



BMT/12/21

ORIGINAL: English

DATE: May 7, 2010

INTERNATIONAL UNION FOR THE PROTECTION OF NEW VARIETIES OF PLANTS
GENEVA

**WORKING GROUP ON BIOCHEMICAL AND MOLECULAR
TECHNIQUES AND DNA PROFILING IN PARTICULAR**

Twelfth Session
Ottawa, Canada, May 11 to 13, 2010

**USE OF A MOLECULAR MARKER BASED SYSTEM FOR IDENTIFICATION OF
VARIETIES IN BRAZIL**

Document prepared by an expert from Brazil



WORKING GROUP ON BIOCHEMICAL AND MOLECULAR TECHNIQUES AND DNA PROFILING IN PARTICULAR

Twelfth Session

Ottawa, Canada, May 11 to 13, 2010

Use of a molecular marker-based system
for identification of varieties in Brazil

Soybean and Rice

Document prepared by :

Luís Pacheco

SNPC/MAPA - Brazil



PVP in Brazil

- ✓ Breeder Testing System
- ✓ PVP Office – technical staff:
 - 1 Coordinator
 - 6 ExaminersExamination of applications and Granting of Plant Breeder's Rights
- ✓ 1 Laboratory – Live Samples Storage (Seed or DNA)
 - 2 Experts
- ✓ Independent Test Lab
 - DNA fingerprints of PBR and NLi varieties
 - Support for identification of varieties



Identification of Genetic Profiles of SSR markers in DNA

- ✓ Objectives - **Identification** of varieties
Enforcement of Seed Law (cooperative work within Departments of MAPA)
Seed Certification
PBR and NLi Varieties
Comparison of “new” and “old” varieties
Post control
- ✓ Control samples
SNPC provided **coded** samples
Doubled samples from SNPC
Negative Control – water / Positive Control – laboratory control sample
2 DNA extractions (bulks of 50 young leaves)
Genetic Analysis by 2 different staff in different days – minimize human error
- ✓ SSR Markers
Highly informative
Extensively validated in scientific literature



Statistical Analysis

- ✓ Loci polymorphism analysed through PCR
Primers marked with fluorescence blue (FAM), green (HEX) and yellow (NED)
- ✓ Alleles Detection
High resolution Capillary Electrophoresis
DNA automated sequencing – ABI Prism 3100
- ✓ Size of Alleles
Estimated by algorithm “Local Southern” – Software Genotyper
Discrete allele sizes - Least Squares Minimization Algorithm – Allelobin
Di, tri and tetranucleotides – variations of 1,5 pb between different runs and
0,5 pb in the same run – positive control in all tests allows to identify deviations
- ✓ Genetic Similarity
Genetic distance between pairs of entries – NTSYSpc 2.10z
Diagonal Matrix of Genetic Distances – UPGMA
Dendrogram of Genetic Distances - NTSYSpc 2.10z



Soybean and Rice - Importance in Brazil

❖ **Soybean** - most important agricultural commodity

Area – 27,1 million ha

Yield – 57,1 million tons

745 varieties in NLI

543 protected by PBR:

~ **50** new varieties protected/year

~ **1/3** of the titles granted in 1 year by SNPC

❖ **Rice** – one of most important foods

Area – 2,9 million ha

Yield – 12,6 million tons

231 varieties in NLI

77 protected by PBR





Numbers of Samples x markers

Soybean: 556 Samples x 15 SSR loci

Rice: 79 Samples x 16 SSR loci



SOYBEAN



SSR Markers - Soybean

- BARC-Satt002 Sequência simples repetitiva de trinucleotídeo; grupo de ligação D2
- BARC-Satt005 Sequência simples repetitiva de trinucleotídeo; grupo de ligação D1
- BARC-Satt042 Sequência simples repetitiva de trinucleotídeo; grupo de ligação A1
- BARC-Satt045 Sequência simples repetitiva de trinucleotídeo; grupo de ligação E
- BARC-Satt038 Sequência simples repetitiva de trinucleotídeo; grupo de ligação G
- BARC-Satt030 Sequência simples repetitiva de trinucleotídeo; grupo de ligação F
- BARC-Satt009 Sequência simples repetitiva de trinucleotídeo; grupo de ligação N1
- BARC-Satt70 Sequência simples repetitiva de trinucleotídeo; grupo de ligação B2
- BARC-Satt100 Sequência simples repetitiva de trinucleotídeo; grupo de ligação C2
- BARC-Satt114 Sequência simples repetitiva de trinucleotídeo; grupo de ligação F
- BARC-Satt586 Sequência simples repetitiva de trinucleotídeo; grupo de ligação F1
- S45035 "Kunitz trypsin inhibitor gene" SSR de dinucleotídeo; inform. de mapa indisponível
- GMABAB - sequência simples repetitiva de dinucleotídeo; grupo de ligação N1
- BARC-Satt431 Sequência simples repetitiva de trinucleotídeo; grupo de ligação J
- BARC-Satt216 Sequência simples repetitiva de trinucleotídeo; grupo de ligação D1

