81	FT-17 (Bandeirantes)*	135	Lancer
28	BR-6 (Nova Bragg)	82	Cristalina	136	FT-440
29	Bossier	83	FT-19 (Macacha)	137	Pampeira*
30	BR-3	84	FT-Abyara*	138	Embrapa 1 (IAS 5 RC)
31	Andrews	85	FT-13 (Aliança)	139	Missões
32	BR-8 (Pelotas)	86	FT-20 (Jaú)	140	RS 5 (Esmeralda)*
33	Emgopa-305 (Caraíba)*	87	FT-Bahia*	141	Prata*
34	BR-29 (Londrina)	88	CAC/BR-43*	142	IAC-16
35	BR-16	89	IAC-5	143	Ocepar 9*
36	BR-9 (Savana)	90	FT-Maracaju	144	Embrapa 19
37	SP/BR-41	91	IAC-4	145	Embrapa 2
38	Doko	92	Davis	146	IPAGRO-20
39	Dourados	93	IAC-2	147	Nova IAC-7*
40	BR-7	94	IAC-6	148	UFV-8 (Monte Rico)
41	Emgopa-302	95	FT-Seriema	149	IAC-11*
42	Cep 20 (Guajuvira)*	96	IAS 1	150	Paranaíba
43	MS/BR-39 (Chapadão)*	97	IAC-Foscarin 31	151	Hood 75
44	Emgopa-304 (Campeira)	98	Pérola	152	IAS 5 (Vagem Escura)
45	Campos Gerais	99	Hardee	153	Cajeme
46	GO/BR-33 (Javaes)	100	BR-32*	154	UFV-14
47	Cep 10	101	UFV-5	155	Santa Rosa
48	Cobb	102	UFV-3	156	UFV/ITM-1
49	Bragg	103	UFV-Araguaia	157	Viçoja
50	Cep 16 (Timbó)*	104	IAC-100	158	UFV-15 (Uberlândia)*
51	FT-12 (Nissei)	105	Industrial	159	IAC-9
52	BR-40 (Itiquira)*	106	Mineira	160	Ocepar 3 (Primavera)
53	FT-4	107	IAS 3 (Delta)	161	UFV-10 (Uberaba)*
54	FT-5 (Formosa)	108	IAC-14	162	IAC-8

Table 1 (cont.)

N.	Cultivar	N.	Cultivar	N.	Cultivar
163	Embrapa 5	218	Soberana	273	IAC-19
164	MT/BR-45 (Paiaguás)*	219	Ocepar 15 (Paracatu)	274	BRAS-60 (Celeste)*
165	Paranagoiana	220	Embrapa 62*	275	BRS 157
166	Ocepar 8*	221	KI-S 702	276	BRS 158 (Milena)*
167	IAC-1	222	Nobre*	277	MA/BR-64 (Parnaíba)
168	Pelicano	223	Embrapa 58*	278	MG/BR-56 (Confiança)*
169	Sertaneja*	224	FT-Saray	279	MA/BR-65 (Sambaíba)
170	Ivaí	225	Cep 26 (Umbú)	280	BRS 137*
171	Planalto	226	FT- Morena*	281	GO/BRS-160 (Goiatuba)*
172	Hill	227	Embrapa 46	282	Emgopa-316 (Rio Verde)
173	UFV-6 (Rio Doce)*	228	Ocepar 16	283	Embrapa 64 (Ponta Porã)
174	RS 7 (Jacui)*	229	Embrapa 59	284	BRS 133*
175	IAC-10	230	MT/BR-50 (Parecis)	285	Ocepar 19 (Cotia)
176	J-200	231	MT/BR-47(Canário)	286	GO/BRS-161 (Catalão)
177	Tiaraju	232	FT- Iramaia	287	MG/BR-58 (Segurança)
178	IAS 5 (Vagem Clara)	233	FT- Líder*	288	BRS 134*
179	IAC-15	234	MT/BR-52 (Curió)	289	Coodetec-203
180	UFV-7 (Juparana)*	235	Embrapa 26	290	MS/BR-57 (Lambari)*
181	União	236	Ocepar 12*	291	MS/BRS-168 (Piapara)
182	Ocepar 13	237	Embrapa 32 (Itaquí)*	292	MT/BR-55 (Uirapuru)
183	UFV-1	238	Ocepar 6*	293	MS/BRS-171 (C. Grande)*
184	Teresina RCH	239	Rainha	294	Cariri RCH
185	Ocepar 4 (Iguaçu)	240	MS/BRS-172 (Tuiuiu)*	295	MS/BRS-169 (Bacuri)
186	São Luiz	241	Embrapa 60*	296	BRS 136
187	BR/IAC-21	242	KI-S 801	297	Coodetec-202*
188	Vitória	243	BRS 154	298	BRS 135*
189	Tropical	244	MS/BRS-167 (Carandá)*	299	GO/BRS-162 (Bela Vista)*
190	Embrapa 48*	245	MG/BRS-66 (Liderança)*	300	BRS 156
191	Embrapa 9 (Bays)*	246	FT-45263	301	Serido RCH
192	Embrapa 47	247	Embrapa 30	302	Emgopa-315 (Rio Vermelho)*
193	MT/BR-53 (Tucano)	248	Ocepar 17	303	MT/BRS- 63 (Pintado)
194	Timbira	249	BRS 153	304	Emgopa-307
195	BR-35 (Rio Balsas)*	250	MS/BRS-173 (Piraputanga)	305	KI-S 602
196	IAC-18	251	FT-Cristalina RCH*	306	FT-Cometa
197	Invicta*	252	KI-S 602 RCH	307	Emgopa-306 (Chapada)
198	Década	253	Embrapa 33 (Cariri RC)*	308	MS/BR-61 (Suruibí)*
199	FT-102	254	BRS 155*	309	BR-27 (Cariri)
200	FT-100	255	Embrapa 31 (Mina)	310	KI-S 601
201	SPS-1	256	FT-106	311	MG/BR-54 (Renascença)
202	BRS 65 (Itapoty)*	257	Coodetec-201	312	MT/BR-49 (Pioneira)*
203	FT-103	258	MS/BR-59 (Mandi)	313	GO/BRS-163 (Jataí)
204	FT-104	259	BR/Emgopa-314*	314	MG/BR-46 (Conquista)*
205	Embrapa 61*	260	FT-2002	315	Emgopa-303*
206	Embrapa 25*	261	Embrapa 34 (Teresina RC)	316	GO/BR-25 (Aruanã)*
207	MT/BR-51 (Xingu)	262	BRS 138	317	Numbaira
208	Cristalina Antiga	263	MT/BRS-159 (Crixás)		
209	FT- Cristal*	264	MG/BRS-68 (Vencedora)		
210	IAC-17	265	Pirapó 78		
211	BRS 66*	266	BRS 132		
212	FT-101*	267	IAC-20		
213	RS 9 (Itaubá)*	268	Embrapa 63 (Mirador)*		
214	Embrapa 4 (BR-4 RC)	269	Ocepar 18		
215	Fepagro-RS 10*	270	MS/BRS-166 (Apaiani)		
216	FT- Cristalina*	271	BRAS-62 (Carla)		
217	MG/BR-48 (Garimpo RCH)*	272	MS/BRS-170 (Taquari)		

*Cultivars used for parentage coefficient analysis.

