



## The genetic base of Brazilian soybean cultivars: evolution over time and breeding implications

Philip Traldi Wysmierski and Natal Antonio Vello

*Departamento de Genética, Escola de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, SP, Brazil.*

### Abstract

Genetic diversity is essential for crop breeding and one way to estimate it is through the concept of genetic base, which can be defined as the number of ancestors and their relative genetic contributions (RGC) to each cultivar. The RGC can be estimated through the coefficient of parentage between the ancestors and cultivars. Previous studies determined that the genetic base of Brazilian soybean was very narrow. The objective of this work was to evaluate the pedigree of 444 Brazilian soybean cultivars to estimate their genetic base. The cultivars were divided according to their release dates and according to their origin (public or private), and the genetic base for each group was also estimated. We found 60 ancestors, of which the top four (CNS, S-100, Roanoke and Tokyo, respectively) contribute 55.3% of the genetic base. Only 14 ancestors have an RGC over 1.0%, and they represent 92.4% of the genetic base. Analysis of the release dates indicated that there has been an increase in the number of ancestors over time, but the four main ancestors were the same over all periods, and their cumulative RGC increased from 46.6% to 57.6%, indicating a narrowing of the genetic base.

**Keywords:** genetic base, *Glycine max*, coefficient of parentage, genetic vulnerability.

Received: May 3, 2013; Accepted: August 5, 2013.

### Introduction

Soybean is one of the most important field crops in the world, and Brazil is the second largest producer and exporter of this commodity, behind the United States. The main countries which import soybean from Brazil are China (22,885,887 t), Spain (2,155,811 t), the Netherlands (1,036,919 t), Japan (548,339 t), Germany (522,354 t) and France (506,775 t). This crop was introduced into Brazil around the end of the nineteenth century. It gained economic importance during the 1970s and has been increasing ever since. In 2011/2012, 25 million ha of soybean were planted, which corresponds to almost half (49.2%) of the total area of field crops in the country, achieving a total production of 66 million metric tons and a mean yield of 2651 kg/ha (Conab, 2013).

The first soybean cultivars planted during the 1960s and 1970s were introduced from the South of the United States, e.g. Bragg, Davis, and Lee. With the growing importance of soybean, breeders began crossing these cultivars among themselves and with other sources, obtaining the first Brazilian cultivars, such as Industrial, Santa Rosa and Campos Gerais.

The frequent crossing of a small number of cultivars can lead to a reduction in genetic diversity. One way to measure the genetic diversity of a crop is through the concept of genetic base, which can be defined as the number of ancestors and their relative genetic contribution (RGC) to a specific group (Cui *et al.*, 2000a). The term ancestor usually refers to a founding stock with no known pedigree (Gizlice *et al.*, 1994). The RGC can be estimated through the mean coefficient of parentage (COP) between all the cultivars and an ancestor. This method has been used in many studies, and it is applicable when there are detailed pedigree records available. The genetic base of Brazilian cultivars was estimated by Hiromoto and Vello (1986), who analyzed 69 cultivars. They found only 26 ancestors, of which the main four ancestors (CNS, Roanoke, Tokyo and S-100) contributed 48.2% of the genetic base, and they concluded that it is narrow. The Brazilian and North American soybean genetic bases also shared six important ancestors. The genetic base for soybean cultivars has been estimated for the United States (Delannay *et al.*, 1983; Gizlice *et al.*, 1994; Sneller, 1994, 2002), China (Cui *et al.*, 2000a), Japan (Zhou *et al.*, 2000) and India (Bharadwaj, 2002). China has the largest genetic base, with 339 ancestors, while other countries have smaller ones (Table 1).

Since the study by Hiromoto and Vello (1986) there have been no large scale attempts to quantify the genetic base of Brazilian soybean cultivars and to analyze its evolu-

Send correspondence to Philip T. Wysmierski. Departamento de Genética, Escola de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Av. Pádua Dias n. 11, 13418-900 Piracicaba, SP, Brazil. E-mail: philiptraldi@yahoo.com.br, philiptraldi@usp.br.

**Table 1** - Comparison among the soybean genetic base of different countries based on the number of cultivars to equal or surpass a specific AGC, number of ancestors, number of cultivars analyzed, mean COP and period analyzed.

Country	AGC (%)			No. ancestors	No. cultivars	Mean COP	Period
	50	80	90				
China	35	190	339	339	651	0.02	1923-1995
Japan	18	53	<sup>a</sup>	74	86	0.04	1950-1988
India	5	<sup>a</sup>	<sup>a</sup>	76	66	0.05	1968-2000
USA Total	6	14	27	80	258	0.13	1947-1988
USA North	4	14	20	67	<sup>a</sup>	0.18	
USA South	3	10	16	39	<sup>a</sup>	0.23	
Brazil	5	9	12	26	69	0.16	1960-1984
Brazil <sup>b</sup>	4	9	13	60	444	<sup>a</sup>	1960-2009

Sources: Hiromoto and Vello, 1986; Gizlice *et al.*, 1994; Cui *et al.*, 2000a,b; Zhou *et al.*, 2000; Bharadwaj *et al.*, 2002.

