USING MOLECULAR MARKERS IN STUDY OF RICE GENETIC DIVERSITY

BUI CHI BUU AND NGUYEN THI LANG

ABSTRACT

Random amplified polymorphism DNA (RAPD) was used as a DNA fingerprinting technique in rice germplasm evaluation. The high efficiency and random coverage of RADP markers were established to analyse the biodiversity of 72 rice germplasm accessions . We examined the amplification products in both their size and their polymorphism. Correlation matrix was carried out using the Genstat program. Cluster analysis using the average-linkage (UPGMA) method was performed using Genstat program, with the similarity matrix as input data, based on Nei's genetic distance. Twenty primers from OPA kit were screened on the total DNA obtained from the leaf tissues of rice . Only ten markers OPAA11, OPAJ01, OPAA13, OPAB17, OPAC14, OPAG08, OPB06, OPAL09, OPAL08, OPAK12 yielded the amplified products. Accurate classification of rice germplasm into the two major clusters and many subclusters can provide essential information for selecting parents in the development of intercluster crossing program. The number of accessions distinguishable individually with the selected primers varied from 19 with OPAJ01, 20 with OPAL08, to 37 with OPAA11. Upland rice landraces such as Jo anh, Koi ame were classed in the same subcluster. The rainfed lowland rice landraces in coastal centre such as Lua con, Lua se, Ven Nghe An have the same subcluster of Oryza officinalis. Glutinous rices such as Nep Som, Nep Oc have the same subcluster of deep water rice in the Mekong Delta like Nam Vang, Lua Lem lun. Glutinous floating rice Nep Co Ba was classified in the same subcluster of normal glutinous rice: Nep Cai Hai Duong. RAPD markers could be very useful for evaluating germplasm because they are easier to detect than RFLPs, but one must proceed cautiously in interpreting RAPD data.

Key words: rice, diversity, RAPD, marker

INTRODUCTION

Plant genetic resource management comprises two phases: (i) germplasm conservation including acquisition of germplasm *in-situ* or *ex-situ*, preserving under controlled conditions, monitoring its viability, maintaining passport and other data, characterization heritable morphological and molecular traits of germplasm; (ii) germplasm management including evaluation, utilization, genetic enhancement (Duvick 1990, Bretting and Widrlechner 1995) as making particular genes more accessible and usable to breeders. The role of genetic markers in genetic enhancement is considered in the context of germplasm management as a whole (Chang 1985, Duvick 1990). The contributions of genetic markers to gene mapping and to plant breeding have thoroughly reviewed in fingerprinting commercial germplasm (Smith and Smith 1992)

Recent advances in molecular biology, principally in the development of the polymerase chain reaction (PCR) for amplifying DNA, DNA sequencing and data analysis, have resulted in powerful techniques which can be used for the screening characterization and

evaluation of genetic diversity. Traits that serve as genetic markers are by definition polymorphic; the more polymorphic the trait, the greater its potential value to germplasm management. The issue of homology may seem trivial for morphological markers, but the increasing use of molecular markers has heightened its importance (Bretting and Widrlechner 1995)

The study of morpho-agronomic variability is the classical way of assessing genetic diversity for plant breeders. For many species, especially rice, it is still the only approach used by breeders. However, with molecular marker techniques, powerful tools have been developed so that genetic resources can be accurately assessed and characterized.

Genetic marker screening is based on the survey of genetic diversity as revealed by variation at specific gene loci and provides information about the amount and distribution of genetic diversity within and among populations. The emphasis of this report will be on DNA-based molecular techniques and how they can applied in assessing the genetic diversity of genetic resources.

Genetic markers should be: easy scored, negligible effects on plant growth, rapidly, safely, and inexpensive scored (Muray et al. 1988, Smith 1989, Chunwongse et al. 1993). Polymorphic DNA is thought to provide ideal genetic markers because (i) nucleotide sequence variation is selectively neutral (Kimura 1983, Nei 1987); (ii) certain complication reducing heritability of protein may be minimized; (iii) three distinct genomes as nuclear, chloroplast, mitochondria, may each involve according to different modes and tempos (Bretting and Widrlechner 1995).