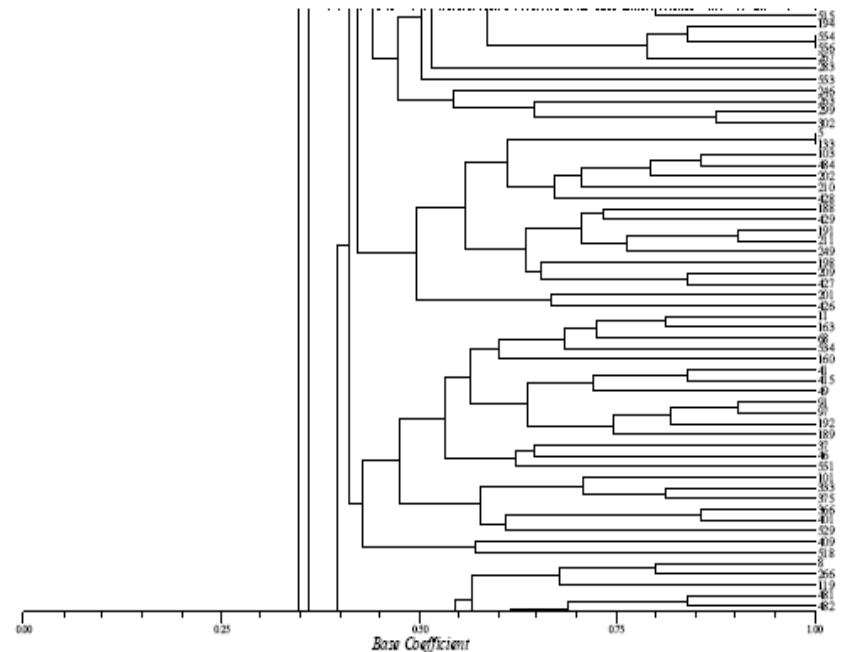
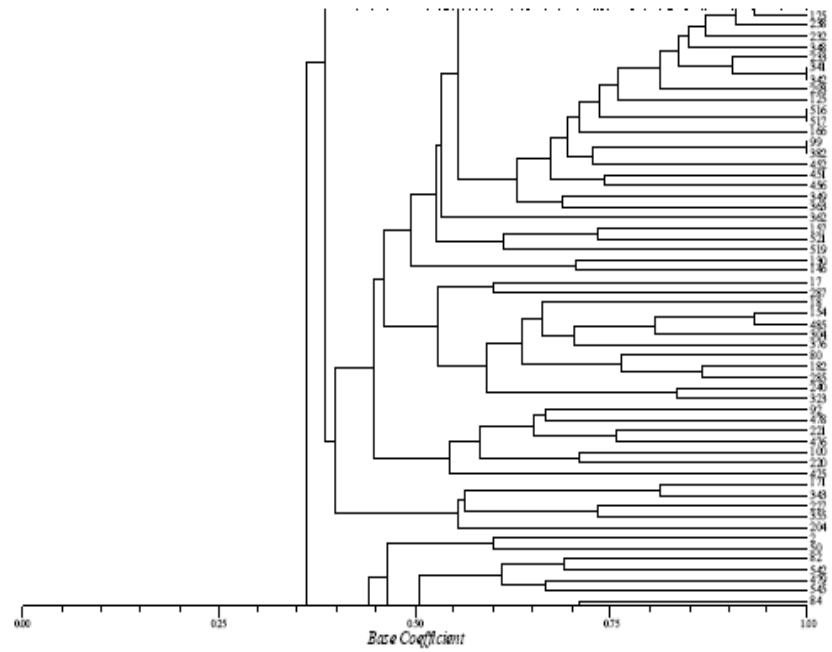
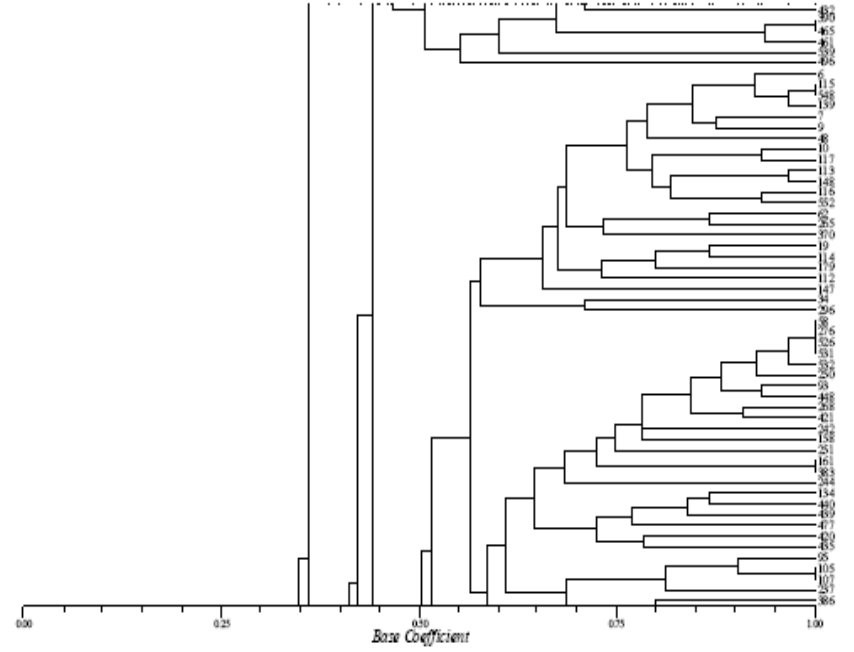
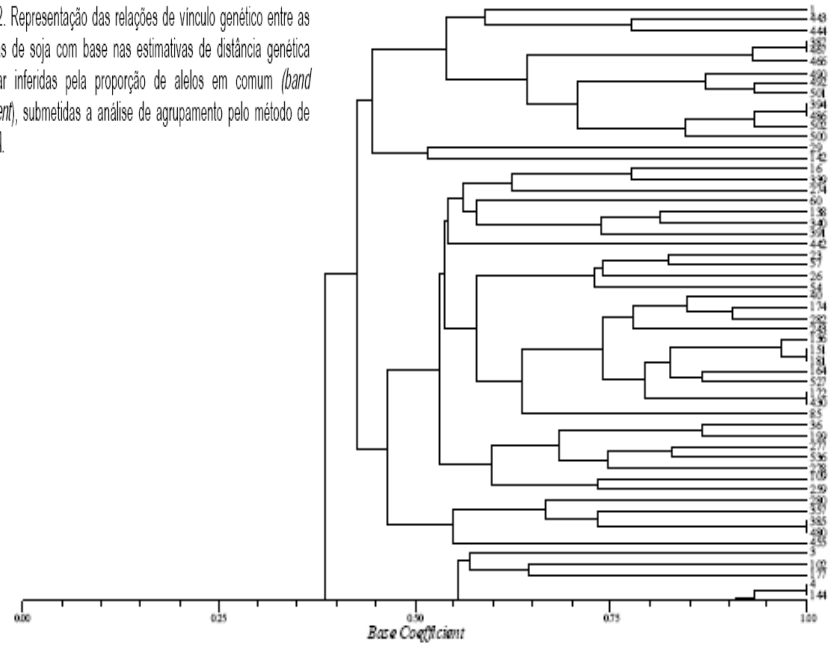