Parentage coefficient

The parentage coefficient, f , (Malécot, 1947) estimates among 100 soybean cultivars released from 1984 to 1998 were based on their respective genealogies (see Vello *et al.*, 1988 for earlier results) and obtained using the PARENT software program (CIAGRI - ESALQ/USP). These cultivars are indicated with an asterisk (*) in Table 1. One hundred cultivars were used because this is the maximum number of genotypes that can be assessed by the PARENT software. The estimates were later used to calculate the correlation between the parentage coefficient and the genetic similarity between the respective pairs of cultivars as measured by AFLP.

Results and Discussion

AFLP analysis

The six primer combinations used to analyze the 317 cultivars generated 394 bands, 78 (19.8%) of those were found to be polymorphic among genotypes (Table 2). The average number of polymorphic markers per primer combination was 13, varying from six to 25. The *EcoRI*-*AAC/MseI*-*CTC* and *EcoRI*-*AAC/MseI*-*CAT* markers generated the highest levels of polymorphism, 18 (34.6%) and 25 (30.1%) respectively.

Maughan *et al.* (1996) detected 274 (36%) AFLP polymorphic bands among *Glycine max* and *Glycine soja* accessions and an average of 18 polymorphic bands per primer combination, considering accessions of both species, with most (31%) of the polymorphism occurring in *Glycine soja* accession and only 17% in *Glycine max*.

In our study, although we considered only adapted *Glycine max* cultivars the level of polymorphism found (19.8%) clearly indicated the genetic similarity among the genotypes. Our results confirm previous findings demonstrating that AFLP is a molecular technique that detects polymorphism in multiple loci, generating a vast number of reproducible markers in a short period of time (Maheswaran *et al.*, 1997) and is a powerful tool for screening highly related genotypes (Powell *et al.*, 1996).

Table 2 - *EcoRI/MseI* AFLP analysis primer combinations and number of bands.

Primers		Number of bands	
<i>EcoRI</i>	<i>MseI</i>	Total	Polymorphics
AAC	CTC	52	18
AAC	CTA	68	11
AAC	CTG	55	8
AAC	CAT	83	25
AAG	CTT	86	10
ACT	CAT	50	6
Total		394	78

According to Zhu *et al.* (1999), the most polymorphic primer combinations were *EcoRI*-*AAC/MseI*-*CTC* (53%) and *EcoRI*-*AAC/MseI*-*CAT* (50%). These same authors reported a greater polymorphism frequency compared to that found by us, probably because Zhu *et al.* (1999) were dealing with adapted and non-adapted *Glycine max* and *Glycine soja* accessions. Their results, however, are in line with those obtained in our study and confirm that these two primer combinations are highly informative for analysis of Brazilian soybean germplasm.

AFLP estimates of genetic similarity

We constructed a similarity coefficient matrix from the genetic similarity calculations for the 317 genotypes. Figure 1 shows the frequency distribution of the coefficients, the average coefficient among all genotypes was 0.61, ranging from 0.17 for the Nobre and Bossier cultivars to 0.97 for the FT-Cristal and FT Cristalina cultivars. We also found that 56.8% of the estimated coefficients had values greater than 0.60, reflecting the high degree of genetic similarity among the cultivars used in this study. However, 9.7% of the coefficients were equal to or less than 0.50, and can be exploited for divergent parent selection.

High genetic similarity among Brazilian soybean cultivars was also detected by Abdelnoor *et al.* (1995), who used RAPD analysis and obtained a mean GS coefficient of 0.82 with a range of 0.69 to 1.00. The most divergent cultivars were Tropical and UFV-6, whereas the most similar cultivars were Ocepar-9 and Paranagoiana. Other studies using AFLP have also shown high similarity among adapted *Glycine max* genotypes. For example, Maughan *et al.* (1996) found similarity values ranging from 0.74 to 1.00. Zhu *et al.* (1999), although observing high similarity coefficients between *Glycine max* and *Glycine soja* accessions (0.60 to 0.94), emphasized the greater similarity of *Glycine max* cultivars. However, Priolli *et al.* (2002) used SSR markers and found GS values ranging from 0.18 to 0.59 in a group of 186 Brazilian soybean cultivars, this

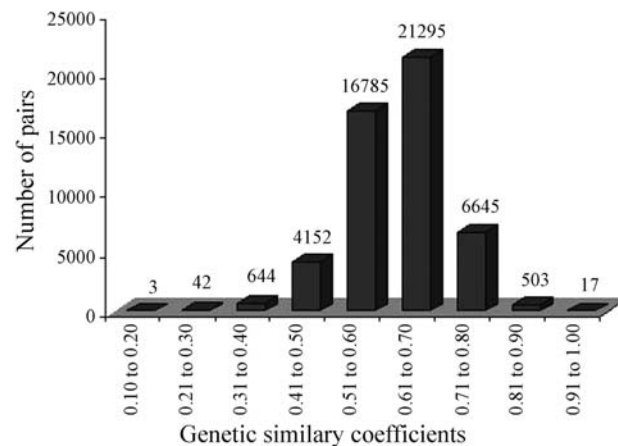


Figure 1 - Frequency distribution of 50,086 AFLP genetic similarity coefficients among 317 soybean cultivars.

level of genetic similarity being lower than that found in our study.

One of the concerns about GS estimates is the number of markers required for sampling the genome. We obtained an average coefficient of variation (CV) of 7.7% by bootstrap analysis with the 78 markers used in our AFLP analysis (Figure 2) and the CV decreased as the sample size increased, indicating that the accuracy of genetic similarity estimates increases if the number of polymorphic loci is increased. Logarithmic transformation of mean CV and sample size established a linear relationship between the two variables and, consequently, the regression equation showed that with 100 polymorphic loci a mean CV of 6.7% would be obtained (Figure 3), an insignificant increase in the precision of the estimates.

Pejic *et al.* (1998) reported a comparative analysis of genetic similarity in maize, using RFLP, RAPD, AFLP, and SSR markers, and found coefficients of variation of 5% and 10% for the four types of markers with 150 polymorphic bands. These authors consider this to be a sufficient number of markers to estimate similarity with high accuracy. Moser and Lee (1994) suggested that species with a

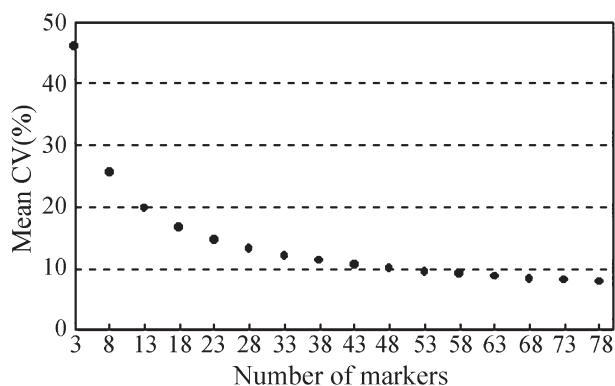


Figure 2 - Mean coefficient of variation of the genetic similarity coefficients among 317 soybean cultivars, according to the number of re-sampled AFLP loci, using the bootstrap technique.