Note: AGC, accumulated genetic contribution; COP, coefficient of parentage.

<sup>a</sup>Data not available.

<sup>b</sup>Data from this study.

tion. Therefore, the objectives of this work are to (i) quantify the genetic base of 444 Brazilian soybean cultivars covering the years from 1943 to 2009, using COP obtained through pedigree records, (ii) quantify the changes that occurred in the genetic base over five periods (before 1971, 1971-1980, 1981-1990, 1991-2000 and 2001-2009), (iii) compare the genetic base of public and private soybean cultivars, and (iv) compare the Brazilian genetic base to that of other countries.

## Material and Methods

The 444 cultivars soybean cultivars analyzed in this study were foreign, public, or private cultivars that were widely grown by farmers during certain periods. We made no distinction between specialty soybean, such as soybean for human consumption, and commodity soybean. Among the 444 cultivars 14 are public transgenic cultivars, which were included in this study. The pedigree data was obtained from various sources in the literature, from the Germplasm Resources Information Network and information from breeders. Ancestors were defined as founding stock with no known pedigree (Gizlice *et al.*, 1994).

Malécot's COP ( $f_{xy}$ ) was estimated based on the equation presented by Falconer and Mackay (1996), as follows

$$f_{xy} = \frac{1}{4}(f_{AC} + f_{AD} + f_{BC} + f_{BD})$$

where  $f$  is the COP between two cultivars X and Y; A and B are the parents of cultivar X and C and D are parents of cultivar Y. We started by assembling the complete pedigrees of the Brazilian soybean cultivars, and traced them back to the ancestors, estimating the probability of the contribution that each ancestor made to each cultivar, that is, the COP between an ancestor and a cultivar. The following assumptions were made: (i) all the ancestors are independent, in other words, the COP between pairs of ancestors is

$f = 0$ ; (ii) in a cross, each parent contributes 50% of the genes; (iii) the cultivars, ancestors, and lines are all homozygous and homogenous; (iv) the COP between a plant selection and its antecedent was considered to be  $f = 1$ ; (v) the COP between a mutant and the original genotype was considered to be  $f = 1$ . The calculations were performed using the computer software Microsoft Excel® (2003).

The relative genetic contribution (RGC) of different ancestors to a given selection was calculated by partitioning the genetic constitution of each cultivar, using the COP as an estimate. The mean RGC of the ancestors was calculated by averaging the RGC of that ancestor from all cultivars. The accumulated genetic contribution (AGC) is the sum, in decreasing order of importance, of the ancestral RGCs. The presence of an ancestor in a pedigree was determined as its frequency. The mean number of ancestors was determined by averaging the number of ancestors each cultivar had. The ancestor/cultivar ratio was also determined.

The cultivars were subdivided into groups according to the year of release or year of registration (when available) in Brazil. The timeframe was divided into five periods: before 1971, 1971-1980, 1981-1990, 1991-2000 and 2001-2009. The cultivars were also subdivided according to origin: public or private. The few cultivars that were developed abroad were excluded from this analysis. For each subgroup, all the above mentioned parameters were also calculated.

## Results and Discussion

### Pedigrees

Although over 700 cultivars were officially registered in Brazil in 2010, according to CultivarWeb, the database for the National Cultivar Registry from the Ministry of Agriculture, Livestock and Food Supply, we were only able to

obtain the pedigrees for 444 cultivars. Due to the Variety Protection Act of 1997, many breeders have not made public the pedigrees of released cultivars, especially more recent ones. A few pedigrees deserve special attention. First, the cultivar FT-Cristalina originated from a natural cross between UFV-1 and an unknown parent. But based on morphological data, some breeders suggest that the unknown parent was Davis (Hiromoto and Vello, 1986), so this pedigree was used in the present work. The same authors reported that cultivar BR-9 (Savana) originated from the bulk LoB 74-2, with a probable pedigree being Davis x (Hill x PI 240664). Another important observation made is that morphological and molecular marker data have not supported the pedigree of the cultivar Lincoln (Gizlice *et al.*, 1994;

Lorenzen *et al.*, 1995), which appears in the genealogy of 28 Brazilian cultivars. So either the genealogy is incorrect, or the accessions maintained in the germplasm bank are not the real ancestors for Lincoln. Therefore, we chose to make Lincoln into an ancestral line.

### Overview of the genetic base

The 444 cultivars analyzed traced back to 60 ancestors (Table 2). The four main ancestors (CNS, S-100, Roanoke and Tokyo), contribute more than half (55.3%) to the genetic base, and CNS contributes alone with almost one-fifth of the genetic base (19.2%). Only the top 14 ancestors have a mean RGC of over 1%, and their AGC is 92.4%. Therefore, the 46 remaining ancestors contribute cumula-

**Table 2** - Identification, mean RGC, AGC, frequency, origin of the ancestors of Brazilian soybean cultivars, mean RGC of ancestors for public and private cultivars, first progeny and CA.

