Genetics and Molecular Biology, 30, 1, 85-88 (2007)

www.sbg.org.br

Short Communication

Genetic relationships between Chinese, Japanese, and Brazilian soybean gene pools revealed by simple sequence repeat (SSR) markers

Naoki Yamanaka[1,2], Hiroyuki Sato[1,3], Zhenyu Yang[4], Dong He Xu[1], Lizandra Lucy Catelli[2,5], Eliseu Binneck[2], Carlos Alberto Arrabal Arias[2], Ricardo Vilela Abdelnoor[2] and Alexandre Lima Nepomuceno[2]

[1]*Japan International Research Center for Agricultural Sciences, Tsukuba, Ibaraki, Japan.*
[2]*EMBRAPA-Soja, Londrina, PR, Brazil.*
[3]*Faculty of Horticulture, Chiba University, Matsudo, Chiba, Japan.*
[4]*Soybean Research Center, Jilin Academy of Agricultural Sciences, Gongzhuling, Jilin, China.*
[5]*Universidade Estadual de Londrina, Londrina, PR, Brazil.*

Abstract

An understanding of the relationship of geographically different soybean gene pools, based on selectively neutral DNA markers would be useful for the selection of divergent parental cultivars for use in breeding. We assessed the relationships of 194 Chinese, 59 Japanese, and 19 Brazilian soybean cultivars (n = 272) using 12 simple sequence repeat (SSR) markers. Quantification Theory III and clustering analyses showed that the Chinese and Japanese cultivars were genetically quite distant to each other but not independent, while Brazilian cultivars were distantly related to the cultivars from the other two countries and formed a cluster that was distant from the other two gene pool clusters. Our results indicated that the Brazilian soybean gene pool is different from the Chinese and Japanese pool. Exchanges of these gene pools might be useful to increase the genetic variability in soybean breeding.

Key words: DNA marker, genetic relationship, genetic resources, *Glycine max*.

Received: October 31, 2005; Accepted: July 17, 2006.

Morphological differences are usually determined by a small number of genes and may not be representative of genetic divergence in the entire genome. Regarding soybeans, an understanding of the relationships between geographically different soybean gene pools, based on selectively neutral DNA markers, would be useful for selection of divergent parental cultivars for use in soybean breeding programs. Additionally, information on soybean diversity revealed by DNA markers may help in understanding the limitations inherent in the genetic base of our breeding materials. In order to overcome restriction of the genetic base of cultivars within the regions, we utilize the genetic resources of these regions efficiently in soybean breeding programs by considering the genetic relationship as well as the agronomic value.

There have been several DNA marker studies on the genetic relationships and diversity between Japanese, Chinese, and North American soybean cultivars (Abe *et al.*, 2003; Ude *et al.*, 2003). Although these United States, China and Japan are important for soybean production and consumption and as a source of soybean genetic variation, Brazil, is a very large-scale soybean producer (either the largest or second largest) and indispensable to meeting the increasing world soybean demand. Brazilian soybean varieties are derived from a limited number of North American ancestors (Hiromoto and Vello, 1986) and since North American cultivars have a narrow genetic base (Gizlice *et al.*, 1994), it can be assumed that, compared with the Chinese and Japanese soybean gene pools, the Brazilian soybean gene pool is similar to that of North American cultivars and shares its low diversity. A study using simple sequence repeat (SSR) markers showed that 437 Brazilian commercial varieties released from 1968 to 2001 might be derived from a very limited number of soybean varieties (Catelli *et al.,* unpublished data) but the genetic relationships between Chinese, Japanese and Brazilian soybean cultivars have not been addressed yet. The objective of the research described in the present paper was to assess the relationship of these three soybean gene pools.

We investigated 194 Chinese, 59 Japanese, and 19 Brazilian soybean cultivars (n = 272) belonging to three genetic pools (Table 1). The Chinese cultivars were randomly

Send correspondence to Naoki Yamanaka. Japan International Research Center for Agricultural Sciences, 1-1 Ohwashi, Tsukuba, Ibaraki 305-8686 Japan. E-mail: naokiy@jircas.affrc.go.jp.

Table 1 - Soybean cultivars (n = 272) used in this study. Chinese cultivars with the same names are distinguished by their China national ID number in square brackets [ZDD number]. Brazilian cultivars have their alternative name in square brackets. The classification of the cultivars are shown in parenthesis as landrace (LR) or developed variety (DV).