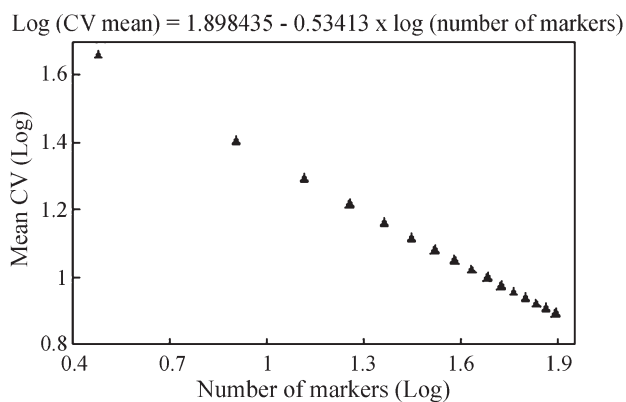


Figure 3 - Regression equation between the natural logarithm of the number of re-sampled AFLP loci and the natural logarithm of the mean coefficient of genetic similarity coefficients using the bootstrap technique.

polymorphism index lower than that found in maize have a lower standard error in these estimates, which means that less loci need be assessed for the same level of precision. In the present study, the results obtained by bootstrap analysis show that the number of AFLP markers was sufficient to characterize soybean cultivars for their genetic similarity.

Genetic diversity of 317 Brazilian soybean cultivars released between 1962 and 1998

The impact of Brazilian genetic breeding programs on soybean genetic diversity over 36 years was investigated using the GS coefficients after grouping the cultivars into three periods according to their release to the market. In the first period we considered 48 cultivars released between 1962 and 1980, in the second period 122 cultivars released between 1981 and 1990, and in the third period 121 cultivars released between 1991 and 1998. We disregarded 26 cultivars because their release date was not available. Table 3 shows that the mean values remained practically constant for the three periods, indicating that the genetic diversity of cultivars developed in the Brazilian breeding programs maintained a similar level of genetic similarity throughout the years. This contrasts with the findings of Kisha *et al.* (1998) who assessed genetic diversity among different USA soybean gene pool using RFLP and found that the diversity among elite cultivars, as compared to ancestral genotypes, was declining over time as a consequence of breeding effects.

In spite of the narrow genetic base found in soybean cultivated in Brazil and the relatively high similarity among cultivars, substantial genetic gains in terms of productivity have been obtained for grain yield and other traits. Similar facts have occurred in the USA breeding programs (Hiro-moto and Vello, 1986). In a soybean breeding program in the Brazilian state of Paraná Toledo *et al.* (1990) estimated mean annual increased productivity due to genetic gains from 1981 to 1986 of 1.8% for an early maturity group and 1.3% for an intermediate maturity group of cultivars. In a study of soybean cultivars widely grown in the southern Brazilian state of Rio Grande do Sul during different periods Rubin and Santos (1996) concluded that there has been a mean genetic gain of 19 kg ha⁻¹ y⁻¹ over the last 40 years, equivalent to 1.1% per year. Rubin and Santos also noted that these gains have been decreasing over the years as a re-

Table 3 - Mean, maximum and minimum genetic similarity coefficients (GS) for soybean cultivars released during different periods in Brazil.

Coefficient	Periods of release		
	1962 to 1980	1981 to 1990	1991 to 1998
Mean	0.62	0.62	0.63
Maximum	0.86	0.97	0.92
Minimum	0.23	0.34	0.23
N. of cultivars	48	122	121

sult of using the same basic germplasm during hybridization. However, this is not necessarily the case, as revealed by the recent releases of new cultivars and the fact that there have been considerable improvements in the performance of agronomic traits due to the correction of defects controlled by qualitative traits and an improvement in grain yield has also been reported in several Brazilian states. For example, in Rio Grande do Sul the BRS 153 cultivar outperformed (in terms of grain yield) the control cultivar BR-16 by 14%, while the cultivar BRS-133 outperformed the same control by 8.5% in the state of Paraná, while in the Brazilian state of Mato Grosso do Sul the cultivar MS/BRS-171 (Campo Grande) outperformed the control cultivar FT-Cristalina by 20% while the cultivar MT/BR-50 (Parecis) outperformed cultivar MT/BR-45 (Paiaguás) by 5% (Congresso Brasileiro de Soja, 1999).

UPGMA grouping of the 317 soybean cultivars based on GS estimates

The GS estimates for the 317 cultivars were used to generate a UPGMA dendrogram (Figure 4). In spite of the large numbers hindering the assessment of similarities between each pair of cultivars, the dendrogram allowed the detection of groups of cultivars with expected genetic similarity corresponding to their genealogy. These results support the previously noted efficiency of AFLP markers for estimating genetic similarity among soybean genotypes. Knowing the genealogy of cultivars was essential for the interpretation of the dendrogram, a fact in line with other studies (Abdelnoor *et al.*, 1995; Diwan and Cregan, 1997; Priolli *et al.*, 2002).

As expected, the genetic similarity (GS) coefficients and the dendrogram showed that cultivars derived from natural mutation had high similarity coefficients. The Paranagoiana and Ocepar 9-SS1 cultivars, mutants of the Paraná cultivar, had similarity coefficients of 0.95. The São Carlos mutant cultivar had a genetic similarity coefficient of 0.83 with respect to the original Davis cultivar, while the UFV-1 mutant cultivar shared a coefficient of 0.67 with the original Viçoja cultivar. The magnitude difference between genetic similarity coefficients observed for the different mutants seems to indicate either that not all mutations giving rise to the mutant genotypes were single point mutations or they that these cultivars may not actually be mutants. However, the Paranagoiana and Ocepar 9-SS1 cultivars are known to be mutants of the Paraná cultivar, this having been demonstrated by the electrophoresis and isoenzyme studies of Derbyshire *et al.* (1990).