No.	Ancestors <sup>a</sup>	RGC <sup>b</sup> (%)	AGC (%)	Freq.	Origin <sup>c</sup>	RGC (%) public <sup>d</sup>	RGC (%) private <sup>d</sup>	First progeny	CA
1	CNS (PI 548445)	19.2	19.2	435	Jiangsu, China	19.1	19.1		
2	S-100 (PI 548488)	14.8	34.0	403	Heilongjiang, China	14.8	14.7		
3	Roanoke (PI 548485)	11.7	45.7	389	Jiangsu, China	11.3	12.8		
4	Tokyo (PI 548493)	9.6	55.3	423	Kanagawa, Japan	8.7	10.6		
5	PI 54610	8.6	63.9	419	Jilin, China	8.1	9.5		
6	PI 60406 <sup>e</sup>	5.0	68.9	280	Fujian, China	5.3	4.3	Improved Pelican (PI 548461)	1/2
7	Arksoy (PI 548438)	4.3	73.2	261	Pyongyang, North Korea	4.0	5.6		
8	Dunfield (PI 548318)	4.2	77.4	312	Jilin, China	4.2	4.3		
9	Haberlandt (PI 548456)	4.2	81.6	309	Pyongyang, North Korea	4.2	4.3		
10	Bilomi No. 3 (PI 240664)	3.0	84.6	101	Luzon, Philippines	3.6	2.2		
11	Biloxi (PI 548444)	2.6	87.2	287	Zhejiang, China	2.8	2.2		
12	Tanner (PI 548490)	2.5	89.7	280	Taiwan	2.6	2.2		
13	Palmetto (PI 548480)	1.7	91.4	149	Jiangsu, China	1.8	1.3		
14	Mogiana <sup>c</sup>	1.0	92.4	54	Brazil	1.5	0.2	Industrial (PI 628857)	1/2
15	Pine Dell Perfection (PI 548481)	0.7	93.1	68	Virginia, USA	0.8	0.8		
16	Laredo (PI 548463)	0.7	93.8	73	Shaanxi, China	0.6	0.6		
17	Mammoth Yellow (PI 548469)	0.7	94.5	72	Japan	0.6	0.6		
18	PI 274454	0.6	95.1	32	Okinawa, Japan	0.8	0.4		
19	PI 171442	0.5	95.6	34	Shaanxi, China	0.4	0.9		
20	Lincoln (PI 548362)	0.4	96.0	28	Illinois, USA	0.2	0.7		
21	Nanda (PI 548474)	0.4	96.4	20	Hwanghae Puk, North Korea	0.5	0.1		
22	Mandarin (PI 548378)	0.3	96.7	19	Heilongjiang, China	0.3	0.2		
23	FC 31745	0.3	97.0	26	Unknown	0.3	0.1		
24	Er-hej-jan (PI 437654)	0.3	97.3	20	China	0.3	0.1		
25	Peking (PI 548402)	0.2	97.5	47	Beijing, China	0.2	0.2		
26	PI 80837	0.2	97.7	3	Kanto and Tosan, Japan	0.2	0.2		
27	Soden-diazu (PI 229358)	0.2	97.9	10	Kanto, Japan	0.3	*		
28	Kedelee STB No. 452 <sup>e</sup>	0.2	98.1	2	Unknown	0.2	0.2	Planalto (PI 628883)	1/2
29	Aliança Preta <sup>e</sup>	0.1	98.2	1	Brazil	0.2	*	IAC 1 (PI 628842)	1/2

**Table 2 (cont.)**

No.	Ancestors <sup>a</sup>	RGC <sup>b</sup> (%)	AGC (%)	Freq.	Origin <sup>c</sup>	RGC (%) public <sup>d</sup>	RGC (%) private <sup>d</sup>	First progeny	CA
30	Chosen <sup>e</sup>	0.1	98.3	1	Unknown	0.2	*	Pioneira (PI 417503)	1/2
31	Asomusume (PI 423959)	0.1	98.4	1	Kumamoto, Japan	0.2	*		
32	Tamahomare (PI 507327)	0.1	98.5	1	Japan, Kanto and Tosan	0.2	*		
33	Patoka (PI 548400)	0.1	98.6	16	Heilongjiang, China	†	0.2		
34	Kuro Daizu (PI 81041) <sup>e</sup>	0.1	98.7	16	Hokkaido, Japan	†	0.2	Perry (PI 548603)	1/2
35	Amarela Comum <sup>e</sup>	0.1	98.8	4	Brazil	0.1	0.1	Ivorá (PI 628860)	1/4
36	Shinano mejiro (PI 507244)	0.1	98.9	4	Japan, Kanto and Tosan	0.1	0.1		
37	Howgyoku (PI 224270)	0.1	99.0	4	Hyogo, Japan	0.1	0.1		
38	Richland (PI 548406)	0.1	99.1	12	Jilin, China	†	0.1		
39	Kanro (PI 84928)	0.1	99.2	3	Pyongyang, North Korea	0.1	*		
40	PLSO-41 (PI 346304)	0.1	99.3	2	India	*	0.3		
41	Mandarin (Ottawa) (PI 548379)	0.1	99.4	2	Heilongjiang, China	0.1	*		
42	Blyvoor (PI 374219)	0.1	99.5	4	North West, South Africa	†	0.2		
43	Kyushu 111 <sup>e</sup>	0.1	99.6	5	Japan	0.1	*	UFVTN 101, UFVTN 102, UFVTN 103, UFVTN 104	1/1 6
44	Kin-du (PI 157440)	0.1	99.7	1	South Korea	0.1	*		
45	Miyako White (PI 227687)	0.1	99.8	2	Okinawa, Japan	0.1	*		
46	PI 230979	0.1	99.9	1	Tokyo, Japan	*	0.2		
47	Shih-Shih (PI 262181)	†	99.9	2	Japan	0.1	*		
48	PI 159925	†	99.9	5	Lima, Peru	0.1	*		
49	PI 88788	†	99.9	6	Liaoning, China	†	†		
50	Goyou Kurakake (PI 86023) <sup>e</sup>	†	99.9	5	Hokkaido, Japan	†	*	UFVTN 101, UFVTN 102, UFVTN 103, UFVTN 104	1/3 2
51	Shirohadaka (PI 86490)	†	99.9	1	Akita, Japan	†	*		
52	PI 71506 <sup>e</sup>	†	99.9	1	Jiangsu, China	†	*	Hutcheson (PI 518664)	1/4
53	V-148 <sup>e</sup>	†	99.9	1	Unknown	†	*	IAC 14 (PI 628849)	1/8
54	A.K. Harrow (PI 548298)	†	99.9	2	China	†	*		
55	No. 171 <sup>e</sup>	†	99.9	1	China	†	*	Capital (PI 548311)	1/2
56	GTS-40-3-2 <sup>f</sup>	†	99.9	14	USA	†	*		
57	TGX-342-351D <sup>e</sup>	†	99.9	3	Nigeria	†	*	BRS 214	1/6 4
58	I-Higo-Wase (PI 205085)	†	99.9	5	Japan	†	*		
59	PI 90763	†	99.9	1	Beijing, China	†	*		
60	Mukden (PI 548391)	†	100.0	2	Liaoning, China	†	*		

