Omonrice 15: 54-62 (2007)

GOLDEN RICE CLUSTERING BY MICROSATELLITES FOR FURTHER BREEDING APPROACH

Pham Trung Nghia and Tran Thi Cuc Hoa

Cuu Long Delta Rice Research Institute, Can Tho, Vietnam

ABSTRACT

The major objective of the golden rice project in Vietnam is to develop elite golden rice varieties that are well-adapted to Vietnamese rice-growing conditions. We employ 61 SSR markers to characterize parental lines including five high yielding varieties, which are widely grown in the Mekong Delta, Viet Nam and seven golden rice lines. At Nei and Li's genetic distance of 0.70, tree analysis generated two clusters: (1) Cluster I includes only five indica plant-type varieties: OM1490, OM2490, OM2817, OM3536 and Khau Lech; (2) Cluster II includes seven golden rice lines which originated from Taipei 309, a japonica plant type rice cultivar. Hence, we select a set of 36 SSR markers that covers all 12 chromosomes of rice genome for use in identifying desirable genotypes for backcrossing program in the golden rice project at the Cuu Long Rice Research Institute.

Key words: golden rice, molecular markers, SSR (simple sequence repeat)

INTRODUCTION

The golden rice project aimed to develop a fortified food for use in areas where there is a shortage of dietary vitamin A. The first golden rice prototype was genetically engineered with three beta-carotene biosynthesis genes: *phytoene synthase* (psy), *lycopene cyclase* (lyc) and a bacterial *phytoene desaturase* (crt1) (Ye *et al.* 2000). This transgenic rice with background of Taipei 309 produced up to 1.6 µg of carotenoids per gram of milled grains under greenhouse conditions. A second version of golden rice, harbored only maize *phytoene synthase* and a bacterial *phytoene desaturase,* produces as 23 times more carotenoid level than the first one. To be of practical success in Asian rice-growing countries, golden rice prototypes have to be bred with locally well-adapted, high yielding rice cultivars.

Advances in molecular markers and its implementation in cereal breeding programs have led to a great understanding of cereal genetics and genomes. In recent years, simple sequence repeats (SSR) or microsatellite markers (Tautz 1989) has been widely used for screening, characterization and evaluation of genetic diversity in cereals. This molecular technique is one of the most effective in detecting polymorphism, accessing diversity and assisting selection of desirable plants for backcross through marker-added selection system.

The major objective of the golden rice project in Vietnam is to develop elite golden rice varieties that are well adapted to Vietnamese rice-growing conditions. Hence, we employ SSR markers to (1) characterize parental lines used in the golden rice-breeding program, and (2) to select a set of polymorphic SSR markers for use in backcrossing program.

MATERIALS AND METHODS

Rice parental lines

Twelve parentals including five high yielding varieties which are widely grown in the Mekong Delta, Vietnam and seven golden rice lines used in research at the International Rice Research Institute (IRRI), Philippines (Table 1).

Table 1: List of rice parentals used in golden rice breeding program.

No.	Variety	Origin (*)
1	OM1490 (OM606/IR44592-62)	CLRRI
2	OM2490 (OM723-11/IR50404)	#
3	OM2817 (OMCS97/IR64)	#
4	OM3536 (TD8/OM1738)	#
5	KHẨU LỆCH	#
6	GR 1, line 1	IRRI
7	GR 1, line 2	#
8	GR 1, line 3	#
9	GR 1, line 4	#
10	GR 1, line 5	#
11	GR 1, line 6	#
12	GR 2, line 1	#

(*) CLRRI: Rice genotypes used at Cuu Long Delta Rice Research Institute, Vietnam
IRRI : Rice genotypes used at IRRI, Philippines.

DNA extraction

Rice leaves were collected from 2 weeks-old rice seedlings grown in greenhouse for DNA extraction. Standard molecular grade chemicals and general techniques for preparing stock solutions, buffers, reagents and equipment were followed according to Sambrook *et al.* (1989). Molecular work was conducted at the genome mapping laboratory (GML), IRRI, Philippines.