The GS estimates for cultivars derived from other cultivars also showed many cases of similarity. Examples were found in the FT-Cristal and FT-Bahia cultivars, which had GS coefficients of 0.97 and 0.87, respectively, compared with the FT-Cristalina cultivar from which they were selected. Similarly, the Ocepar-8 cultivar showed high similarity (0.90), when compared to the Paraná cultivar from

which it was selected. However, in some instances, low similarity was found between putatively related cultivars, as in the case of the IAS 5 cultivar, which had been separated into two types according to pod color (dark pod IAS 5 and pale pod IAS 5) by researchers at the Embrapa National Soybean Research Center (Personal Communication). Our AFLP analysis indicated that these two cultivars must diverge in other genes besides those defining pod color because the AFLP GS coefficient between them was only 0.63. However, we also found that the FT-2 cultivar, derived from a selection within the IAS 5 cultivar, was closer (GS = 0.78) to the pale pod IAS 5 cultivar than to the dark pod IAS 5 cultivar (GS = 0.65). These data indicate that the AFLP technique is highly discriminating for cultivar differentiation even among closely related genotypes. Diwan and Cregan (1997) analyzed soybean genotypes using 20 microsatellite markers but were unable to separate the Ilini genotype from the its ancestral AK Harrow genotype.

Among cultivars derived from the same cross-sister cultivars, there were several cases of agreement between GS coefficients and their allocation in the same group. Cultivars MT/BR-50 (Parecis), MT/BR-51 (Xingu), MT/BR-52 (Curió) and MT/BR-53 (Tucano), derived from the BR 83-9520-1 (2) x FT-Estrela cross, had genetic similarity coefficients greater than 0.80. There was a similar situation for the UFV-2, UFV-3, UFV-4 and UFV-Araguaia cultivars, derived from the Hardee x IAC-2 cross, whose coefficients were greater than 0.83. Among the FT-5, FT-10, FT-14 and FT-15 cultivars, derived from the FT-9510 x Santa Ana cross, the GS coefficients varied from 0.71 to 0.93. Cultivars Embrapa 59, Embrapa 60, Embrapa 61, and BRS 66, all derived from the FT-Abyara x BR 83-147 cross had GS coefficients greater than 0.75. Additionally, the BRS 133, BRS 135, BRS 158, MS/BR-57 (Lambari) and MS/BRS-171 (Campo Grande) cultivars, also selected from the FT-Abyara x BR 83-147 cross, had GS coefficients lower than 0.65 as compared to their Embrapa 59, Embrapa 60, Embrapa 61, and BRS 66 sister cultivars. We attributed these differences in similarity among the sister cultivars to selection effects. Abdelnoor *et al.* (1995) also found similarities at several levels between cultivars derived from the same cross.

Cultivars developed from backcrosses had variable GS coefficients, as compared to their recurrent parents. For example, BR-6 (Nova Bragg) and BR-13 (Maravilha) cultivars were obtained from backcrosses (three to Nova Bragg and four to Maravilha) with the Bragg cultivar and had genetic similarity coefficients greater than 0.75 in relation to the Bragg cultivar. The Embrapa 1 cultivar, obtained from six backcrosses to the IAS 5 cultivar, had a genetic similarity coefficient of 0.68 when compared with the dark pod IAS 5 cultivar and 0.54 when compared with the pale pod IAS 5. The Embrapa 4 cultivar, derived from six backcrosses to the BR-4 cultivar, had a similarity coefficient of 0.61 only with BR-4 cultivar. The lower than expected ge-

netic similarity between backcross progeny and respective recurrent parents found in our study may be explained by the work of Muehlbauer *et al.* (1988), who suggested that these types of effects are caused by the introgression of other markers in the same linkage groups as the transferred gene. Another possibility may be the lower selection pressure applied to recover the genetic characteristics of the recurrent parent.

Dendrogram analysis did not allow the separation of cultivars into groups based on the geographic distribution of their release sites or recommended planting sites, although the RFLP analyses of Kisha *et al.* (1998) showed a clear separation between soybean cultivars from the north and south of the USA and greater similarity among the genotypes from the south. The results of Kisha *et al.* (1998) were probably due to the fact that the cultivars from each region in the USA were derived from distinct ancestral groups, whereas in Brazil there are no such expected differences because the cultivars developed at different locations were derived from the same ancestral group (Romeu Kiihl - Personal Communication).

In most of the cases discussed above not only was the similarity indicated by the GS coefficients greater than that displayed in the dendrogram but some cultivars with high GS coefficients were placed in different groups in the dendrogram. For example, this occurred with the Pirapó 78 and Nova IAC-7 cultivars, which, even with coefficients of 0.82, were allocated to distinct groups of the respective parental cultivars Paraná and IAC-7. Possible causes for such discrepancies could be the large number of very closely genetically related genotypes analyzed and the low cophenetic correlation obtained for the original coefficients compared to that estimated by grouping ($r = 0.60$). Powell *et al.* (1996) obtained a cophenetic correlation of 0.78 among *Glycine max* accessions using AFLP markers, but the value rose to 0.96 when accessions of this species were considered together with *Glycine soja* accessions. In Powell's study, the lower cophenetic correlation observed among *Glycine max* accessions may have been due to the greater genetic similarity of the genotypes of this species. This explanation may be extended to the results of our study in which the cultivars had very similar GS coefficients which may have interfered when the estimates were grouped in the dendrogram.

Grouping of the 62 soybean cultivars based on GS estimates

To simplify the interpretation of the dendrogram, a new dendrogram was constructed using only 62 of the 317 cultivars (Figure 5). The cultivars used included those with ambiguous results regarding their grouping in the previous 317-cultivar dendrogram, as well as those with similarity coefficients of different magnitudes. Cophenetic correlation is also a parameter that expresses the consistency of a cluster by calculating the correlation between the den-