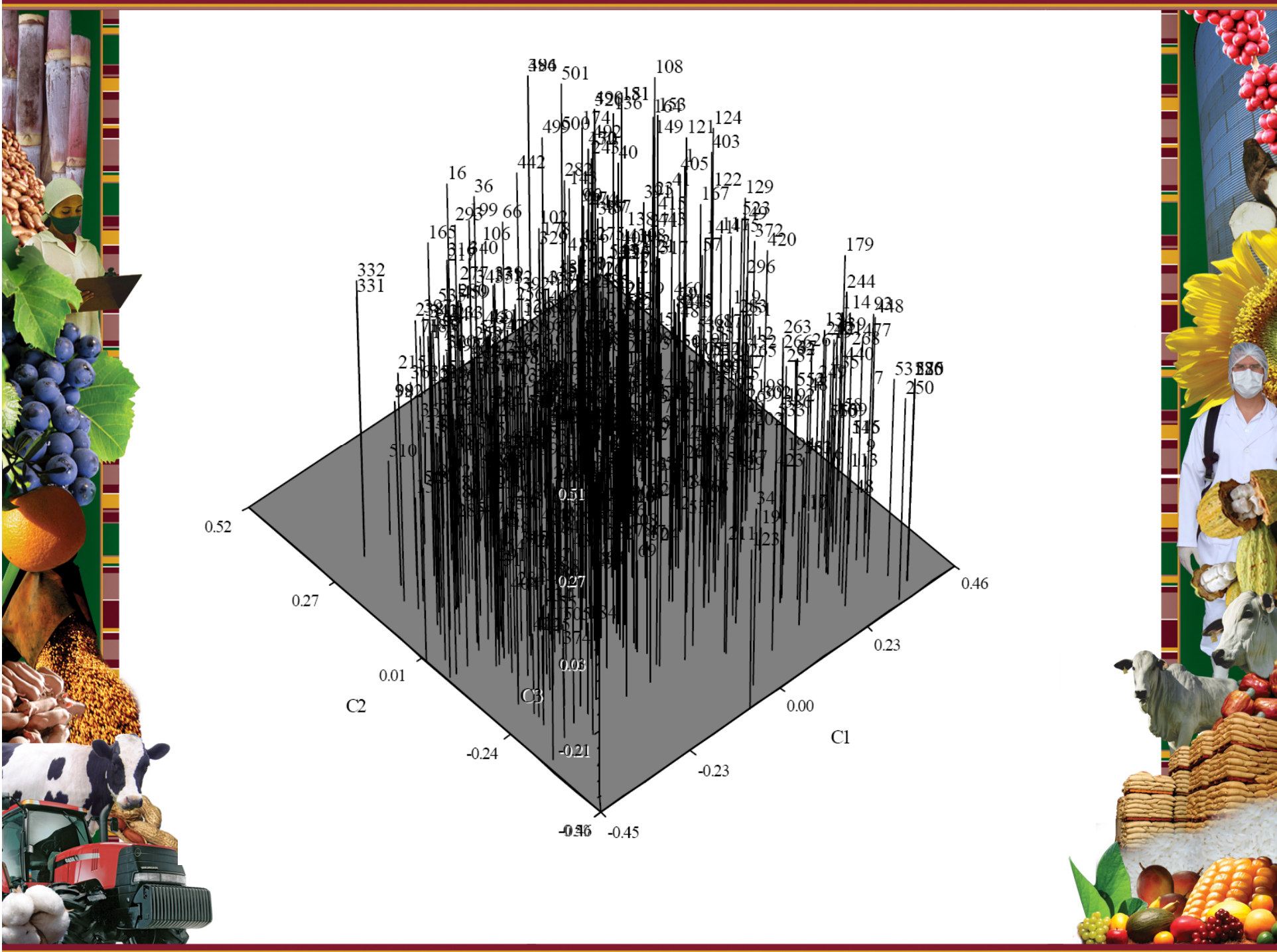
Tabela 1. Perfil genético multiloco de amostras de soja fornecidas pelo LADIC/MAPA, analisadas em 15 locos microssatélites. São apresentados os valores em pares de base (pb) de cada um dos dois alelos observados em cada loco nas amostras analisadas, após conferência dos genótipos em regime de análise de prova e contraprova.