DNA suitable for PCR analysis was prepared using a simplified procedure according to Mc Couch (1988). A piece of young rice leaf (2cm long) was collected and placed in labeled 2.0 ml microtube in icebox. The leaf was ground in liquid nitrogen using a blue rod and then added 800µl of extraction buffer (50mM Tris pH 8.0, 25mM EDTA pH 8.0, 300 mM NaCl, 1% SDS) by pipetting. The lysate was deproteinized using 400µl of chloroform, mixed well and spinned at the top speed of benchtop centrifuge. The aqueous supernatant was transferred to a new 1.5ml tube and DNA precipitated using an equal volume of absolute ethanol. DNA was then washed with 70% ethanol, air-dried and resuspended in 50µl of TE buffer.

DNA quality was checked using 1% agarose gel. Seven microliliters of DNA sample plus 3µl loading buffer (Tris 1M pH = 8.0, glycerol, EDTA 0.5M pH = 8.0, xylene cyanol 0.2%, bromphenol blue 0.2% and distilled water) was run at 70-80v, 60mA for 45 minutes or until loading buffer dye moved far from the wells. The gel was then stained with ethium bromide and visualized under UV light. DNA concentration was measured by the spectrophotometry method.

SSR analysis

PCR assay: 61 SSR primer pairs (RM Map Pairs, Research Genetics, Hansville, AL) were used to survey polymorphism of parental lines (table 2). The PCR reaction was as follows:

Components for PCR reaction

Reagent	Stock concentration	Adding sequence	Volume (µl)	Final concentration
Sterile dH_2O	---	1	12.5	
PCR buffer with 15 mM $MgCl_2$	10X	2	2.0	1X
dNTP mix	1mM	3	2.0	0.1Mm
Forward Primer	5 µM	4	1.0	0.25 µM
Reverse Primer	5 µM	5	1.0	0.25 µM
Taq polymerase	5 U/ µl	6	0.5	2.5U/ 20 µl
DNA	20ng/ µl	7	1.0	1ng/µl
Total			20	

Table 2: Polymorphic data resulted from 61 SSR primer pairs.

No.	Primer name	Chrom. No.	Allele No.	No.	Primer name	Chrom. No.	Allele No.	No.	Primer Name	Chrom. No.	Allele No.
1	RM5	1	4	21	RM489	3	2	**41**	RM3403	1	5
2	RM17	12	1	22	RM515	8	2	**42**	RM3428	11	4
3	RM38	8	4	23	RM526	2	4	**43**	RM3701	11	3
4	RM168	3	2	24	RM551	4	4	**44**	RM3836	4	4
5	RM174	2	4	25	RM553	9	1	**45**	RM4413	9	4
6	RM207	2	2	26	RM555	-	2	**46**	RM4552	-	3
7	RM248	7	6	27	RM559	4	3	**47**	RM4601	11	1
8	RM256	8	1	28	RM590	10	1	**48**	RM5361	5	2
9	RM258	10	2	29	RM600	1	1	**49**	RM5432	8	3
10	RM276	6	2	30	RM1024	5	3	**50**	RM6051	9	2
12	RM279	2	5	31	RM1150	6	4	**51**	RM6094	11	4
12	RM282	3	3	32	RM1155	4	6	**52**	RM6150	10	4
13	RM401	4	2	33	RM1227	-	5	**53**	RM6369	8	2
14	RM424	2	2	34	RM1230	3	2	**54**	RM6410	12	1
15	RM429	7	6	35	RM3120	8	5	**55**	RM6648	1	4
16	RM436	7	1	36	RM3170	5	8	**56**	RM6574	7	4
17	RM447	8	4	37	RM3215	8	3	**57**	RM6832	3	4
18	RM449	1	3	38	RM3252	1	1	**58**	RM6840	1	4
19	RM455	7	1	39	RM3332	4	2	**59**	RM6959	3	4
20	RM461	6	3	40	RM3346	3	2	**60**	RM7110	7	1
								61	RM8216	12	2

Reactions were overlaid with mineral oil and processed in a Programmable Thermal Controller programmed for 35 cycles of 1 min at 94^0C, 1 min at 55^0C and 2 min at 72^0C, with a final extension at 72^0C for 5 min. After amplification, six microliters of each reaction were run on polyacrylamide gel.