The importance role of genetic diversity assessment in plant genetic resource management was highlighted (Kresovich and McFerson 1992). When genetic marker data can be interpreted by locus/allele models, allelic diversity can be described by expected heterozygosity $H = 1 - \Sigma_i\Sigma_{j\to m} p^2_{ij} / m$, where p_{ij} is frequency of the i^{th} allele at the j^{th} of m loci (Nei 1973, 1987; Brown and Weir 1983).

MATERIALS AND METHODS

DNA of 72 local rice varieties from Vietnam were extracted (table 1)

DNA amplification and RAPD

The optimal reaction for RAPD analysis was set up under the following conditions: 1X reaction buffer, 0.5 µl Taq DNA polymerase. 0.3 µM of the 10 mer random primer, 150µm dNTPs and 25-50ng template DNA for total volume of 25 µl. Amplification conditions were set up using a programmable thermalcycler. The arbitrary primer kits OPA were purchased from Operon Technologies. A total of 20 primers were screened in this study.

The amplification products were separated on 1.5 % agarose gels in 0.5X TBE buffer. The banding patterns were visualized under UV light and photograped using a polaroid camera. One kilobase ladder was used as DNA standard.

Primer screening

20 primers from OPA kit were screened on the total DNA obtained from the leaf tissues of rice to yield amplification products.

Data analysis

Each informative RAPD band was scored independently as 1 for "presence" and 0 for "absence"

Correlation matrix was carried out using the Genstat programme. Similarity matrix was generated based on the simple-matching coefficient, using the presence / absence data for individual RAPD fragment between pairs of rice accessions. Cluster analysis using the average-linkage (UPGMA) method was performed using Genstat programme, with the similarity matrix as input data.

Nei's distance

$$D = \ln\left(\frac{\sum_m \sum_i P_{1mi} P_{2mi}}{\left[\sum_m \sum_i P_{1mi}^2\right]^{1/2} \left[\sum_m \sum_i P_{2mi}^2\right]^{1/2}}\right)$$

where m = summed over loci

i = over alleles at the m^{th} locus

P_{1mi} = frequency of the i^{th} allele at the m^{th} locus in population 1

RESULTS AND DISCUSSION

PCR amplification of total genomic DNA using 20 random 10-mer primers yielded scorable amplification products (Table 2). The amplification produced obtained with each of these primers was resolved on 1.5% agarose gels. The size of amplification produced scored in 1.5 % agarose gels ranged between 100-2000bp. The number of accession distinguishable individually with selected primers varied. Collectively, these ten primers were sufficient to distinguishable all the cultivars and accessions analyzed in the study. These are probably sufficient to identify the 72 distinct cultivars.

The RAPD analyses generally detect the occurrence of a single allele, whereas isozyme, RFLP, and other DNA techniques can distinguish among many alleles at specific loci (William et al. 1993). The DNA fragments produced via arbitrary priming are generally inherited a simple dominant-Mendelian fashion, with fragment absence recessive. In this respect, RAPD markers may be inferior to codominant genetic markers, although the frequency of alleles coding for fragment occurrence or absence may be estimated by maxium-likelihood procedures (Edward 1992), and nucleotide divergence can be estimated from RADP data via relevant statistical analyses (Clark and Lanigan 1993).

In addition to RAPDs, PCR technique can amplify specific genetic loci containing variable numbers of tandemly repeated nucleotide sequences (Nakamura et al. 1987) of about 10 to 50 base pairs [minisatellites] (Jeffrey et al. 1985).