Country & entry	Name (classification)	Country & entry	Name (classification)	Entry & country	Name (classification)
Japan (n = 59)		**China (continued)**		**China (continued)**	
1	Sakamoto wase (LR)	92	CaiZhongPu (00163) (LR)	185	ZhuYanDou (LR)
2	Tsurunoko (LR)	93	WuCangDou (LR)	186	HuangMoSiDou (LR)
3	Akita ani (LR)	94	BaiHuangDou (LR)	187	TieFen N. 3 (DV)
4	Hakusan dadacha (LR)	95	BaiQi [00181] (LR)	188	HuangDaDou (LR)
5	Kawanagare (LR)	96	DaBaiMei [00195] (LR)	189	DaHuangDou [0075] (LR)
6	Goyou daizu (LR)	97	MuLanShangDengDou (LR)	190	TieJia (LR)
7	Hakuhou (LR)	98	XiaoJinHuang [00199] (LR)	191	TianEDan (LR)
8	Kinako mame (LR)	99	JianYeDou (LR)	192	HuangTieJia (LR)
9	Egou daizu (LR)	100	ShangZiDuLuDou (LR)	193	TieJiaHei (LR)
10	Mizukuguri (LR)	101	XiaoHeiQi [00208] (LR)	194	PingDingXiang [00787] (LR)
11	Shoufuku (LR)	102	YiWoFeng (LR)	195	ZiHuaChuoZi (LR)
12	Hitashi mame (LR)	103	HuLanLiuQi (LR)	196	DaJinHuang [00792] (LR)
13	Hadaka (LR)	104	TieJiaZi [00217] (LR)	197	FuShou (LR)
14	Kosa mame (LR)	105	XiaoLiDou N. 9 (LR)	198	XiaoLiHuang [00802] (LR)
15	Chuu teppou (LR)	106	HunZaDong (LR)	199	274-2 (LR)
16	Chasengoku 81 (LR)	107	GanNanNiuMaoHuang (LR)	200	XiaoHuangQi [00808] (LR)
17	Hajinomi (LR)	108	DuLuDou [00234] (LR)	201	FengDiHuang (LR)
18	Ichigou wase (LR)	109	BaiQiDaDouWang (LR)	202	DaLi (LR)
19	fuufu daizu (LR)	110	BaiQiXiaoJinHuang (LR)	203	XiaoJinYuan (LR)
20	Hotoyoshi (LR)	111	KeSanDaJinHuang (LR)	204	KaiLiHuang (LR)
21	Sangou wase (LR)	112	HaGuang1657 (DV)	205	XiaoHeiQi [00820] (LR)
22	Chashouryu (LR)	113	GuangShiShouDaDou (LR)	206	HuangKe (LR)
23	Gin daizu (LR)	114	HeiNong N. 22 (DV)	207	RuYiDou (LR)
24	Shakkin nashi (LR)	115	Ha70-5004 (DV)	208	DaLiHuang [00830] (LR)
25	Asahiro (LR)	116	DeDuYiWoFeng (LR)	209	DaLiHuang832 (LR)
26	Fukusen nari (LR)	117	XiaoJinHuang [00280] (LR)	210	HuangDou (LR)
27	Kairyou kimusume (LR)	118	HuangQiDaDouWang (LR)	211	XiaoBaiQi [00840] (LR)
28	Hanayome β (LR)	119	SongShu N. 1 (LR)	212	6-5 (LR)
29	Hokkaihadaka (DV)	120	Gong 500 (LR)	213	DaYangDouv (DV)
30	Isuzu (DV)	121	MaoYang (LR)	214	JinZhou4-1 (LR)
31	Kitami shiro (DV)	122	QingMeShiDou (LR)	215	ShouJianDou (LR)
32	Kitamusume (DV)	123	Jilin N. 1 (DV)	216	DaJinHuang [00861] (LR)
33	Nagaha hadaka 1 (DV)	124	ZhaoFeng N. 2 (DV)	217	DaBaiQi [00864] (LR)
34	Tokachikuro (DV)	125	ZhaoFeng N. 3 (DV)	218	BaYueMang (LR)
35	Waseshiroge (DV)	126	JiTi N. 2 (DV)	219	KaiHuangDou (LR)
36	Ani (DV)	127	JiTi N. 3 (DV)	220	YouHuLu (LR)
37	Geden shirazu 1 (DV)	128	JiTi N. 4 (DV)	221	DaJinYan (LR)
38	Hatsukari (DV)	129	QuenXuan N. 1 (DV)	222	XiaoJinHuang [00880] (LR)
39	Okushirome (DV)	130	ManChangJin [00376] (DV)	223	ErLiHuang (LR)
40	Shirome nagaha (DV)	131	HuangBaoZhu (DV)	224	TieJiaBaoQi (LR)
41	Dekisugi 1 (DV)	132	YuangBaoJin (DV)	225	DongDaLi (LR)
42	Ibaragi mame 7 (DV)	133	GuoYu B4 (DV)	226	DaBaiQi [00939] (LR)
43	Shoromeyutaka β (DV)	134	BuoYu B5 (DV)	227	164-4-19 (LR)
44	Tamahomare (DV)	135	GuoYu98 (DV)	228	DaBaiQi [00945] (LR)
45	Kokeshijiro (DV)	136	Ha N. 1 (LR)	229	HeiQi [00947] (LR)