Note: RGC, relative genetic contribution; AGC, accumulated genetic contribution; CA, contribution from ancestor to first progeny.

<sup>a</sup>PI numbers from GRIN (Germplasm Resources Information Network) are provided when available.

<sup>b</sup>† indicates that this ancestor contributes less than 0.1% to the genetic base.

<sup>c</sup>The primary geographical origins are China, Japan, North Korea and South Korea; the secondary ones are other locations mentioned (based on GRIN (Germplasm Resources Information Network)).

<sup>d</sup>\* indicates that the ancestor does not contribute to the public or private pool.

<sup>e</sup>no known seed sources available.

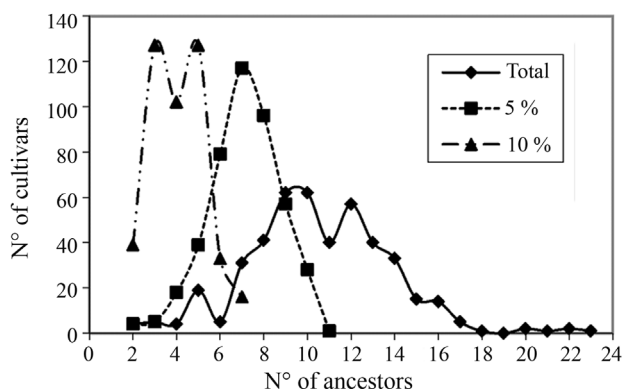
<sup>f</sup>Developed through particle gun bombardment of line A5403 (PI 556861).

tively only 7.6% to the genetic base. For comparison, Hiro-moto and Vello (1986) reported 26 ancestors for a group of 69 cultivars, and the AGC of the top four ancestors was 48.2%. These were the same top four ancestors found in this study, but with a different order. These results indicate that the genetic base of Brazilian soybean is still narrow, despite the incorporation of new ancestors.

For those ancestors that no longer exist (or we do not know from where seeds are available) we provide a first progeny, as defined by Gizlice *et al.* (1994), for screening options due to the missing ancestor. For the top 20 ancestors, only two are unavailable (PI 60406 and Mogiana).

On average, there were 10.52 ancestors for each cultivar, ranging from 2 to 23. However, this increase must be taken with caution, because it is mainly due to the incorporation of low-contribution ancestors. For example, the mean number of ancestors for the total period is 10.52 (Table 3), a high value when compared to the Chinese (3.79) and Japanese (3.20) values (Cui *et al.*, 2000a; Zhou *et al.*, 2000). But when we only consider ancestors with RGCs of 5% and 10% to cultivars, this number drops to 7.11 and 4.08, respectively (Figure 1), demonstrating that many ancestors have only small contributions.

The frequency of ancestors also demonstrates how narrow the genetic base is. The main ancestor, CNS, is present in the pedigree of 435 cultivars (98.0%). The other top ancestors (S-100, Roanoke and Tokyo) also have very high frequencies. CNS and S-100 are the most common ancestors because their cross resulted in cultivar Lee and the line D49-2491, a sister of Lee and an ancestor of Bragg. Lee and Bragg were used as parents in many early cultivars developed in Brazil. They are present in 118 and 135 cultivars, respectively. As the RGC decreases so does the frequency, both being highly correlated ( $r = 0.88$ ), to the extreme that some ancestors (11) only contribute to one cultivar. Notwithstanding, some ancestors have low RGCs but have higher than expected frequencies. This usually occurs because the ancestor was used in backcrosses to incorporate



**Figure 1** - Number of ancestors that contribute to cultivars irrespective of RGC (circle), with RGC equal to or greater than 5% (square) and 10% (triangle).

simple traits. One clear example is the ancestor GTS-40-3-2, which was used to incorporate the *CP4 EPSPS* gene, conferring resistance to glyphosate, into the 14 transgenic cultivars analyzed here (Sneller, 2003).