Polyacrylamide gel electrophoresis (PAGE): 8% polyacryramide gel. Glass plates were washed with lab detergent and rinsed with water. The inner surfaces of the plate set were cleaned carefully with 95% ethanol and wiped with lint-free cloth. The gel solution was prepared in a beaker with continuous stirring; then, carefully poured into the glass plate assembly starting from one corner until it reaches the top portion of the short plate. The comb was gently inserted so that its teeth were completely submerged in the gel solution . The gel was run in TAE buffer for 2 to 2:30 hrs at 100 volts. DNA marker of 100bp ladder (50ng/µl) was used for size determination. The acrylamide gel was removed and soaked in ethidium bromide solution for 15 to 20 minutes. Bands in the ethidium bromide stained gels were detected and photographed under UV light. Allelic bands were scored as 1 for the presence of the band or 0 for the absence. Data were entered directly into an excel spreadsheet.

Data analysis

Pair–wise comparisons of the lines based on the presence or absence of unique and shared polymorphic products were used to calculate genetic similarity coefficients by Nei and Li distance measure method (Nei and Li 1979) in the numerical taxonomy and multivariate analysis system (NTSYSpc 2.1) (Rohlf 1990). The unweighted pair group method (UPGMA) was used to cluster parental lines into genetic groups following SHAN grouping method in the NTSYSpc 2.1 software program.

RESULTS AND DISCUSSION

Molecular Diversity Analysis Using SSR (Microsatellite) Markers

The genetic diversity of a population depends upon the number and frequency of alleles (allelic composition). Initial survey using 61 random selected SSR primer pairs was done on the 12 rice

genotypes. These SSR markers were selected based on polymorphic information content and their distribution in the rice genome (Table 2). 60 SSR markers gave amplified polymorphic products among the parental lines except RM 256 locus showed a unique amplified product. Two thirds of the SSR markers demonstrated single locus variation of one, two or three alleles. The highest number of alleles (eight alleles) was scored from locus RM3170 (Table 2). These selected SSR primers generated an average of 3.0 alleles per locus. Some examples of DNA polymorphisms are shown in figures 1 and 2 as exhibited by primers RM17, RM276, RM3428 and RM3836, respectively on the 12 rice varieties resolved by polyacrylamide gel stained with ethidium bromide.

Nei and Li's genetic distances were computed for all the 144 possible combinations of the 12 parental lines. Estimates of the genetic distances among 12 varieties ranged from 0.42 (58% similarity) to 0.70 (30% similarity). The lowest genetic distances were exhibited by (8) GR1 line 3 and (9) GR1 line 4 as these two are genetically closely related and could not distinguished at the cut off 0.42. The dendrogram generated by UPGMA showed the genetic relationships among the 12 varieties studied (Figure 3). As expected, at genetic distance of 0.70, tree analysis generated 2 clusters: (1) Cluster I includes only 5 elite varieties (OM1490, OM2490, OM2817, OM3536 and Khau Lech, Vietnam). These genotypes belong to indica plant type; (2) Cluster II includes only seven golden rice lines which originated from Taipei 309, a japonica plant type. Therefore, SSR markers were able to group the varieties into their plant type groups and these two groups are quite distant from each other as their estimate of genetic similarity is of only 30%. The diversity among elite varieties used in golden rice breding program at Cuu long delta rice research institute were demonstrated by a range in distance from 0.44 (56% similarity) to 0.59 (41% similarity). At the cut off 0.59, the cluster I was futher separated into 2 small groups; one contains 3 elite rice lines (OM1490, OM2490 and OM2817); the other has OM3536 and Khau Lech.