Table 1: List of 72 local varieties used for clustering analysis

No	Acc. no.	Designation	No	Acc. no.	Designation
1	32078	Thom	37	32179	Nang bet
2	32079	Trang cut D11	38	32191	Nang tay C
3	32082	Bang cha	39	32194	Nang thau
4	32084	Ba sao	40	32195	Nang thuot
5	32085	Ba se chum	41	32200	Nanh chon
6	32093	Bay danh	42	32209	Nep ca ro
7	32097	Bup tra bong	43	32212	Nep co ba
8	32098	Cai don	44	32218	Nep do
9	32099	Ca nhan	45	32223	Nep lem
10	32101	Canh nong sa bo	46	32226	Nep mo
11	32103	Dalat lua ray	47	32228	Nep mui
12	32111	Gay xe	48	32229	Nep muong
13	32114	Jo anh	49	32231	Nep non tre
14	32116	Koi ame	50	32233	Nep quan
15	32117	Koi con	51	32235	Nep ruoi huong
16	32118	Koi goum	52	32237	Nep sap
17	32119	Koi ke	53	32239	Nep som
18	32120	Koikon	54	32243	Nep thap
19	32123	Lua ba ba	55	32248	Nep trang
20	32124	Lua ba bong	56	32250	Nha trang
21	32126	Lua can	57	32274	Trang quang bay
22	32127	Lua can	58	47461	Bau
23	32129	Lua con	59	47463	Bau huong Hai Duong
24	32121	Lua doi	60	47483	Chiem loc Nghe An
25	32133	Lua don	61	47506	Chiem 3
26	32135	Lua lem lun	62	47531	Nep cai Hai Duong
27	32138	Lua lu trang	63	47532	Nep oc
28	32139	Lua mau nau	64	47533	Nep sap
29	32141	Lua rang do	65	47539	Re quang Ha Tinh
30	32143	Lua re do	66	47544	Re thom Ha Dong
31	32145	Lua se	67	47550	Sai duong
32	32152	Mbrabrung	68	-	Ven Nghe An
33	32155	Mdie ke no	69	-	*Oryza officinalis*
34	32160	Nam vang	70	-	IR54 (check)
35	32166	Nang co	71	-	IR64 (check)
36	32177	Nanh moi	72	-	Azona (check)

Nep : glutinous rice

Table 2. Ten primers and the characteristics of their amplification products.

Primer	Sequence	No .of band detected	Size of product (kb)
OPAA11	CAATCGCCGT	5	1.4
OPAJ01	ACGGGTCAGA	8	2.0
OPAA13	GAGCGTCGCT	5	0.1
OPAB17	CCTGTACCGA	8	1.6
OPAC14	GTCGGTTGTC	5	1.5
OPAG08	AAGAGCCCTC	4	1.3
OPB06	TGCTCTGCCC	6	1.8
OPAL09	CAGCGAGTAG	6	1.6
OPAL08	GTCGCCCTCA	7	1.2
OPAK12	AGTGTAGCCC	5	2.0
Total		59	

Primer screening

To identify primers that detect polymorphism, 20 primers from OPA kit were screened on the total DNA obtained from the leaf tissues of rice . Of these 20 primers, ten failed to yield amplification products. The remaining ten markers *OPAA*11, *OPAJ*01, *OPAA*13, *OPAB*17, *OPAC*14, *OPAG*08, *OPB*06, *OPAL*09, *OPAL*08, *OPAK*12 (table 2). The size of the yielded reproducible fragments and at least 59 loci were scorable. The size of the fragments ranged from 100 to 2000 bp. Fig 2, 3 and 4 show the fragments in rice amplified DNA obtained with *OPAJ*01, *OPAA*11 and *OPAL*08 primers, respectively.

Cluster analysis

DNA markers can provide information on genetic diversity of the germplasm. The random amplified polymorphism DNAs (RAPDs) were used to survey DNA sequence variation of 72 local varieties. Accurate classification of rice germplasm into the two major clusters and many subclusters can provide essential information for selecting parents in the development of intercluster crossing program.

PCR amplification of total genomic DNA using ten random 10-mer primers yielded scorable amplification products.

Based on computing genetic distance from gene frequencies, we can read the distance matrix phylogeny programs FITCH and KITSH. Then bootstrap is used in phylogeny estimation.

The number of accessions distinguishable individually with the selected primers varied from 19 with *OPAJ*01, 20 with *OPAL*08, to 37 with *OPAA*11.

The reliability of RAPD data for the classification of rice germplasm was tested by subjecting the data to unweighted pair group method analysis of arithmetic means (UPGMA) in order to explore the possibility of classifying the cultivars using RAPD analysis.

A usefulness of this technique for germplasm characterization depends on ability to sample any portion of the genome, study markers on all the linkage groups, detect genetic differences among distinct genotypes, classify the accessions into specific groups (Bhat et al.1995)

The accession included (Table 1) were chosen for their distinctiveness as well as for the close similarities making them difficult to distinguish using morphological markers. The two varieties varieties Lua Thom (acc.1) from Mekong Delta and Lua Re Do (acc.30) from coastal central areas have the same cluster (Figure 1). The remaining cluster can be classified into many subclusters. Upland rice landraces such as Jo anh, Koi ame were classified in the same subcluster. The rainfed lowland rice landraces in coastal centre such as Lua con, Lua se, Ven Nghe An have the same subcluster of *Oryza officinalis*. Glutinous rices such as Nep Som, Nep Oc have the same subcluster of deep water rice in the Mekong Delta like Nam Vang, Lua Lem lun. Glutinous floating rice Nep Co Ba was classified in the same subcluster of normal glutinous rice: Nep Cai Hai Duong.