SATT005	1	181.58	181.58	SATT038	1	168.17	168.17	SATT586	1	216.64	216.64
SATT005	2	178.51	178.51	SATT038	2	171.20	171.20	SATT586	2	203.40	203.40
SATT005	3	156.86	156.86	SATT038	3	168.05	168.05	SATT586	3	210.15	210.15
SATT005	4	156.90	156.90	SATT038	4	168.20	168.20	SATT586	4	199.84	199.84
SATT005	5	153.88	153.88	SATT038	5	168.14	168.14	SATT586	5	210.25	210.25
SATT005	6	178.51	178.51	SATT038	6	168.14	168.14	SATT586	6	210.23	210.23
SATT005	7	153.85	153.85	SATT038	7	168.26	168.26	SATT586	7	210.30	210.30
SATT005	8	153.81	153.81	SATT038	8	168.23	168.23	SATT586	8	196.88	196.88
SATT005	9	153.74	153.74	SATT038	9	168.03	168.03	SATT586	9	203.33	210.09
SATT005	10	156.89	156.89	SATT038	10	168.04	168.04	SATT586	10	210.18	210.18
SATT005	11	153.67	153.67	SATT038	11	171.18	171.18	SATT586	11	210.21	210.21
SATT005	12	156.92	156.92	SATT038	12	168.10	168.10	SATT586	12	210.23	210.23
SATT005	13	153.82	153.82	SATT038	13	168.07	168.07	SATT586	13	235.74	235.74
SATT005	14	154.32	154.32	SATT038	14	168.18	168.18	SATT586	14	236.39	236.39
SATT005	15	156.98	156.98	SATT038	15	168.14	168.14	SATT586	15	210.27	210.27
SATT005	16	144.28	144.28	SATT038	16	168.29	168.29	SATT586	16	200.12	200.12
SATT005	17	156.99	156.99	SATT038	17	171.37	171.37	SATT586	17	235.81	235.81
SATT005	18	156.91	156.91	SATT038	18	171.13	171.13	SATT586	18	199.98	210.04
SATT005	19	153.85	153.85	SATT038	19	168.00	168.00	SATT586	19	210.13	210.13
SATT005	20	153.85	153.85	SATT038	20	168.03	168.03	SATT586	20	203.24	203.24
	⋮				⋮				⋮		
SATT005	550	156.95	156.95	SATT038	550	167.57	167.57	SATT586	550	235.42	235.42
SATT005	551	157.03	157.03	SATT038	551	188.94	188.94	SATT586	551	209.99	209.99
SATT005	552	178.97	178.97	SATT038	552	167.58	167.58	SATT586	552	209.98	209.98
SATT005	553	135.46	135.46	SATT038	553	167.58	167.58	SATT586	553	199.83	199.83
SATT005	554	153.65	153.65	SATT038	554	167.60	167.60	SATT586	554	203.23	203.23
SATT005	555	153.80	153.80	SATT038	555	167.61	167.61	SATT586	555	196.61	196.61
SATT005	556	153.68	153.68	SATT038	556	167.61	167.61	SATT586	556	203.23	203.23
SATT005	Gen	156.88	156.88	SATT038	Gen	171.19	171.19	SATT586	Gen	235.68	235.68
SATT100	1	141.72	141.72	SATT009	1	217.78	217.78	SATT431	1	231.48	231.48
SATT100	2	141.70	141.70	SATT009	2	239.87	239.87	SATT431	2	231.54	231.54
SATT100	3	138.59	138.59	SATT009	3	221.10	221.10	SATT431	3	231.59	231.59
	⋮				⋮				⋮		



Figura 2. Representação das relações de vínculo genético entre as amostras de soja com base nas estimativas de distância genética par-a-par inferidas pela proporção de alelos em comum (*band coefficient*), submetidas a análise de agrupamento pelo método de UPGMA.