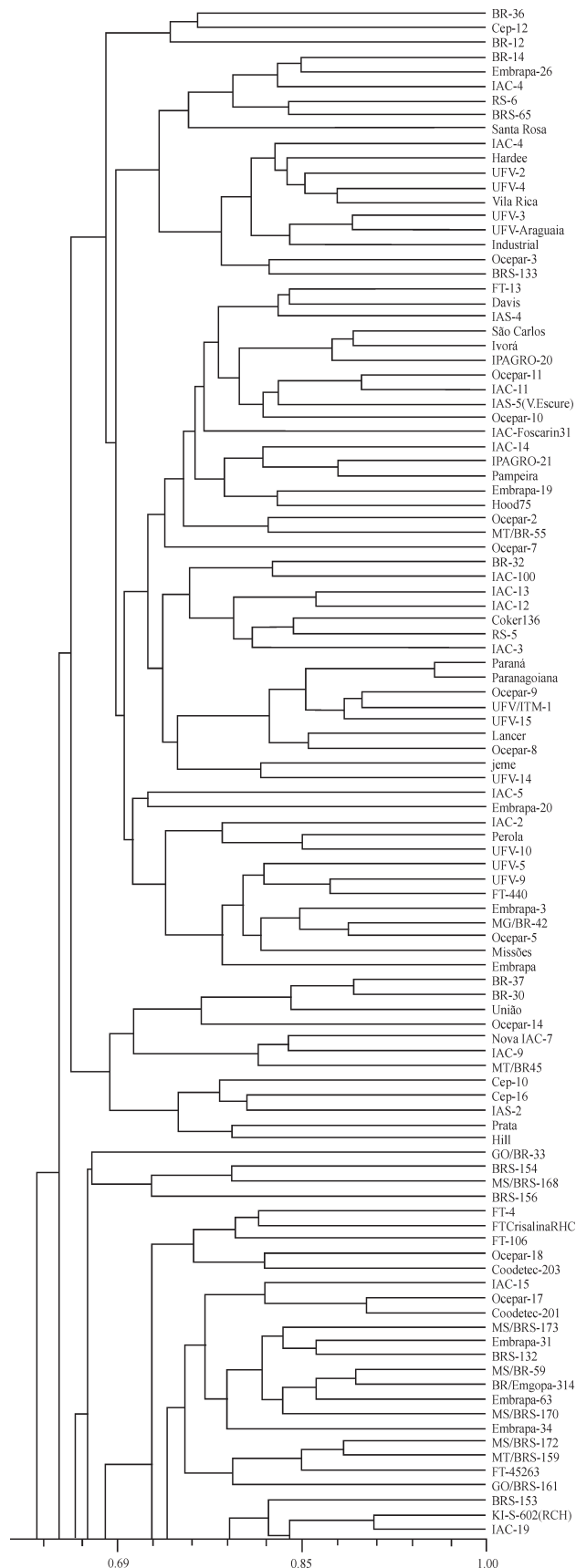


Figure 4a - UPGMA dendrogram based on AFLP similarity coefficients of 317 soybean cultivars.

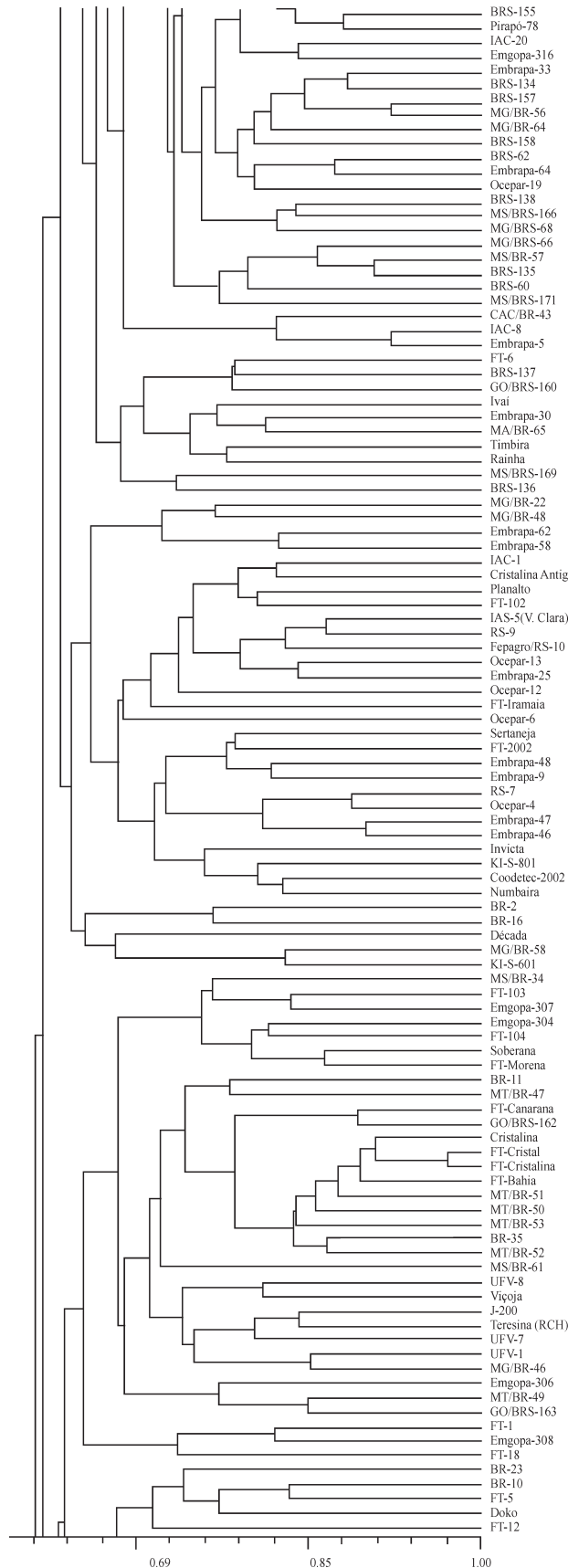


Figure 4b - UPGMA dendrogram based on AFLP similarity coefficients of 317 soybean cultivars (cont.)

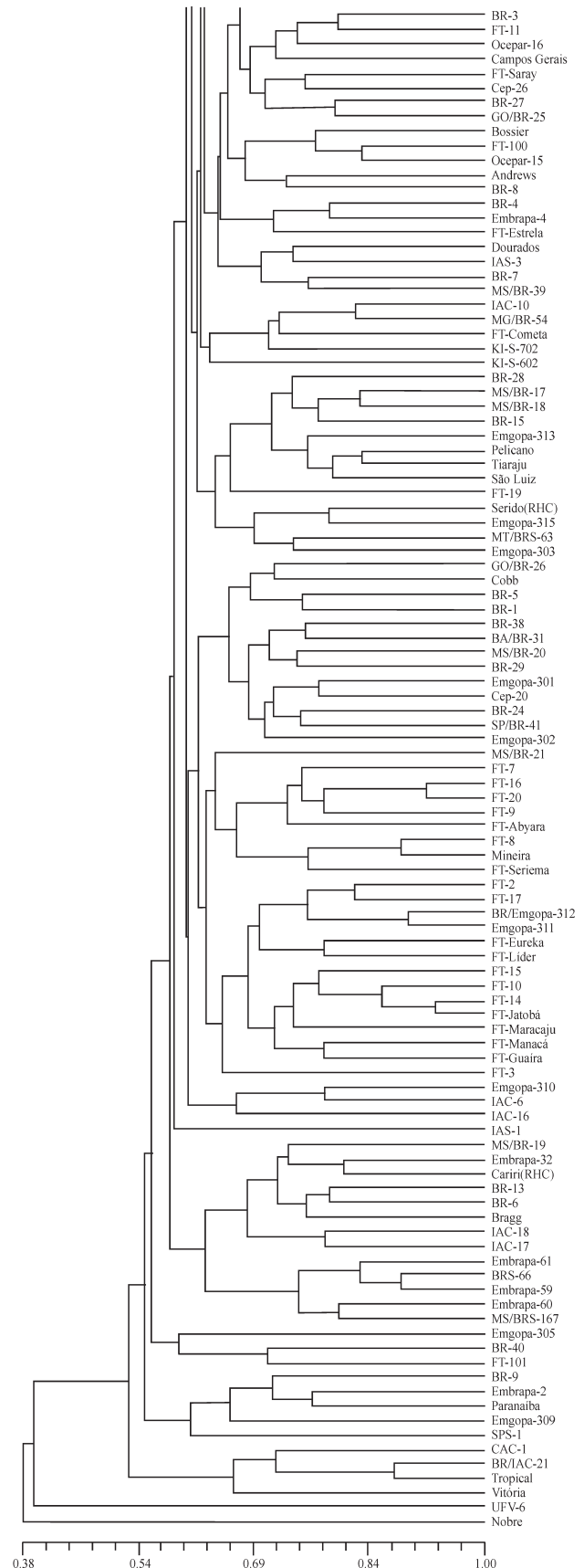


Figure 4c - UPGMA dendrogram based on AFLP similarity coefficients of 317 soybean cultivars (cont.)