Genotype identification based on SSR markers

For the effective management of parental lines and the selection of desirable plants for backcrossing in golden rice breeding program, we hereby attempted to select useful SSR markers for variety identification. Some genotypes could be differentiated from all other genotypes with a selection of these SSR markers. For instance, five Vietnamese elite rice varieties could be distinguished from each other by a combination of 4 SSR markers: RM17, RM276, RM3428 and RM3836 (Figures 1& 2). As shown in table 3, all 12 parental lines used in the study could be differentiated from each other by a set of 22 SSR markers. We also select a set of 36 SSR markers that covers all 12 chromosomes of rice genome for use in identifying desirable genotypes for backcrossing program (Table 4).

From these results, the following recommendations are presented:

1. Continuous analysis for some backcross generations of golden rice crosses using a set of 36 SSR markers.
2. Extensive molecular marker analysis may be conducted by considering more primers for its relevant application and efficient attainment of breeding objectives in golden rice breeding. More emphasis will be given to added traits come from elite recurent parents

ACKNOWLEDGMENTS

We thank the Bill and Melinda Gates project for funding the golden rice program in Viet Nam and Dr. Virk P. for the support and help in Pham Trung Nghia's on-the-job training at the International Rice Research Institute, Philippines.

Table 3. DNA fingerprinting profile of 12 rice parental lines using 22 SSR primers. Each block represents the amplification profile of the SSR primer indicated on the left side (the first score row stands for the lowest band position)

No	SSR PRIMER	PARENTAL ACCESSIONS												AMPLIFIED ON
		1	2	3	4	5	6	7	8	9	10	11	12	
1	RM5	1	1	1	1	1	0	0	0	0	0	0	1	Chromosome 1
		0	0	0	0	0	1	1	1	1	1	1	0	
		0	0	0	0	0	0	0	0	0	0	0	1	
		0	0	0	0	0	0	0	0	0	0	0	1	
		0	0	0	0	0	0	0	0	0	0	0	1	
2	RM449	0	0	1	1	1	1	1	1	1	1	1	1	Chromosome 1
		1	0	1	0	1	1	0	0	0	0	0	0	
		0	1	0	0	0	0	0	0	0	0	0	0	
		0	0	0	1	1	0	0	0	0	0	0	0	
3	RM3252	1	0	0	0	0	1	1	1	1	1	1	1	Chromosome 1
		0	1	1	1	1	0	0	0	0	0	0	0	
4	RM6648	0	0	0	1	1	1	1	1	1	1	1	1	Chromosome 1
		1	1	1	1	1	0	0	0	0	0	0	0	
		0	0	0	0	0	1	1	1	1	1	1	1	
		0	0	0	1	1	0	0	0	0	0	0	0	
5	RM526	0	0	0	1	1	1	1	1	1	1	1	1	Chromosome 2
		1	1	1	1	1	0	0	0	0	0	0	0	
6	RM489	1	1	1	1	0	0	0	0	0	0	0	0	Chromosome 3
		1	1	1	1	0	0	0	0	0	0	0	0	
		0	0	0	1	1	0	0	0	0	0	0	0	
		0	0	0	0	0	1	1	1	1	1	1	1	
7	RM1320	0	0	0	0	1	0	0	0	0	0	0	0	Chromosome 3
		0	0	0	1	0	1	1	1	1	1	1	1	
		1	1	1	0	0	0	0	0	0	0	0	0	
		0	1	0	0	0	0	0	0	0	0	0	0	
8	RM551	1	1	1	1	1	0	0	0	0	0	0	0	Chromosome 4
		1	1	1	1	1	1	1	1	1	1	1	1	
		0	0	0	0	0	1	1	1	1	1	1	1	
		0	0	0	0	0	0	0	0	0	1	0	1	
		0	0	0	1	1	0	0	0	0	1	1	1	
		0	0	0	1	1	0	0	0	0	0	0	0	
9	RM559	1	1	1	1	1	0	0	0	0	0	1	1	Chromosome 4
		0	0	0	0	1	1	1	1	1	1	1	1	
		0	0	0	1	1	0	0	0	0	0	0	0	
10	RM3836	0	0	0	0	0	1	1	1	1	1	1	0	Chromosome 4
		0	1	0	1	1	0	0	0	0	0	0	1	
		1	0	1	1	1	0	0	0	0	0	0	0	
		0	0	0	1	0	0	0	0	0	0	0	0	
		0	0	0	1	0	0	0	0	0	0	0	0	
11	RM1024	1	1	1	1	0	0	0	0	0	0	0	0	Chromosome 5
		0	0	0	0	1	1	1	1	1	1	1	1	
		0	0	0	0	1	1	1	1	1	1	1	1	
		1	1	1	1	1	0	0	0	0	0	0	0	