As seen above, most of the ancestors have very low mean RGCs. We believe this occurs because breeders usually use new lines as parents to introduce specific genes, usually involving backcrossing. Therefore, parents used as gene donors through backcrossing lead to an increase in the number of ancestors, but do not contribute significantly to increase the genetic base of the cultivated species.

A database search on GRIN (Germplasm Resources Information Network) allowed us to identify a few notable characteristics of some ancestors (Table 4). Some low-contributing ancestors (but not all) may have been incorporated into the genetic base for these traits. For example, the ancestor I-Higo-Wase (PI 205085), which has a mean RGC of 0.0036% and contributes to only five cultivars, has the *lx3* allele, which confers an absence of lipoxygenase-3 in seeds. This ancestor was used to incorporate this trait into five soybean cultivars developed for human consumption. Another ancestor used in this manner is PI 88788, with a mean RGC of 0.0299%. This ancestor is resistant to soybean cyst nematode, races 3 and 14. Shirohadaka (PI 86490), with the small seed trait, was used to incorporate this characteristic into a cultivar aimed at producing natto, a fermented soybean dish appreciated in Japan.

The geographical origin of the 60 ancestors is provided (Table 2). The primary geographical origins were considered to be the countries where soybean may have originated, or are considered as centers of diversity (based on information from GRIN). When the primary geographical origin was unknown, we used the secondary geographical origin, which is the country where the ancestor was collected or developed. When no information was available, the origin was left as unknown. Most of the ancestors are from Asia (78%), mainly from China and Japan. One assumption made in this work is that ancestors are unrelated; however, this may not be the case, especially with ancestors originating in the same area (*e.g.* Jiangsu, China). Therefore, the relationship between ancestors and cultivars is probably underestimated and the real genetic base may be narrower than the estimated genetic base (Hiromoto and Vello, 1986; Mikel *et al.*, 2010). Indeed, Gizlice *et al.* (1993), using multivariate analysis of ten traits, estimated genetic similarity coefficients for 14 ancestors of North American soybean cultivars, which ranged from 0.00 to 0.88, demonstrating that some are indeed similar and may have some degree of relatedness.

### Period analysis

The results from the different periods (Table 3) show that as soybean grew in importance in Brazil the number of cultivars increased. Before 1971 there were 17 cultivars and in the last period (2001-2009) there were 116 cultivars,



**Table 3** - Analysis of all the periods, indicating number of cultivars and ancestors in each period, average number of ancestors per cultivar (anc./cv), AGC of the top 4 ancestors, number of ancestors that correspond from 50% to 90% of the genetic base and ratio of ancestors per released cultivar (Anc./cv ratio) in each period.

Period	No. cultivars	No. ancestors	Average no. of anc./cv	AGC of top 4 ancestors (%)	Genetic base					Anc./cv ratio
					50	60	70	80	90	
Before 1971	17	18	4.88	46.60	5	6	8	10	13	1.06
1971-1980	47	26	6.68	54.13	4	5	7	10	13	0.55
1981-1990	112	41	9.11	52.54	4	5	7	9	13	0.37
1991-2000	152	40	11.27	56.80	4	5	7	9	12	0.26
2001-2009	116	52	13.20	57.59	4	5	6	9	13	0.45
Total	444	60	10.52	55.26	4	5	7	9	13	0.14

Note: AGC, accumulated genetic contribution.

not counting cultivars with unavailable pedigrees. We also clearly detect an increase in the number of ancestors over time, also supported by an increase in the mean number of ancestors per cultivar, increasing from 4.88 to 13.20. But here one must take into account that many ancestors have only low contributions (Figure 1).

The genetic base of Brazilian soybean cultivars has decreased over the past few decades (Table 3). The top four ancestors in each period were always CNS, S-100, Roanoke and Tokyo, and with the exception of CNS, which remained as the ancestor with the highest RGC throughout all periods, the ranking of the other three changed. In the final period (2001-2009), CNS alone represented more than one-fifth of the genetic base (20.7%). These results also support the fact that the genetic base has decreased over time. Delannay *et al.* (1983), studying the North American genetic base also observed a similar effect, with the number of ancestors increasing over time, but accompanied by an increase of the main ancestors' RGC. On the other hand, Cui *et al.* (2000a) observed that the Chinese soybean genetic base increased over time, with the incorporation of new ancestors and the RGC of ancestors became more uniformly distributed.