Results - Soybean

- ✓ No subdivisions of samples in clusters
clear genetic similarities between some groups of samples BUT
apparent narrow genetic base of brazilian soybean germplasm limits
precise reconstruction of relations of genetic linkages of all accessions
- ✓ Group of selected markers allowed differentiation of samples
- ✓ Soybean: autogamous and diploid - bulk samples
genotype in each loci is typically homozygote (one allele duplicated)
When two distinct alleles are observed: evidence of residual heterozygosity or
mixed lines – in this case, additional tests are needed
- ✓ Genetic Similarity - conclusions (based on the multiloci genetic profiles)
525 unique genetic profiles – 94,77% of samples
Undistinguishable samples
Very similar samples
Evaluated case by case – EDVs, parentals, siblings, etc.



RICE



SSR Markers - Rice

- RM252 Sequência repetitiva em tandem de dinucleotídeo CT; cromossomo 4
- RM224 Sequência complexa repetitiva em tandem do tipo (AAG)₈(AG)₁₃; cromos. 11
- OG44 Sequência repetitiva em tandem de dinucleotídeo AG; cromossomo 3
- RM248 Sequência repetitiva em tandem de dinucleotídeo CT; cromossomo 7
- OS19 Sequência repetitiva em tandem de dinucleotídeo AG; cromossomo 6
- OG106 Sequência repetitiva em tandem de dinucleotídeo AG; cromossomo 9
- OG81 Sequência repetitiva em tandem de dinucleotídeo AG; cromossomo 1
- RM263 Sequência repetitiva em tandem de dinucleotídeo CT; cromossomo 2
- RM420 Sequência repetitiva em tandem de tetranucleotídeo AAAT; cromossomo 7
- RM418 Sequência repetitiva em tandem de trinucleotídeo ATT; cromossomo 7
- RM475 Sequência repetitiva em tandem de tetranucleotídeo TACT; cromossomo 2
- RM171 Sequência repetitiva em tandem de tetranucleotídeo GATG; cromossomo 10
- RM481 Sequência repetitiva em tandem de dinucleotídeo CAA; cromossomo 7
- OG61 Sequência repetitiva em tandem de dinucleotídeo AG; cromossomo 6
- RM335 Sequência repetitiva em tandem de trinucleotídeo CTT; cromossomo 4
- RM105 Sequência repetitiva em tandem de trinucleotídeo CCT; cromossomo 9

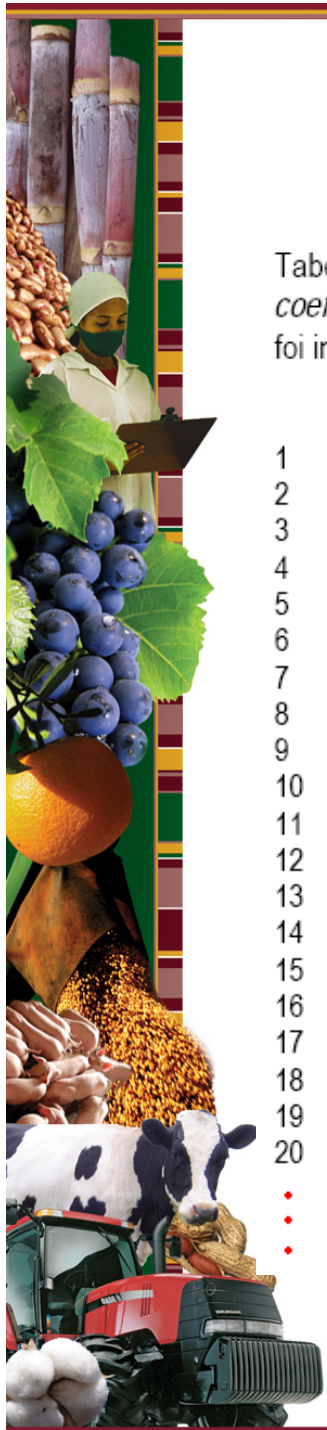
Tabela 1. Perfil genético multiloco de amostras de arroz fornecidas pelo LADIC/MAPA, analisadas em 16 locos microssatélites. São apresentados os valores em pares de base (pb) de cada um dos dois alelos observados em cada loco nas amostras analisadas, após conferência dos genótipos em regime de análise de prova e contraprova.