No	SSR PRIMER	PARENTAL ACCESSIONS												AMPLIFIED ON
		1	2	3	4	5	6	7	8	9	10	11	12	
		1	1	1	1	1	0	0	0	0	0	0	0	
12	RM276	0	0	0	0	0	1	1	1	1	1	1	1	Chromosome 6
		1	1	0	1	1	0	0	0	0	0	0	0	
		0	0	0	1	1	0	0	0	0	0	0	0	
		0	0	1	0	0	0	0	0	0	0	0	0	
13	RM455	1	0	0	0	1	1	1	1	1	1	1	1	Chromosome 7
		1	0	1	1	0	0	0	0	0	0	0	0	
		0	1	0	0	0	0	0	0	0	0	0	0	
		0	0	0	1	1	0	0	0	0	0	0	0	
14	RM38	0	0	0	0	0	0	0	1	1	1	1	1	Chromosome 8
		1	1	1	1	1	1	1	1	1	1	1	1	
		1	0	0	0	1	0	1	0	0	0	0	0	
		1	0	0	0	1	0	1	1	1	0	0	0	
		1	0	0	0	1	0	1	1	1	0	0	0	
15	OM515	1	1	1	0	0	0	0	0	0	0	0	0	Chromosome 8
		1	1	1	1	1	1	1	1	1	1	1	1	
		0	0	0	1	1	1	1	1	1	1	0	0	
		0	0	0	0	0	0	0	0	0	0	1	1	
16	OM5432	0	0	0	0	0	0	0	0	0	0	0	1	Chromosome 8
		0	0	0	0	0	0	0	1	1	0	1	1	
		0	0	0	0	0	0	0	0	1	0	0	0	
		1	1	1	1	1	1	0	0	1	1	0	0	
		0	0	0	0	0	0	1	1	0	0	1	0	
		0	0	0	0	0	0	0	0	0	0	0	1	
17	OM553	1	1	0	0	0	0	0	0	0	0	0	0	Chromosome 9
		0	0	1	1	1	1	1	1	1	1	1	0	
		0	0	0	0	0	0	0	0	0	0	0	1	
18	OM258	1	0	1	1	1	0	1	0	0	0	0	0	Chromosome 10
		0	0	0	0	0	0	0	0	1	0	1	0	
		0	0	0	0	0	0	0	1	0	1	0	1	
		0	0	0	0	0	0	0	1	1	0	0	0	
19	OM3701	0	0	0	0	0	0	0	0	1	1	1	0	Chromosome 11
		0	0	0	0	0	0	0	0	0	0	0	1	
		1	1	1	1	1	1	1	1	0	0	0	0	
		0	0	0	0	0	0	0	0	0	0	0	1	
		1	1	1	1	1	0	1	1	1	0	1	0	
		1	1	1	1	1	0	1	1	1	0	1	1	
20	OM6094	1	1	1	1	1	1	1	1	1	1	1	1	Chromosome 11
		1	1	1	1	1	1	1	1	0	0	0	0	
		1	0	0	1	1	1	1	1	0	0	0	0	
		0	0	0	0	0	0	1	1	0	0	0	0	
21	OM17	1	1	0	1	0	1	1	1	1	1	1	1	Chromosome 12
		0	0	1	0	1	0	0	0	0	0	0	0	
22	OM8216	1	1	1	1	0	0	0	0	0	0	0	0	Chromosome 12
		0	0	0	0	0	1	1	1	1	1	1	1	
		0	0	0	0	1	1	1	1	1	1	1	1	
		0	0	0	0	1	0	0	0	0	0	0	0	

Table 4. Selected SSR primers for use in backcross breeding in the Golden Rice Project at CLRRI, Vietnam. All primers have annealing temperature at 55°C.