#### Public vs. private cultivars

Of the 444 cultivars analyzed, 301 were of public and 112 of private origin, and 31 were of foreign or unknown origin. The public cultivars had 58 ancestors (Table 2), close to the total number of ancestors (60), with only two ancestors contributing exclusively to the private cultivars, PI 346304 and PI 230979. The top four ancestors for public cultivars are identical to the total ancestors of the total group of cultivars, and their AGC is 53.9%. The top 13 ancestors for this group contributed with 90.5% to the genetic base, and are also the same top 13 ancestors for the total group of cultivars, with minor changes in rank.

On the other hand, the private cultivars had only 37 ancestors. This low number of ancestors may reflect the lower number of cultivars in this group and the use of a restricted number of cultivars in private breeding programs. However, the genetic base of this group does not significantly differ from the public and the total group. For example, the top four ancestors for private cultivars are the same as those for the public group and general group of cultivars, CNS, S-100, Roanoke and Tokyo in this order, but with a slightly higher AGC (57.2%). Also, the top 13 ancestors for this group, which represent 93.1% of the genetic base, are also the same ancestors for the other two groups, with minor changes in rank.

#### Comparison with other genetic bases

When comparing the results from this study to the soybean genetic bases of other countries (Table 1), one can see that Brazil still has a very narrow genetic base. Hiro-moto and Vello (1986) reported that six of the main ances-

**Table 4** - Ancestors, identified by name and number based on mean RGC and notable traits.

No.	Ancestors	Some notable traits <sup>a</sup>
1	CNS (PI 548445)	<i>Rj2</i> allele, inefficient nodulation of b7, b14 and b122 <i>Rhizobium</i> strains; <i>Rps2</i> allele, resistant to <i>Phytophthora</i> rot, races 1 and 2; <i>Rpv1</i> allele, resistant to Peanut Mottle Virus, M-2; <i>rxp</i> allele, resistant to bacterial pustule.
2	S-100 (PI 548488)	Allele <i>Ncl1</i> , chloride exclusion; allele <i>Dt1</i> , indeterminate growth; resistant to <i>Phytophthora</i> rot race 1.
3	Roanoke (PI 548485)	Resistant to frog eye leaf spot.
4	Tokyo (PI 548493)	Allele <i>Rsv1-t</i> , resistant to Soybean Mosaic Virus, strains G1, G2, G4, G5 and G6.
5	PI 54610	Resistant to frog eye leaf spot races 2 to 5
7	Arksoy (PI 548438)	Allele <i>Rps1-c</i> , resistant <i>Phytophthora</i> rot, races 1 to 3, 6 to 11, 13, 15, 17 and 21
9	Haberlandt (PI 548456)	Resistant to frog eye leaf spot
10	Bilomi No. 3 (PI 240664)	Long juvenile period
18	PI 274454	Resistant to insects (e.g. <i>Diabrotica speciosa</i> and <i>Cerotoma</i> sp.) <sup>b</sup>
19	PI 171442	Allele <i>Rps3</i> ; resistant <i>Phytophthora</i> rot, races 1 to 5, 8, 9, 11, 13, 14, 16, 18
23	FC 31745	Allele <i>Rps1-b</i> ; resistant <i>Phytophthora</i> rot, races 1, 3 to 9, 13 to 15, 17, 18, 21, 22
24	Er-hej-jan (PI 437654)	Resistant to soybean cyst nematode (alleles <i>rgh1</i> and <i>Rgh4</i> ) <sup>c</sup>
25	Peking (PI 548402)	Resistant to soybean cyst nematode (alleles <i>rgh 1</i> to 3 and <i>Rhg 4</i> ) <sup>c</sup>
26	PI 80837	Resistant to <i>Phomopsis</i> spp. <sup>d</sup> ; resistant to <i>Cercospora kikuchii</i> <sup>e</sup> ; allele <i>Pd1</i> dense pubescence <sup>f</sup>
27	Soden-diazu (PI 229358)	Resistant to insects (e.g. <i>Diabrotica speciosa</i> and <i>Cerotoma</i> sp.) <sup>b, g</sup>
31	Asomusume (PI 423959)	Resistant to <i>Phytophthora</i> rot, race 1
33	Patoka (PI 548400)	Allele <i>rcs1</i> , susceptible to frog eye leaf spot, race 1.
36	Shinano mejiro (PI 507244)	Resistant to <i>Phytophthora</i> rot races 1 to 8, 17 and 25
40	PLSO-41 (PI 346304)	Resistant to CCMV (Cowpea Chlorotic Mottle Virus)
43	Kyushu 111	Triple null for lipoxygenases 1, 2 and 3 (alleles <i>lx1</i> , <i>lx2</i> , and <i>lx3</i> )
44	Kin-du (PI 157440)	Allele <i>ti</i> , absence of Kunitz trypsin inhibitor <sup>f</sup>
45	Miyako White (PI 227687)	Resistant to insects <sup>h</sup>
46	PI 230979	Resistant to <i>Phytophthora</i> rot, races 1 to 4
47	Shih-Shih (PI 262181)	Resistant to <i>Phytophthora</i> rot, races 1, 3 and 4; vegetable and conventional soybean
48	PI 159925	Allele <i>j</i> , long juvenile period
49	PI 88788	Resistant to soybean cyst nematode, races 3 and 14 <sup>c</sup>
50	Goyou Kurakake (PI 86023)	Allele <i>lx2</i> , absence of lipoxygenase-2
51	Shirohadaka (PI 86490)	Small seeds
52	PI 71506	Resistant to soybean aphid ( <i>Aphis glycines</i> ) <sup>i</sup> ; poor germination under high temperatures <sup>j</sup>
56	GTS-40-3-2	Source of <i>CP4 EPSPS</i> gene <sup>k</sup>
58	I-Higo-Wase (PI 205085)	Allele <i>lx3</i> , absence of lipoxygenase-3
59	PI 90763	Resistant to soybean cyst nematode, races 1 to 3 and 5 <sup>c</sup>
60	Mukden (PI 548391)	Resistant to <i>Phytophthora</i> rot (allele <i>Rps1</i> ), races 1, 2, 10, 13, 16