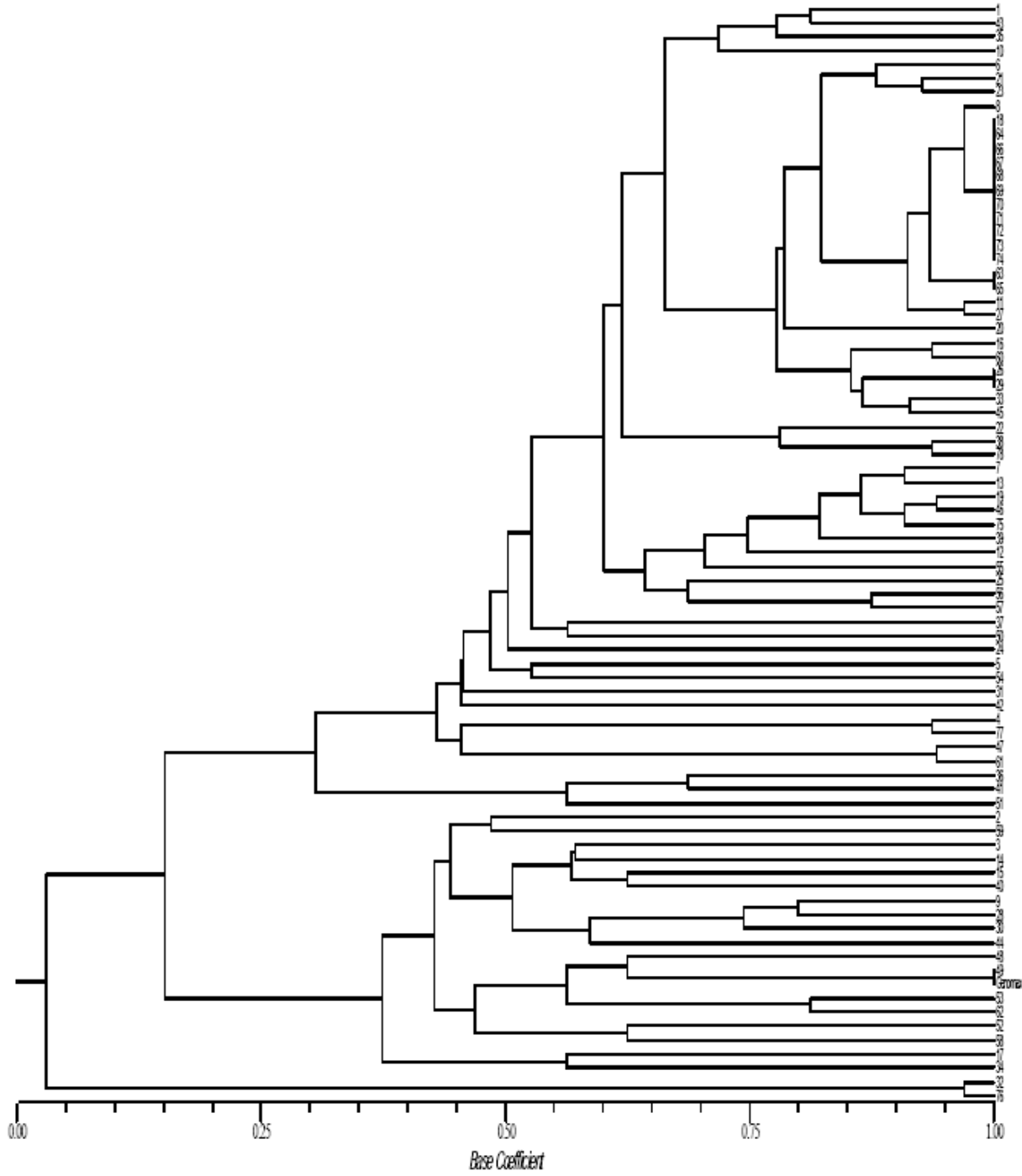
Loco	LADIC	alelo 1	alelo 2	Loco	LADIC	alelo 1	alelo 2
RM248	1	99.83	99.83	RM335	1	113.71	113.71
RM248	2	83.36	83.36	RM335	2	136.86	136.86
RM248	3	83.41	83.41	RM335	3	158.61	158.61
RM248	4	93.65	93.65	RM335	4	99.93	99.93
RM248	5	99.82	99.82	RM335	5	146.01	146.01
RM248	6	89.69	99.86	RM335	6	113.55	113.55
RM248	7	99.86	99.86	RM335	7	99.67	99.67
RM248	8	89.55	89.55	RM335	8	146.12	146.12
RM248	9	83.10	93.63	RM335	9	108.11	108.11
RM248	10	99.83	99.83	RM335	10	113.69	113.69
RM248	11	89.45	89.45	RM335	11	146.23	146.23
RM248	12	81.05	99.79	RM335	12	99.93	146.26
RM248	13	89.46	89.46	RM335	13	99.58	99.58
RM248	14	83.42	83.42	RM335	14	158.63	158.63
RM248	15	83.38	83.38	RM335	15	155.42	155.42
RM248	16	89.53	89.53	RM335	16	145.98	145.98
RM248	17	99.78	99.78	RM335	17	140.02	140.02
RM248	18	89.68	89.68	RM335	18	146.18	146.18
RM248	19	89.64	89.64	RM335	19	99.70	99.70
RM248	20	89.63	89.63	RM335	20	146.18	146.18
	⋮				⋮		
RM248	69	89.69	89.69	RM335	69	146.17	146.17
RM248	70	89.67	89.67	RM335	70	146.10	146.10
RM248	71	89.71	89.71	RM335	71	146.09	146.09
RM248	72	89.65	89.65	RM335	72	146.17	146.17
RM248	73	89.68	89.68	RM335	73	146.16	146.16
RM248	74	89.64	89.64	RM335	74	146.18	146.18
RM248	75	99.85	99.85	RM335	75	99.61	99.61
RM248	76	79.13	79.13	RM335	76	139.88	139.88
RM248	77	93.79	93.79	RM335	77	99.68	99.68
RM248	78	89.69	89.69	RM335	78	99.69	99.69
RM248	Genomax	93.69	93.69	RM335	Genomax	108.12	108.12
OG44	1	150.17	150.17	RM420	1	184.03	184.03
OG44	2	163.01	163.01	RM420	2	183.87	183.87
OG44	3	150.29	150.29	RM420	3	184.05	184.05