No.	SSR primers	Chrom. No.	Repeat motif	No.	SSR primers	Chrom. No.	Repeat motif
1	RM449	1	(AG)12	19	RM455	7	(TTCT)5
2	RM490	1	(CT)13	20	RM429	7	(TG)10
3	RM14	1	(GA)18	21	RM436	7	(TAA)6
4	RM526	2	(TAAT)5	22	RM5432	8	(TC)16
5	RM29	2	(GA)7	23	RM72	8	(TAT)5C(ATT)15
6	RM535	2	(AG)11	24	RM502	8	(TG)10
7	RM489	3	(TAT)8	25	RM553	9	(CT)10
8	RM114	3	(GA)7	26	RM316	9	(GT)8-TG)9(TTTG)4(TG)4
9	RM1230	3	(AG)15	27	RM2144	9	(AT)22
10	RM551	4	(AG)18	28	RM258	10	(GA)21(GGA)3
11	RM559	4	(AACA)6	29	RM590	10	(TCT)10
12	RM3836	4	(GA)22	30	RM222	10	(CT)18
13	RM1024	5	(AC)13	31	RM473	11	(TCTA)14
14	RM421	5	(AGAT)6	32	RM3428	11	(CT)18
15	RM289	5	G11(GA)16	33	RM4B	11	(AG)16
16	RM276	6	(AG)8A3(GA)33	34	RM17	12	(GA)21
17	RM8101	6	(TA)28(GTTA)9	35	RM8216	12	(TAA)25
18	RM340	6	(CTT)8T3(CTT)14	36	RM19	12	(ATC)10

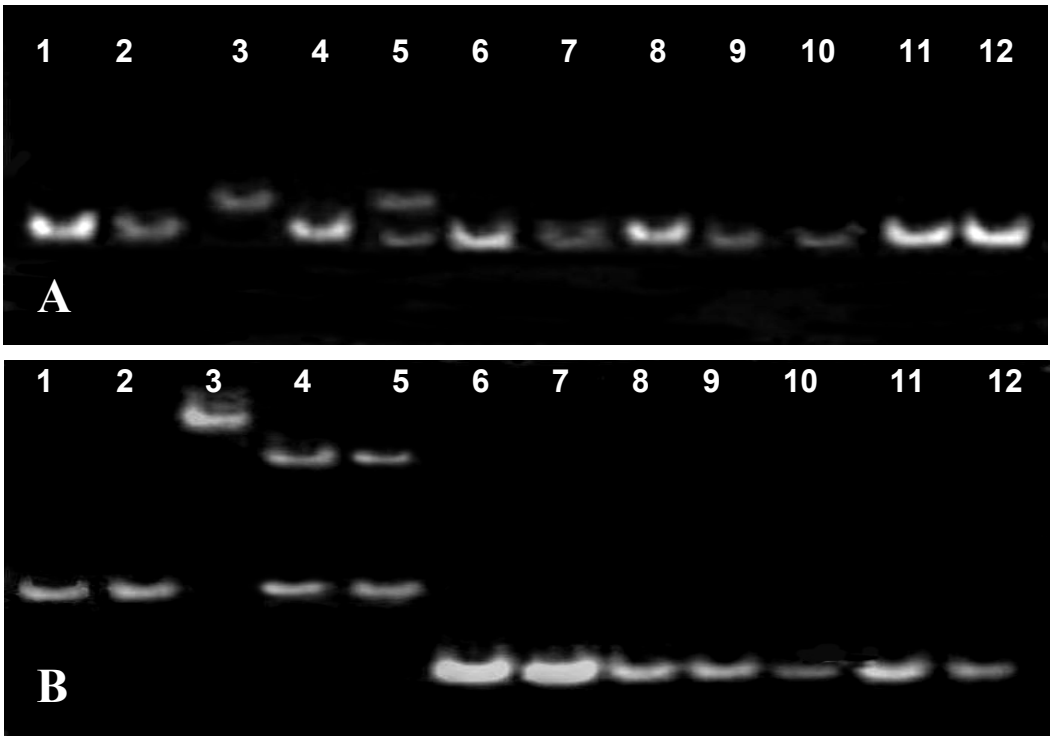

Figure 1. Microsatellite DNA profiles of parental rice varieties used in the golden rice-breeding program detected by (A) SSR primer pair RM17 and (B) SSR primer pair RM 276.