dendrogram-derived similarities and the matrix similarities. In BioNumerics, the value is calculated for each cluster (branch) thus estimating the faithfulness of each subcluster of the dendrogram. The cophenetic correlation was obtained for the whole dendrogram from the cophenetic correlation at the roots (Figure 5). In the 63-cultivar dendrogram it was easier to visualize groups and there was a small increase in the general cophenetic correlation values from 0.60 to 0.70. In practical terms, these results suggest that parental selection based on genetic diversity may be more effective when soybean breeders use smaller groups of genotypes to calculate coefficients and if genotypes are previously selected based on their qualitative and quantitative agronomic traits.

Among cultivars considered discrepant in the previous dendrogram (Figure 4) there were pairs of cultivars that were reallocated within the same group, indicating good agreement with their original coefficients. An example of this was the group gathering the Embrapa 32 (Itaquí), Embrapa 33 and Cariri RCH cultivars, all derived from backcrosses to the BR-27 (Cariri) cultivar. A similar situation was observed for cultivar MS/BRS-172 (Tuiuiú) and its recurrent parent (the FT-Cristalina cultivar) with four backcrosses (0.77). The MT/BR-45 (Paiaguás) cultivar, from the cross between Doko x IAC-7, had a GS of 0.79 with cultivar IAC-7 forming a group with cultivars IAC-7 and Nova IAC-7. Another example of improved grouping occurred between the sister cultivars MT/BR-55 (Uirapuru) and MT/BRS-159 (Crixás) in relation to cultivar MT/BR-50, this new group having GS coefficients higher than 0.70.

However, some cultivars remained in different groups in spite of sharing high similarity coefficients, probably reflecting deficiencies in the grouping approach, as shown by the moderate cophenetic correlation value. This could be seen in several cases, such as the groups containing cultivar BR/IAC-21, derived from five backcrosses to the IAC-8 cultivar, with GS coefficients of 0.40; the Coodetec 201 cultivar, derived from five backcrosses to the Ocepar 4 (Iguaçu) (0.75); and the two pairs of sister cultivars, BR-4 and BR-5 (0.75) and BRS 153 and BRS 154 (0.71).

In spite of such relationship, some cultivars were allocated to different groups because of the low similarity obtained from the AFLP markers, examples being cultivar BRS 137 and the recurrent parent Dourados cultivar, with a similarity coefficient of 0.65, as well as the Embrapa-20 (Doko RC) cultivar, derived from four backcrosses to the Doko cultivar, with a similarity coefficient of 0.63. A similar finding was observed between cultivar UFV-6 and its sister cultivars UFV-5, UFV-9, and UFV-10.

Factors that may affect genetic similarity estimates

Results contradictory to theoretical expectations, based on the genealogy of each cultivar are difficult to explain with certainty, but several sources of error may be found in such studies, e.g. the use of seeds containing ge-

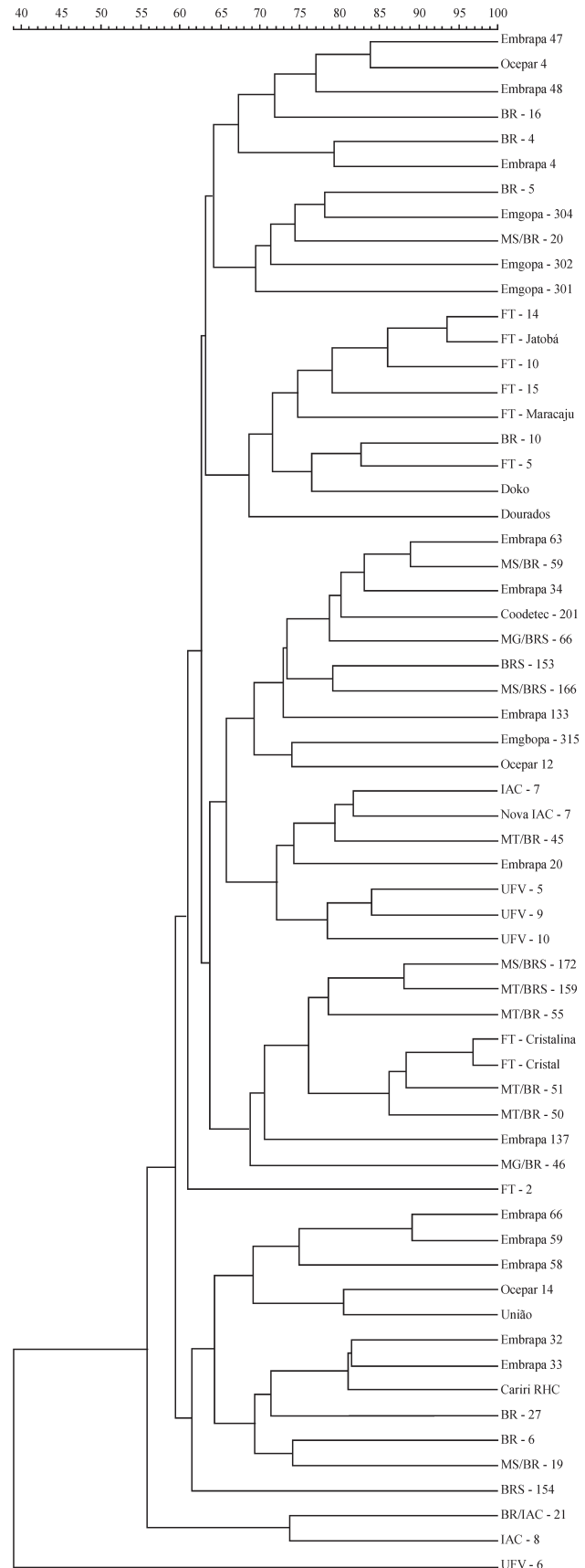


Figure 5 - UPGMA dendrogram based on AFLP similarity coefficients of 62 soybean cultivars and cophenetic correlation at the roots.

netic material not originating from the stated parent plants (extraneous DNA), quality of chemicals and reagents used in the analysis, imprecision in the AFLP analysis, and mistakes in reading and interpreting the polymorphic fragment data. In the present study extraneous DNA was probably the most important source of error, because leaf samples were not taken from a single plant since the use of genetic material taken from a single plant could result in atypical genetic data for that cultivar, this being especially true in the case of sister lines.