These cultivars were placed in different subclusters along with those previously classified thus helping in the identification of their genomic composition.

The random amplified polymorphic DNAs (RAPDs) are very simple to detect because they do not require DNA sequence information or synthesis of specific primers. However, because the fragments are amplified based on homology to a very short, random DNA sequence used to prime the PCR, there is some uncertainty about the genetic relationship of fragments from different genotypes and about the genome origin of the fragments. RAPD markers could be very useful for evaluating germplasm because they are easier to detect than RFLPs, but one must proceed carefully in interpreting RAPD data.

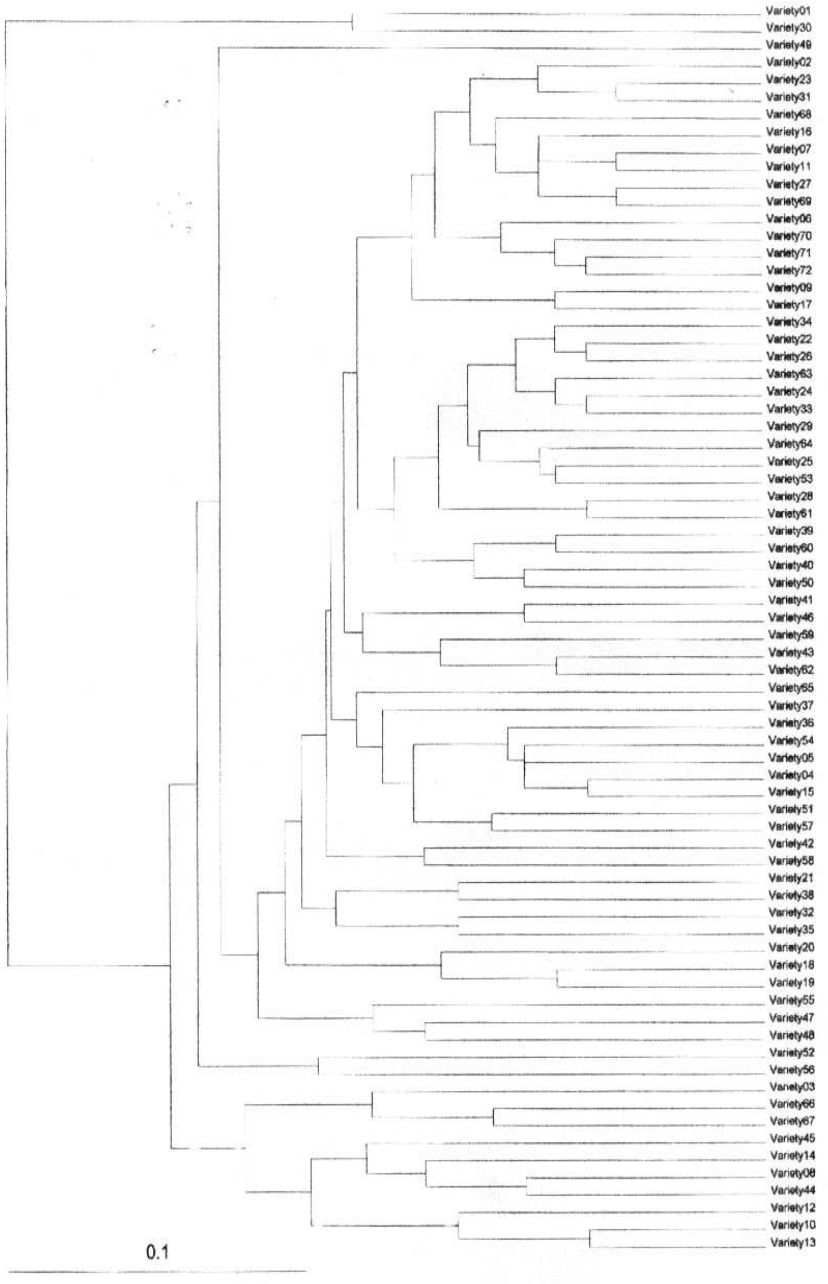

Fig. 1 Phenogram resulting from the analysis of 59 RAPD alleles depicting relationship between 72 local rice accessions. The key to abbreviations is in table 1