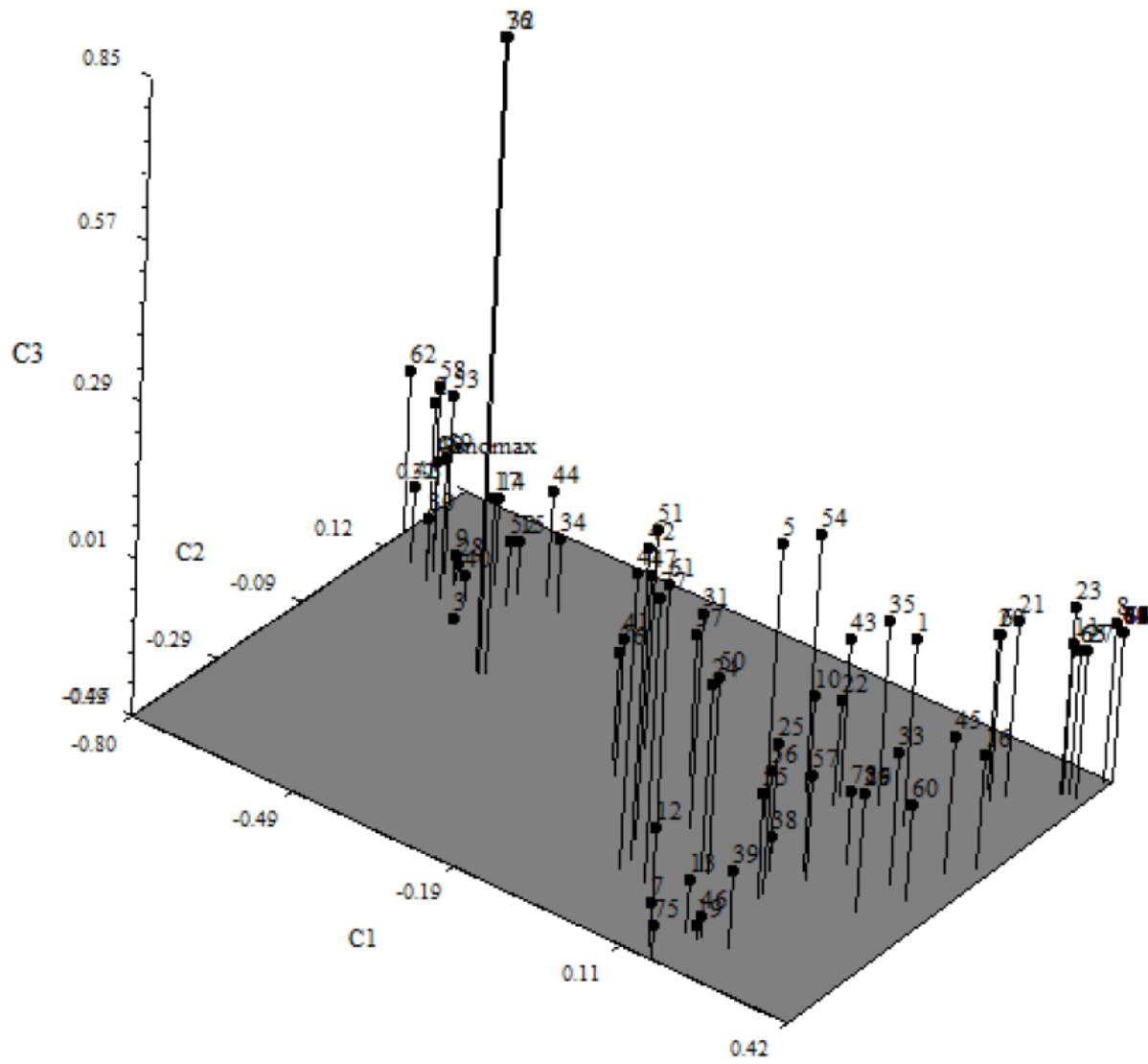


Tabela 2. Estimativas de distância genética par-a-par com base na proporção de alelos em comum (*band coefficient*) entre as 78 amostras de arroz utilizadas na análise. A amostra de arroz “Genomax” também foi incluída na análise.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
1	1.00																							
2	0.06	1.00																						
3	0.23	0.40	1.00																					
4	0.44	0.06	0.11	1.00																				
5	0.50	0.13	0.11	0.50	1.00																			
6	0.85	0.06	0.17	0.30	0.42	1.00																		
7	0.55	0.06	0.28	0.42	0.42	0.53	1.00																	
8	0.67	0.06	0.17	0.36	0.55	0.76	0.53	1.00																
9	0.17	0.40	0.53	0.11	0.11	0.17	0.22	0.17	1.00															
10	0.74	0.06	0.21	0.40	0.46	0.72	0.61	0.61	0.16	1.00														
11	0.67	0.06	0.17	0.36	0.48	0.82	0.53	0.88	0.17	0.61	1.00													
12	0.46	0.10	0.24	0.41	0.46	0.45	0.75	0.50	0.19	0.57	0.50	1.00												
13	0.44	0.06	0.23	0.44	0.38	0.48	0.91	0.55	0.17	0.51	0.55	0.67	1.00											
14	0.19	0.56	0.57	0.06	0.13	0.18	0.18	0.18	0.46	0.17	0.18	0.15	0.13	1.00										
15	0.13	0.50	0.57	0.06	0.13	0.18	0.24	0.24	0.51	0.17	0.18	0.21	0.19	0.56	1.00									
16	0.69	0.06	0.23	0.44	0.56	0.73	0.67	0.85	0.17	0.69	0.85	0.62	0.69	0.19	0.19	1.00								
17	0.19	0.31	0.29	0.06	0.25	0.24	0.30	0.18	0.46	0.23	0.18	0.26	0.25	0.31	0.50	0.19	1.00							
18	0.69	0.06	0.17	0.38	0.50	0.79	0.55	0.97	0.17	0.63	0.91	0.51	0.56	0.19	0.25	0.88	0.19	1.00						
19	0.48	0.06	0.28	0.42	0.36	0.53	0.88	0.59	0.22	0.61	0.59	0.75	0.91	0.18	0.24	0.73	0.24	0.61	1.00					
20	0.63	0.06	0.23	0.38	0.50	0.67	0.48	0.79	0.23	0.63	0.79	0.46	0.50	0.19	0.25	0.81	0.13	0.81	0.55	1.00				







Results - Rice

- ✓ Subdivisions of samples in two clusters
probably subgroups indica (irrigated) and japonica (dry lands),
except 2 samples (LADIC 32 and 76) - hybrids between the two subgroups?
- ✓ Group of selected markers allowed differentiation of samples