Several other factors may influence genetic similarity estimates and should be taken into account in studies of this nature. Firstly, the number of markers used can affect the variance in similarity estimate because a marker represents an independent genomic sample (Powell *et al.*, 1996), although in our study the 78 polymorphic markers used were found to be adequate for the analysis of the 317 cultivars.

Secondly, the distribution of markers in the genome is also an important factor to consider in diversity studies because a good coverage of the genome improves its representation efficiency as well as the comparison between individuals. It is normally assumed that markers are randomly distributed in the genome (Williams *et al.*, 1990) and there is evidence that AFLP provides a wide coverage of the plant genome (Maheswaran *et al.*, 1997).

Thirdly, the genetic similarity coefficient used may influence how similarity results are interpreted and grouped. For example, while the coefficient of Nei and Li (Nei and Li, 1979) does not consider the absence of bands as evidence of similarity between individuals the simple matching or common distance coefficient (S_{SM}) of Sokal and Michener (1958) does, which may cause the S_{SM} coefficient to overestimate genetic similarity because the absence of amplification in a dominant marker band common to two genotypes does not necessarily represent genetic similarity among the genotypes (Duarte *et al.*, 1999). In respect to Jaccard's coefficient (S_J) the Nei and Li coefficient differs only by the double weight it assigns to the occurrence of bands in both of any two analyzed genotypes (Duarte *et al.*, 1999; Mohammadi and Prasanna, 2003) and it thus seems that the Nei and Li coefficient is best suited to the type of analysis discussed in this paper.

Genetic modifications taken as *de novo* variability may also have an effect on this kind of study, with such modifications occurring because of intra-genic recombination, unequal crossing over, transposon activity and DNA methylation, although these factors are not thought to have been an important source of variability in our study.

Genetic similarity estimates based on the parentage coefficient (f) and correlation between the GS and f coefficients of genetic similarity

The 4,950 parentage coefficients (f) of the 100 soybean cultivars, released from 1984, varied from $f = 0$ to $f = 1$, with a mean of 0.21. When $f = 0$ there is no parentage

between cultivars pairs, and this occurred in this study in 294 (5.94%) of the GS estimates. The maximum f coefficient ($f = 1$) was found in 14 (0.28%) of the cultivars pairs. Vello *et al.* (1988) estimated the parentage coefficient of each of the 69 cultivars recommended for the 1983/84 growing season and observed a mean coefficient of 0.16. In the USA, cultivars released between 1947 and 1988 showed mean f coefficients of 0.18 for the cultivars from the north and 0.23 for the cultivars from the south (Gizlice *et al.*, 1993), values considered high. Sneller (1994) reported a similar finding when comparing elite cultivars from the north and south of the United States, as well as a surprisingly low f coefficient (0.10) between northern and southern regions, suggesting that soybean breeders have kept distinct gene pools. Analysis of Chinese soybean cultivars showed a very low f coefficient (0.02), indicating a high degree in genetic diversity of the germplasm of this country (Cui *et al.*, 2000).

In our study the correlation coefficient between GS and the parentage coefficients (f) was low ($r = 0.12$) for several reasons. One reason could be the fact that the two types of coefficients are not based on the same type of genetic similarity, because the f coefficient is a mathematical determination based on probabilities, while AFLP molecular markers detect similarities directly at DNA level. Another reason is that the conditions relating to the use of the f coefficient are not always fulfilled when dealing with the germplasm used in plant breeding. In fact, the violation of these assumptions seems to be critical for using this coefficient in the plant breeding context and this may lead to low correlation between f coefficients and genetic marker data. This happens because the f calculation assumes that each parent transfers 50% of its genetic material to its offspring, without considering the effects of selection and genetic drift (Barret *et al.*, 1998). Only in cases where cultivars are derived from selection within cultivars or by mutation must the f coefficient be equal to 1 (*i.e.* full similarity) because there is no possibility of a derived cultivar possessing different genes. However, the assumption that $f = 0$ (*i.e.* no similarity) when the ancestors of a cultivar do not have parentage in common is not always true and may possibly lead to an underestimation of the relationship between two genotypes. For example, Gizlice *et al.* (1993) used multivariate analysis to calculate similarity coefficients and found coefficients varying from 0 to 0.88 among North American soybean ancestors. Thus, lack of precise knowledge regarding the genealogy of a cultivar may interfere negatively with f coefficient estimates, a fact that does not occur with molecular markers because they do not require previous knowledge of genealogy for the calculation of genetic similarity.

The correlation between genetic similarity values obtained using markers and parentage (f) coefficients has been investigated in many studies with variable results. Very low correlation values ($r = 0.33$) were obtained for

wheat using RFLP (Barbosa-Neto *et al.*, 1996), while high ($r = 0.97$) values were obtained in maize with the same type of marker (Smith and Smith, 1991). Using AFLP markers, Barret and Kidwell (1998) found a correlation coefficient of 0.42 for wheat, which they explained in a similar manner to that outlined in the previous paragraph.

In this paper, we have presented the first global analysis of genetic similarity in Brazilian soybean germplasm and have shown that AFLP markers are a very rapid, effective and reliable tool for this type of analysis. These findings not only highlight the capacity of the AFLP technique but should also help Brazilian soybean breeders in the selection of parent-plants for their crossing programs.

Acknowledgments

The authors acknowledge the financial support received from the Brazilian agency Fundação Mato Grosso (FMT) and are grateful to P.A.V. Barroso for the bootstrap and E. Binneck for the cophenetic analysis. A.L.V. Bonato received a doctorate scholarship from the Brazilian agency CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico).