46	Shinmejiro (DV)	137	GongJiao 5013-12-2 (DV)	230	XiaoJinHuang [00948] (LR)
47	Tachisuzunari (DV)	138	ShiLiHuang [00424] (LR)	231	QingDou (LR)
48	Hime daizu (DV)	139	ShiLiHuang [00427] (LR)	232	274-1 (LR)
49	Tamanishiki (DV)	140	ShiLiQi (LR)	233	Jin8-14 (LR)
50	Akisengoku (DV)	141	DaLiHuang [00438) (LR)	234	DaGaoTui (LR)
51	Akishirome (DV)	142	XiaoLiHuang [00450] (LR)	235	DaBaiMei[(01011] (LR)
52	Asomasari (DV)	143	XiaoBaiDou (LR)	236	DaLiZi (LR)
53	Shirosaya 1 (DV)	144	SuoYiLing [0046]) (LR)	237	DaBaiQi [01048] (LR)
54	Hyuuga (DV)	145	SuoYiLing [00465] (LR)	238	BaiQi [01059] (LR)
55	Fukuyutaka (DV)	146	ZhiHuaChuoZhi [00467] (LR)	239	HuangQi (LR)
56	Higomusume (DV)	147	ZhiHuaChuoZhi [00468] (LR)	240	XiaoBaiQi [01075] (LR)
57	Kogane daizuv (DV)	148	DuLuDou [00483] (LR)	241	DaHuangDou (01098) (LR)
58	Matsuura (DV)	149	DongFengDuLuDou (LR)	242	YangHuangDou (LR)
59	Houchiouv (DV)	150	DuLuDou [00489] (LR)	243	WanDouSou [01107] (LR)
		151	XiAnDuLuDou (LR)	244	WanDouSou [01108] (LR)
China (n = 194)		152	TieJiaSiLiHuang (LR)	245	HeiQi [01121] (LR)
60	MoshidouGong 503 (DV)	153	TieJiaZi [00506] (LR)	246	XiaoHuangQi [01122] (LR)
61	HeiNong N. 3 (DV)	154	XiaoHeiQi [00508] (LR)	247	6611 (DV)
62	HeiNong N. 4 (DV)	155	XiaoLanQI (LR)	248	QingPi (LR)
63	HeiNong N. 6 (DV)	156	PingDingXiang [00522] (LR)	249	XiaoQingDou (LR)
64	HeiNong N. 17 (DV)	157	BaiLuDou (LR)	250	PingDingXiang [12012] (LR)
65	FengShou N. 3 (DV)	158	DaHeiGan (LR)	251	BaoPiQing (LR)
66	FengShou N. 4 (DV)	159	BengPi (LR)	253	DaQingDou (LR)
67	FengShou N. 13 (DV)	160	ZhaQiDou (LR)		
68	HeFeng N. 5 (DV)	161	HouDingKui (LR)	**Brazil (n = 19)**	
69	HeFeng N. 16 (DV)	162	HuiTieJia (LR)	254	Bossier (DV)
70	MuFeng N. 4 (DV)	163	DongFengJinYuan (LR)	255	BR14 [Modelo] (DV)
71	NenFeng N. 4 (DV)	164	CaiZhongPu [00552] (LR)	256	BRS153 (DV)
72	NenLiang N. 7 (DV)	165	PingDingXiang [00561] (LR)	257	BRSMS Tuiuiú (DV)
73	ManChangJin [00078] (LR)	166	DaJinHuang [00570] (LR)	258	CD208 (DV)
74	XiBiWa (LR)	167	BaiHuChou [00573] (LR)	259	CD209 (DV)
75	ZaoTieJiaQing (DV)	168	BaiHuChou [00575] (LR)	260	EMBRAPA20 [DokoRC] (DV)
76	DaLiDou (LR)	169	BianDaDou (LR)	261	Emgopa311 (DV)
77	PingJinDing (LR)	170	HeQiFengDiHuang (LR)	262	FT16 (DV)
78	AiHuiBenDiZhong (LR)	171	XiaoYangDou (LR)	263	FT101 (DV)
79	AnDaBaiMei (LR)	172	HeQiPingDingXiang (LR)	264	FT2000 (DV)
80	NenJiangPingDingXiang (LR)	173	YiTongManCangJi (LR)	265	IAS3(Delta) (DV)
81	TongBeiXiaoJinHuang (LR)	174	YuShuManChangJi (LR)	266	IAS5 [Vagemescura] (DV)
82	XiaoBaiMei (LR)	175	DaNiuMaoHuang (LR)	267	MT/BR50 [Parecis] (DV)
83	XiaoDou (LR)	176	NiuMaoHuang [00621] (LR)	268	Suprema (DV)
84	SiLiJin (LR)	177	LianQi (LR)	269	UFV3 (DV)
85	RiBenDi (LR)	178	NiuMaoHuang [00649] (LR)	270	UFV8 [MonteRio] (DV)
86	QingGang9-1 (LR)	179	NiuMaoHuang [00652] (LR)	271	UFV10 [Uberaba] (DV)
87	AnDa37-1 (LR)	180	NiuMaoHuang [00653] (LR)	272	UFV19 [Triângulo] (DV)
88	DaTieJiaoQing (LR)	181	LuPiDou (LR)		
89	MiShanTieJiaQing (LR)	182	XiaoQiDouv		
90	ZhaoDong50 (LR)	183	QingZaDou (LR)		
91	XiaoBaiPi (LR)	184	HeiDou (LR)		