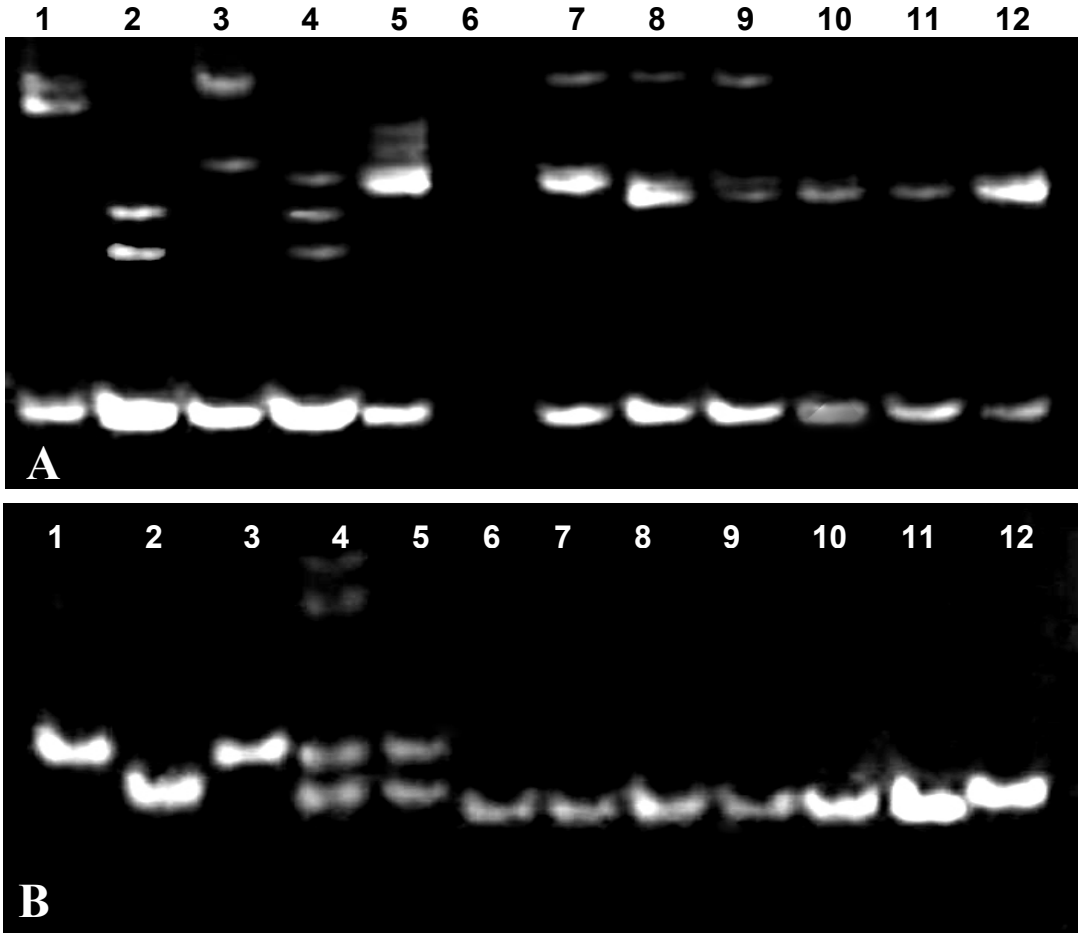

Figure 2. Microsatellite DNA profiles of parental rice varieties used in the golden rice breeding program detected by (A) SSR primer pair RM3428 and (B) SSR primer pair RM3836