- ✓ Rice: autogamous and diploid - bulk samples
genotype in each loci is typically homozygote (one allele duplicated)
When two distinct alleles are observed: evidence of residual heterozygosity or
mixed lines – in this case, additional tests are needed
- ✓ Genetic Similarity - conclusions (based on the multiloci genetic profiles)
66 unique genetic profiles – 83,54% of samples
Undistinguishable samples (genetic similarity > 0,95)
Very similar samples
Evaluated case by case – EDVs, parentals, siblings, etc.



Practical uses of the system

- PVP enforcement and Seed Inspection
- Identification of characteristics
- Identify mixture in seed samples
- To check purity of DNA samples
- To Check declared parental information
- Resulting in improvement of the effectiveness and efficiency of the DUS testing process



Example



BRS Carla
Depigmentation due to
high temperature and
hydric stress

BRS Celeste
Depigmentation due to
high temperature and
hydric stress





BRS Celeste - typical (black) and altered color of hilum



BRS Carla - typical (brown) and altered color of hilum

References

Soybean

- Diwan and Cregan, 1997 – Theor. Appl. Genet. 95:723-733
- Cregan et al., 1994 – Meth Molec.& Cell. Biol. 5:49-61
- Song et al., 1999 – Plant Varieties and Seeds 12:207-220
- Idury and Cardon, 1997 – Genome Research 11:1104-1109
- Band coefficient, Lynch, M. 1990 – Molecular Biology and Evolution, 7:478-484

Rice

- Beló et al. (2000) Genetics and Molecular Biology 23:433
- Brondani et al. (2001) Hereditas 134:59-71
- Coburn et al. (2002) Crop Science 42:2092-2099
- Blair et al. (2002) Theor. Appl. Genet. 105:449-457
- McCouch et al. (2002) DNA Research 9:199-207
- Pessoa Filho et al (2007) BMC Plant Biology (online), v. 7, p. 23, 2007





Thank you!
-Luís Gustavo Asp Pacheco

**Examiner - National Plant Variety Protection Office
Ministry of Agriculture, Livestock and Food Supply**

luis.pacheco@agricultura.gov.br

Photo by Roy Kaltschmidt of LBN