selected from the Northeast China soybean genetic resources database and included soybean landraces and developed varieties which have been well characterized in relation to their main traits (Yamanaka and Okabe, 2005). The Japanese cultivars were selected from the Ministry of Agriculture, Forestry and Fisheries of Japan (MAFF) Genebank according to the geographic regions where the developed varieties and landraces originated. The Brazilian cultivars were selected from an unweighted pair group method using arithmetic averages (UPGMA) dendrogram previously constructed with 22 SSR markers and 437 Brazilian cultibars released from 1968 to 2001 (Catelli et al. unpublished data). To reflect the genetic diversity of this gene pool, the Brazilian soybean cultivars used in the present study were chosen by randomly selecting one cultivar from each of 19 clusters (coefficient of similarity = 0.44 for cluster assessment) in the UPGMA dendrogram. For analysis we selected 12 SSR markers (Table 2) developed by Cregan *et al.* (1999) and carried out SSR amplification and the gel electrophoresis according to the methods of Hossain *et al.* (2000). The sizes of the amplified bands were calculated for every ten base-pairs and the alleles for each SSR marker were then decided according to the band sizes in all the cultivars and then used for analysis. Two types of analyses were used to evaluate the genetic relationships between the three gene pools. The data was analyzed by Hayashi's Quantification Theory III (QT III; Hayashi, 1956), which quantified the category data (*e.g.* haplotype data) based on the allelic patterns obtained with the 12 SSR markers to schematically represent the relationships among the cultivars and classify them. This type of analysis corresponds to principal component analysis (PCA) using the values of interval or ratio scales and similar results can be obtained by the Correspondence analysis (CA) or Dual scaling (DS).