References

- Abdelnoor RV, Barros EG and Moreira MA (1995) Determination of genetic diversity within Brazilian soybean germplasm using random amplified polymorphic DNA techniques and comparative analysis with pedigree data. *Braz J Genet* 18:265-273.
- Barbosa-Neto JF, Sorrels ME and Cisar G (1996) Prediction of heterosis in wheat using coefficient of parentage and RFLP-based estimates of genetic relationship. *Genome* 39:1142-1149.
- Barrett BA and Kidwell KK (1998) AFLP-based genetic diversity among wheat cultivars from the Pacific Northwest. *Crop Sci* 38:1261-1271.
- Barrett BA, Kidwell KK and Fox PN (1998) Comparison of AFLP and pedigree-based genetic diversity assessment methods using wheat cultivars from the Pacific Northwest. *Crop Sci* 38:1271-1278.
- Barroso PAV, Geraldi IO, Vieira MLC, Pulcinelli CE, Vencosvky R and Dias CTS (2003) Predicting performance of soybean populations using genetic distances estimated with RAPD markers. *Genet Molec Biol* 26:343-348.
- Bonato ALV, Calvo ES, Arias CAA, Toledo JFF and Geraldi IO (2006) Prediction of genetic variability through AFLP-based measure of genetic distance in soybean (*Glycine max* L. Merrill). *Crop Breed Appl Biotechnol* 6:30-39.
- Brown-Guedira GL, Thompson JA, Nelson RL and Warburton ML (2000) Evaluation of genetic diversity of soybean introductions and North American ancestors using RAPD and SSR markers. *Crop Sci* 40:815-823.
- Congresso Brasileiro de Soja (1999) Anais... Embrapa Soja, Londrina, 533 pp (Embrapa Soja Documentos 124).
- Cox TS, Kiang YT, Gorman MB and Rodgers DM (1985) Relationship between coefficient of parentage and genetic similarity indices in the soybean. *Crop Sci* 25:529-532.
- Cui Z, Carter Jr TE and Burton JW (2000) Genetic diversity patterns in Chinese soybean cultivars based on coefficient of parentage. *Crop Sci* 40:1780-1793.
- Delannay X, Rodgers DM and Palmer RG (1983) Relative genetic contribution among ancestral lines to North American soybean cultivars. *Crop Sci* 23:944-949.
- Derbyshire E, Carvalho MTV and Bonato ER (1990) Comparison of natural variants of the soybean cultivar Paraná by isoenzyme analysis. *Braz J Genet* 13:83-87.
- Diwan N and 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.
- Duarte JM, Santos JB dos and Melo LC (1999) Comparison of similarity coefficients based on RAPD markers in the common bean. *Genet Molec Biol* 22:427-432.
- Gizlice Z, Carter Jr TE, Gerig TM and Burton JW (1996) Genetic diversity patterns in North American public soybean cultivars based on coefficient of parentage. *Crop Sci* 33:753-765.
- Gizlice Z, Carter TE and Burton JW (1993) Genetic diversity in North American soybean: I. Multivariate analysis of founding stock and relation to coefficient of parentage. *Crop Sci* 33:614-620.
- Helms T, Orf J, Vallad G and McClean P (1997) Genetic variance, coefficient of parentage, and genetic distance of six soybean populations. *Theor Appl Genet* 94:20-26.
- Hiramoto DM and Vello NA (1986) The genetic base of Brazilian soybean (*Glycine max* (L.) Merrill) cultivars. *Braz J Genet* 9:295-306.
- Kisha TJ, Diers BW, Hoyt JM and Sneller CH (1998) Genetic diversity among soybean plant introductions and North American germplasm. *Crop Sci* 38:1669-1680.
- Maheswaran M, Subudhi PK, Nandi S, Xu JC, Parco A, Yang DC and Huang N (1997) Polymorphism, distribution, and segregation of AFLP markers in a doubled haploid rice population. *Theor Appl Genet* 94:39-45.
- Malécot G (1947) *Les Mathématiques de l'Hérédité*. Masson et Cie, Paris, 80 pp.
- Manjarrez-Sandoval P, Carter Jr TE, Webb DM and Burton JW (1997) RFLP Genetic similarity estimates and coefficient of parentage as genetic variance predictors for soybean yield. *Crop Sci* 37:698-703.
- Maughan PJ, Saghai-Marooof MA and Buss GR (1996) Amplified fragment length polymorphism (AFLP) in soybean: Species diversity, inheritance, and near-isogenic line analysis. *Theor Appl Genet* 93:392-401.
- Mohammadi SA and Prasanna BM (2003) Analysis of genetic diversity in crop plants - Salient statistical tools and considerations. *Crop Sci* 43:1235-1248.
- Moser H and Lee M (1994) RFLP variation and genealogical distance, multivariate distance, heterosis, and genetic variance in oats. *Theor Appl Genet* 87:947-956.
- Muehlbauer GJ, Specht JE, Thomas-Compton MA, Staswick PE and Bernard RL (1988) Near-isogenic lines - A potential resource in the integration of conventional and molecular marker linkage maps. *Crop Sci* 28:729-735.
- Nei M and Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc Natl Acad Sci USA* 76:5269-5273.
- Oliveira KM, Laborda PR, Garcia AAF, Paterniani MEZ and Souza AP (2004) Evaluating genetic relationships between

- tropical maize inbred lines by means of AFLP profiling. *Hereditas* 140:24-33.
- Pejic I, Ajmone-Marsan P, Morgante M, Kozumplick V, Castiglioni P, Taramino G and Motto M (1998) Comparative analysis of genetic similarity among inbred lines detected by RFLPs, RAPDs, SSRs and AFLPs. *Theor Appl Genet* 97:1248-1255.
- Powell W, Morgante M, André C, Hanafey M, Vogel J, Tingey S and Rafalski A (1996) The comparison of RFLP, RAPD, AFLP, and SSR (microsatellite) markers for germplasm analysis. *Molec Breed* 2:225-238.
- Priolli RHG, Mendes Junior CT, Arantes NE and Contel EPB (2002) Characterization of Brazilian soybean cultivars using microsatellite markers. *Gen Mol Biol* 25:185-193.
- Rohlf FJ (1997) NTSYS-pc: Numerical Taxonomy and Multivariate Analysis System. EXETER Software, New York.
- Rubin SAL and Santos OS (1996) Progresso do melhoramento genético da soja no Estado do Rio Grande do Sul: I. Rendimento de grãos. *Pesq Agropec Gaúcha* 2:139-147.
- Saghai-Maroo MA, Sloman KM, Jorgensen RA and Allard RW (1984) Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. *Proc Natl Acad Sci USA* 81:8014-8018.
- Smith JSC and Smith OS (1991) Restriction fragment length polymorphisms can differentiate among U.S. maize hybrids. *Crop Sci* 31:893-899.
- Sneller CH (1994) Pedigree analysis of elite soybean lines. *Crop Sci* 34:1515-1522.
- Sneller CH, Miles J and Hoyt JM (1997) Agronomic performance of soybean plant introductions and their genetic similarity to elite lines. *Crop Sci* 37:1595-1600.
- Sokal RR and Michener CD (1958) A statistical method for evaluating systematic relationships. *Univ. Kansas Sci Bull* 38:1409-1438.
- Thompson JA and Nelson RL (1998) Utilization of diverse germplasm for soybean yield improvement. *Crop Sci* 38:1362-1368.
- Tivang JG, Nienhuis J and Smith OS (1994) Estimation of sampling variance of molecular marker data using bootstrap procedure. *Theor Appl Genet* 89:259-264.
- Toledo JFF, Almeida LA, Kiihl RAS and Menosso OG (1990) Ganho genético em soja no Estado do Paraná, via melhoramento. *Pesq Agropec Bras* 25:89-94.
- Ude GN, Kenworthy WJ, Costa JM, Cregan PB and Alvernaz 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.
- Vello NA, Hiromoto DM and Azevedo-Filho AJBV (1988) Coefficient of parentage and breeding of Brazilian soybean germplasm. *Braz J Genet* 11:679-697.
- Williams JGK, Kubelik AR, Livak KJ, Rafalski JA and Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucl Acids Res* 18:6531-6535.
- Zhu SL, Monti LM, Avitabile A and Rao R (1999) Evaluation of genetic diversity in Chinese soybean germplasm by AFLP. *Plant Genet Res Newsletter* 119:10-14.

Associate Editor: Everaldo Gonçalves de Barros