Sources: <sup>a</sup> GRIN (2010); <sup>b</sup> Lara *et al.* (1999); <sup>c</sup> Brucker *et al.* (2005); <sup>d</sup> Jackson *et al.* (2005); <sup>e</sup> Jackson *et al.* (2006); <sup>f</sup> Palmer *et al.* (2004); <sup>g</sup> Clark *et al.* (1972); <sup>h</sup> Rose *et al.* (1989); <sup>i</sup> Hill *et al.* (2004); <sup>j</sup> Smith *et al.* (2008); <sup>k</sup> Parrott and Clemente (2004).

Note: RGC, relative genetic contribution.

tors of that time were shared with the North American soybean genetic base. In this study, a total of 26 ancestors were shared, based on Gizlice *et al.* (1994). When only considering the main ancestors of both countries, then the Brazilian and North American soybean genetic bases share only six ancestors (CNS, S-100, Roanoke, Tokyo, PI 54610 and Dunfield), the same number as reported by Hiromoto and Vello (1986). Another point is that the top five ancestors described here (CNS, S-100, Roanoke, Tokyo and PI

54610) are the exact same top five ancestors for the soybean genetic base of the southern USA. This, as Hiromoto and Vello (1986) argue, is due to the fact that the Brazilian soybean genetic base was partially derived from cultivars from this part of the USA.

When compared to the Chinese (Cui *et al.*, 2000a) and Japanese (Zhou *et al.*, 2000) genetic base, Brazil apparently shares only 9 (S-100, Peking, Mandarin (Ottawa), A.K. Harrow, CNS, Richland, Mukden, Lincoln and Man-

darin) and 7 (Kuro Daizu, Mandarin (Ottawa), CNS, S-100, A.K. Harrow, PI 54610 and Tokyo) ancestors, respectively, based on ancestors names. And these shared ancestors have very low contributions to soybean cultivars from China and Japan. Therefore, as reported previously by Yamanaka *et al.* (2007), the Brazilian gene pool is distinct from the Chinese and Japanese gene pools, and they can represent sources for increasing the Brazilian soybean genetic base. However, since ancestor names and designations can vary or become lost, it is likely that the Brazilian, Chinese and Japanese genetic bases share more ancestors than currently observed.

### Breeding implications

Based on the data presented here, breeders interested in expanding the genetic base of their programs may choose parents with the most divergent pedigrees from the present genetic base. As shown above, cultivars from Japan and China may be good choices because of the difference between their genetic bases and Brazil's genetic base. Cui *et al.* (2000a) reported that Chinese breeders have used this approach with some success, incorporating US cultivars into Chinese breeding programs, also because of their genetic divergence.

Another potential use for the ancestors of a genetic base is described by Carter Jr *et al.* (2004). They propose that breeders can screen the ancestors for a given trait, instead of the cultivars, because there are many more cultivars than ancestors. After the identification of the ancestors with a particular trait, the cultivars with the highest probability of having that trait can be identified, based on their RGC with the identified ancestors. This would allow researchers to screen much more rapidly for a given trait and would also save valuable resources in cases where the trait is difficult or expensive to evaluate. This approach has been used to screen many traits in the North American genetic base, including foliar resistance to ozone injury (Burkey and Carter Jr, 2008) and tolerance to herbicides (Carter Jr *et al.*, 2004).

### Acknowledgments

The authors would like to thank the reviewers for their valuable comments. This received financial support and scholarships from CAPES (Coordination for the Improvement of Higher Education Personnel) and CNPq (National Council for Scientific and Technological Development).

### References

Bharadwaj CH, Satyavathi CT, Tiwari SP and Karmakar PG (2002) Genetic base of soybean (*Glycine max*) varieties released in India as revealed by coefficient of parentage. *Indian J Agric Sci* 72:467-469.

Burkey KO and Carter Jr TE (2008) Foliar resistance to ozone injury in the genetic base of U.S. and Canadian soybean and

prediction of resistance in descendent cultivars using coefficient of parentage. *Field Crops Res* 111:207-217.

Brucker E, Carlson S, Wright E, Niblack T and Diers B (2005) *Rhg1* alleles from soybean PI 437654 and PI 88788 respond differentially to isolates of *Heterodera glycines* in the greenhouse. *Theor Appl Genet* 111:44-49.