Table 2 - The 12 simple sequence repeat (SSR) markers used in this study. The number of alleles, size range, and diversity index were calculated using all the cultivars from the three germplasm sources.

SSR marker	Map position[1]	N. of alleles	Range (bp)	Diversity index $(H)^2$ $(H)^1$
Satt365	C2	11	60-350	0.776
Satt489	C2	7	210-270	0.783
Satt153	O	6	170-220	0.620
Satt373	L	12	180-300	0.778
Satt597	B1	3	140-160	0.478
Satt513	L	7	90-170	0.579
Satt545	A1	8	130-210	0.792
Satt587	I	5	140-180	0.222
Sct_065	J	4	110-160	0.405
Satt599	A1	4	150-170	0.659
Satt299	L	9	170-250	0.836
Satt174	A1	6	120-180	0.716

[1]Map positions of SSR markers were obtained from Cregan *et al.* (1999).
[2]Diversity index (H) was calculated for each SSR marker as $1 - \Sigma p_{ij}^2$, where p_{ij} is the frequency of the *j*th allele of marker *i*.

QT III analysis was carried out by the Japanese web-based Black-Box statistics service (http://aoki2.si.gunma-u.ac.jp/BlackBox/BlackBox.html. as of 11/04/2005). In addition, UPGMA cluster analysis was performed with the PHYLIP computer program (Felsenstein, 1989). For cluster analysis, the genetic distance was calculated using the MSAT2 computer (Minch *et al.*, 1997) and the 1 - *P* distance measurement (where *P* is the proportion of shared alleles for the 12 SSR markers) and 1,000 bootstrap re-samplings before subjecting the genetic matrix to UPGMA cluster analysis.

All 12 SSR markers produced amplified bands in all 272 cultivars, with the number of alleles ranging from 3 to11 (Table 2). The diversity index (H) of the SSR markers ranged from 0.222 to 0.836, indicating that these markers present large differences in information for characterization of the cultivars studied. As a result, these SSR markers could distinguish 261 cultivars, *i.e.* more than 95% of the cultivars could be identified.

In both QT III and cluster analyses, the Chinese and Japanese cultivars were not classified as independent but these two groups were quite distant from each other. However, Brazilian cultivars were distantly related to the Chinese and Japanese cultivars and formed a cluster that was distant from the gene pools from Japan and China (Figures 1 and 2). Cluster analysis showed seven clusters, with the first cluster containing only Japanese and Chinese cultivars, the second to the sixth cluster Chinese and two Japanese cultivars, and seventh cluster all the Brazilian cultivars (Figure 2). In the cluster analysis the distances between cultivars belonging to the three gene pools was 0 to 0.691 for both Chinese and Japanese cultivars but only 0 to 0.520 for Brazilian cultivars. Abe *et al.* (2003) have shown that the genetic relationship revealed by SSR markers corresponds well to the geographical separation between Japanese and Chinese soybean gene pools and classified these

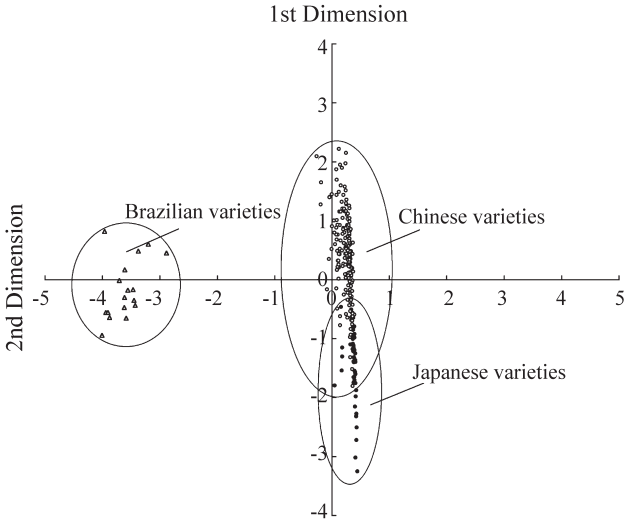

Figure 1 - Two-dimension Quantification Theory III analysis scatter plot of 272 soybean cultivars. Circles indicate the distributions of the three genetic resources.