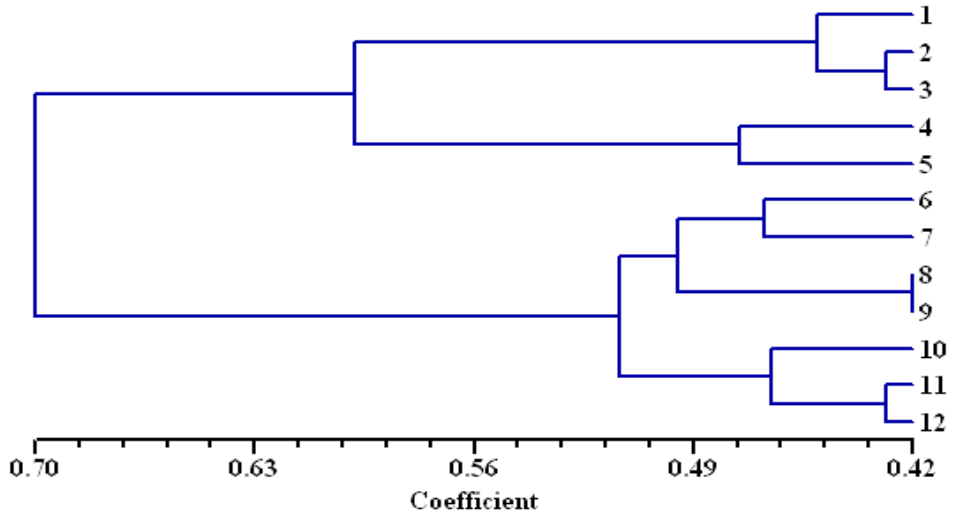

Figure 3: Dendrogram 12 parental rice lines for use in the Golden Rice Breeding program based on genetic distance coefficients generated from 61 SSR primers (184 alleles).

REFERENCES

McCouch S.R. 1988. Molecular mapping of rice chromosomes. Theor. Appl. Genet. 76: 815-829.

Nei M and W Li. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Natl. Acad. Sci. USA 76: 5269–5273.

Rohlf FJ. 1990. NTSYS-pc. Numerical Taxonomy and Multivariate Analysis System. New York: Applied Biostatistics Inc. 175p.

Sambrook J, EF Fritsch and T Maniatis. 1989. Molecular cloning: A laboratory manual, 2nd edition. Cold Spring Harbor, NewYork.

Tautz D. 1989. Hypervariability of simple sequences as a general source for polymorphic DNA markers. Nucleic Acids Res.17: 6463–6471.

Ye X, S Al-Babili, A Kloti, J Zhang, P Lucca, P Beyer, and I Potrykus. 2000. Engineering the provitamin A (beta-carotene) biosynthetic pathway and (carotenoid free) rice endosperm. Science 287: 303-305.

PHÂN NHÓM CÁC DÒNG LÚA VÀNG BẰNG MICROSATELLITE MARKER

Mục đích của dự án "hạt lúa vàng" thực hiện tại Viện Lúa Đồng bằng sông Cửu Long là nhằm tạo ra các giống lúa có hạt gạo giàu vitamin A, thích hợp với điều kiện canh tác ở các vùng trồng lúa Việt Nam. Phương pháp đánh dấu phân tử SSR marker được ứng dụng nhằm đánh giá các giống lúa dùng làm bố mẹ trong các chương trình lai tạo của dự án hạt lúa giàu vitamin A. Dựa trên phương pháp tính toán khoảng cách di truyền của Nei và Li, năm giống lúa cao sản của Việt Nam nằm trong cùng một nhóm, khác biệt với nhóm lúa chuyển gen giàu vitamin A có nguồn gốc từ giống Taipei 309 ở khoảng cách di truyền là 0,7. Từ kết quả thí nghiệm này, 36 SSR marker được tiếp tục sử dụng nhằm chọn lựa con lai thích hợp cho các chương trình lai tạo hồi giao.