Carter Jr TE, Nelson RL, Sneller CH and Cui Z (2004) Genetic diversity in soybean. In: Boerma HR and Specht JE (eds) Soybeans: Improvement, Production and Uses. 4<sup>th</sup> edition. American Society of Agronomy, Crop Science Society of America and Soil Science Society of America, Madison, pp 302-416.

Clark WJ, Harris FA, Maxwell FG and Hartwig EE (1972) Resistance of certain soybean cultivars to bean leaf beetle, striped blister beetle and bollworm. *J Econ Entomol* 65:1669-1672.

Cui Z, Carter Jr TE and Burton JW (2000a) Genetic base of 651 Chinese soybean cultivars released during 1923 to 1995. *Crop Sci* 40:1470-1481.

Cui Z, Carter Jr TE and Burton JW (2000b) 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 contributions among ancestral lines to North American soybean cultivars. *Crop Sci* 23:944-949.

Falconer DS and Mackay TFC (1996) Small populations III: pedigree populations and close inbreeding. In: Falconer DS and Mackay TFC (eds) *Introduction to Quantitative Genetics*. 4th edition. Longman, Essex, pp 82-99.

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.

Gizlice Z, Carter TE and Burton JW (1994) Genetic base for North American public soybean cultivars released between 1947 and 1988. *Crop Sci* 34:1143-1151.

Hill CB, Li Y and Hartman GL (2004) Resistance to the soybean aphid in soybean germplasm. *Crop Sci* 44:98-106.

Hiramoto DM and Vello NA (1986) The genetic base of Brazilian soybean (*Glycine max* (L.) Merrill) cultivars. *Brazil J Genet* 9:295-306.

Jackson EW, Fenn P and Chen P (2005) Inheritance of resistance to *Phomopsis* seed decay in soybean PI 80837 and MO/PSD-0259 (PI 562694). *Crop Sci* 45:2400-2404.

Jackson EW, Fenn P and Chen P (2006) Inheritance of resistance to purple seed stain caused by *Cercospora kikuchii* in PI 80837 soybean. *Crop Sci* 46:1462-1466.

Lara FM, Elias JM, Baldin ELL and Barbosa JC (1999) Feeding preference of *Diabrotica speciosa* (Germ.) and *Cerotoma* sp. for soybean genotypes. *Sci Agric* 56:947-951.

Lorenzen LL, Boutin S, Young N, Specht JE and Shoemaker RC (1995) Soybean pedigree analysis during map-based molecular markers. I: Tracking RFLP markers in cultivars. *Crop Sci* 35:1326-1336.

Mikel MA, Diers BW, Nelson RA and Smith HH (2010) Genetic diversity and agronomic improvement of North American soybean germplasm. *Crop Sci* 50:1219-1229.

Palmer RG, Pfeiffer TW, Buss GR and Kilen TC (2004) Qualitative genetics. In: Boerma HR and Specht JE (eds) Soybeans: Improvement, Production and Uses. 4th edition. American Society of Agronomy, Crop Science Society of America and Soil Science Society of America, Madison, pp 137-233.



- Parrott WA and Clemente TE (2004) Transgenic soybean. In: Boerma HR and Specht JE (eds) Soybeans: Improvement, Production and Uses. 4th edition. American Society of Agronomy, Crop Science Society of America and Soil Science Society of America, Madison, pp 137-233.
- Rose RL, Sparks TC and Smith CM (1989) The influence of resistant soybean (PI 227687) foliage and coumestrol on the metabolism of xenobiotics by the soybean looper, *Pseudoplusia includens* (Walker). *Pestic Biochem Physiol* 34:17-26.
- Smith JR, Mengistu A, Nelson RL and Paris RL (2008) Identification of soybean accessions with high germinability in high-temperature environments. *Crop Sci* 48:2279-2288.
- Sneller CH (1994) Pedigree analysis of elite soybean lines. *Crop Sci* 34:1515-1522.
- Sneller CH (2003) Impact of transgenic genotypes and subdivision on diversity within elite North American soybean germplasm. *Crop Sci* 43:409-414.
- Yamanaka N, Sato H, Yang Z, Xu DH, Catelli LL, Binneck E, Arias CAA, Abdelnoor RV and Nepomuceno AL (2007) Genetic relationships between Chinese, Japanese, and Brazilian soybean gene pools revealed by simple sequence repeat (SSR) markers. *Genet Mol Biol* 30:85-88.
- Zhou X, Carter Jr TE, Cui Z, Miyazaki S and Burton JW (2000) Genetic base of Japanese soybean cultivars released during 1950 to 1988. *Crop Sci* 40:1794-1802.

## Internet Resources

- Conab - National Supply Company, <http://www.conab.gov.br> (January 16, 2013).
- CultivarWeb - Ministry of Agriculture, Livestock and Supply (MAPA) [http://extranet.agricultura.gov.br/php/snpc/cultivarweb/cultivares\\_registradas.php](http://extranet.agricultura.gov.br/php/snpc/cultivarweb/cultivares_registradas.php) (September 24, 2010).
- Germplasm Resources Information Network (GRIN, 2010) <http://www.ars-grin.gov/cgi-bin/npgs/html/crop.pl?51> (September 8, 2010).

*Associate Editor: Márcio de Castro Silva Filho*

License information: This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.