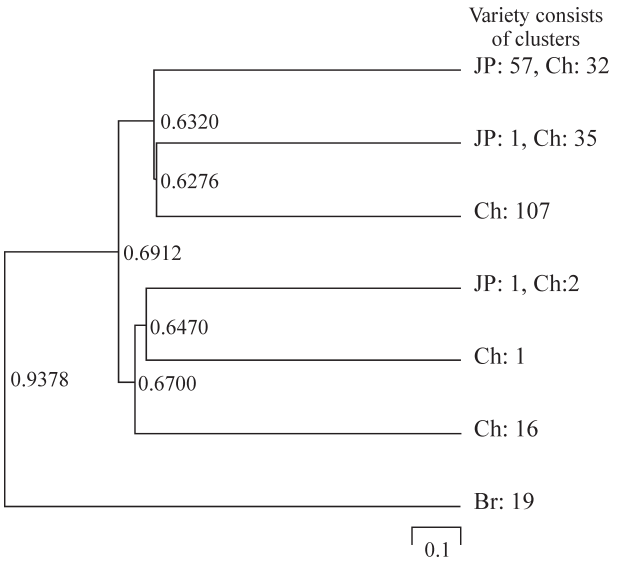

Figure 2 - Unweighted pair group method with averages (UPGMA) dendrogram constructed for the 272 soybean cultivars used in this study. The genetic distance and absolute Bootstrap value from 1,000 replications at each branching point are also shown. Each cluster was formed based on the genetic distances being equal to 0.600, and was labeled the origin (Jp = Japan, Ch = China and Br = Brazil) and the number of cultivars present in the cluster.

gene pools as independent, which was not the case in our study – possibly because of the different materials used. However, we also showed that the Brazilian varieties were distantly related and classified in a completely different group from the Japanese and Chinese gene pools and, furthermore, the genetic diversity within each gene pool was lower than that observed among different gene pools. A similar phenomenon was reported by Ude *et al.* (2003) who showed that for Japanese, Chinese, North American soybean cultivars and North American ancestral lines, the genetic distances between different regional gene pools was greater than the variation within gene pools. Thus, geographically different genetic resources of soybean are quite different genetically even though they are not completely different. A DNA marker study by Thompson and Nelson (1998) revealed that introgression of genetic diversity from exotic genetic resources can contribute to increasing the yield of current USA cultivars. Therefore it can also be expected that exchanging soybean genetic resources between Japan, China and Brazil would expand their genetic base and increase variability especially when Brazilian varieties are used for soybean breeding in Japan and China or when Chinese cultivars are used in breeding programs in Brazil. These findings are useful as selection criteria to determine parental cultivars in addition to considering only the agronomic characteristics of the cultivar. Furthermore, the data presented in this paper are useful for the identification of cultivars and for checking the introduction of gene segments into breeding lines when the cultivars described in this paper are used in breeding programs.

Acknowledgments

This study was supported by JIRCAS international projects, “Comprehensive Studies on Soybean Improvement, Production and Utilization in South America” and “Development of Stable Production and Utilization of Major Food Resources in China”.

References

Abe J, Xu DH, Suzuki Y, Kanazawa A and Shimamoto Y (2003) Soybean germplasm pools in Asia revealed by nuclear SSRs. Theor Appl Genet 106:445-453.

Cregan PB, Jarvik T, Bush AL, Shoemaker RC, Lark KG, Kahler AL, Kaya N, VanToai TT, Lohnes DG, Chung J and Specht JE (1999) An integrated genetic linkage map of the soybean. Crop Sci 39:1464-1490.

Felsenstein J (1989) PHYLIP - Phylogeny Inference Package (Version 3.2). Cladistics 5:164-166.

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

Hayashi C (1956) Theory and examples of quantification III. Proc Institute Statistical Math 4:19-30.

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

Hossain KG, Kawai H, Hayashi M, Hoshi M, Yamanaka N and Harada K (2000) Characterization and identification of (CT)n microsatellites in soybean using sheared genomic libraries. DNA Res 7:103-110.

Thompson JA and Nelson RL (1998) Utilization of diverse germplasm for soybean yield improvement. Crop Sci 38:1362-1368.

Ude GN, Kenweorthy 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.

Yamanaka N and Okabe A (2005) Evaluation and utilization of soybean genetic resources in Northeast China and development of novel soybean breeding materials using new technology. JIRCAS Working Report N. 42:77-87.

Internet Resources

Minch E, Ruiz-Linares A, Goldstein D, Feldman M and Cavalli-Sforza LL (1997) Microsat: A computer program for calculating various statistics on microsatellite allele data. http://hpgl.stanford.edu/projects/microsat/ (20/04/2006).

Ministry of Agriculture, Forestry and Fisheries of Japan (MAFF) Genebank, http://www.gene.affrc.go.jp/index.html (25/4/2002).

The database of soybean genetic resources in Northeast China, http://www.jircas.affrc.go.jp/DB/guide-eng.html (19/11/2004).

Associate Editor: Everaldo Gonçalves de Barros