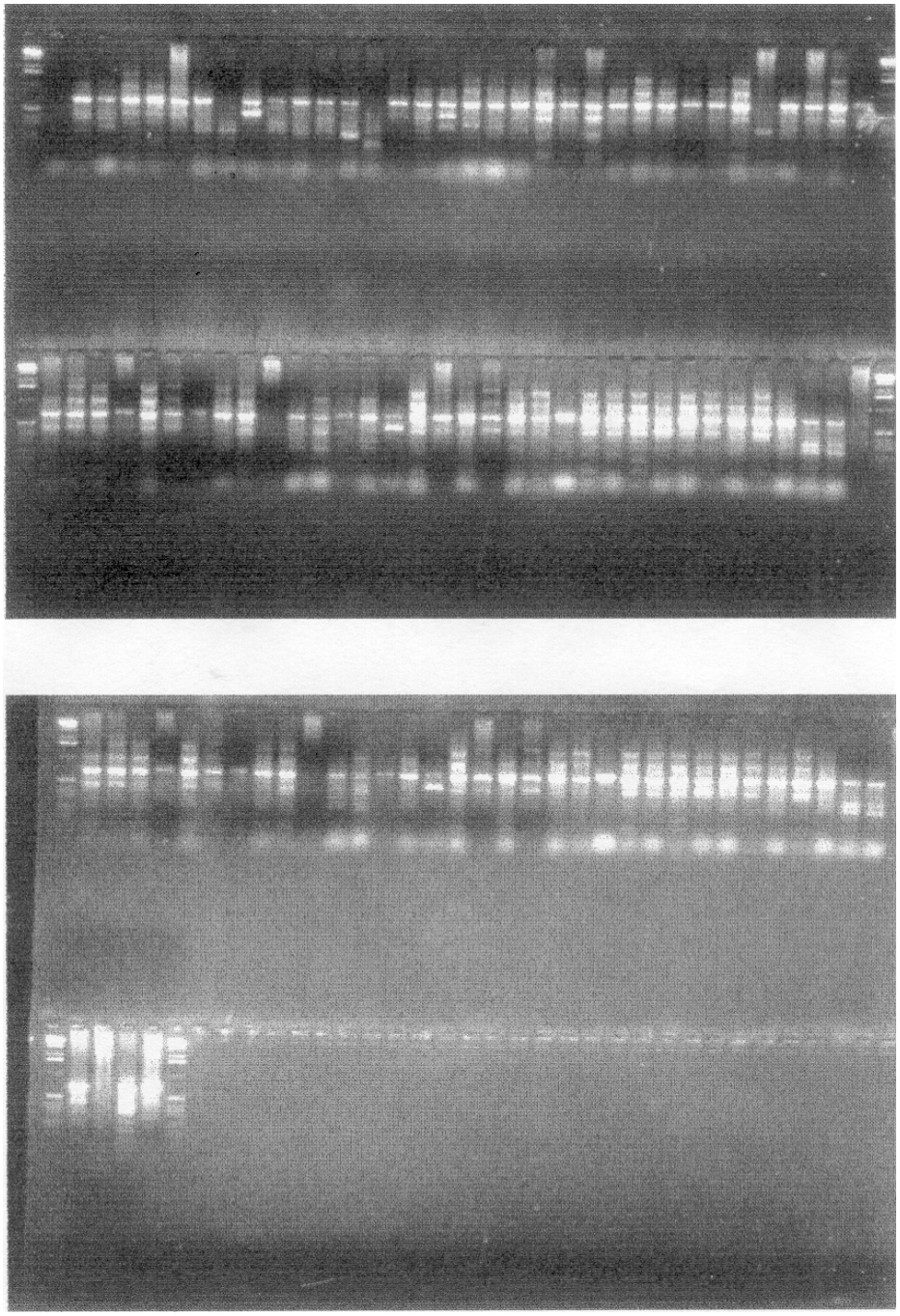

Fig. 2: OPAJ01

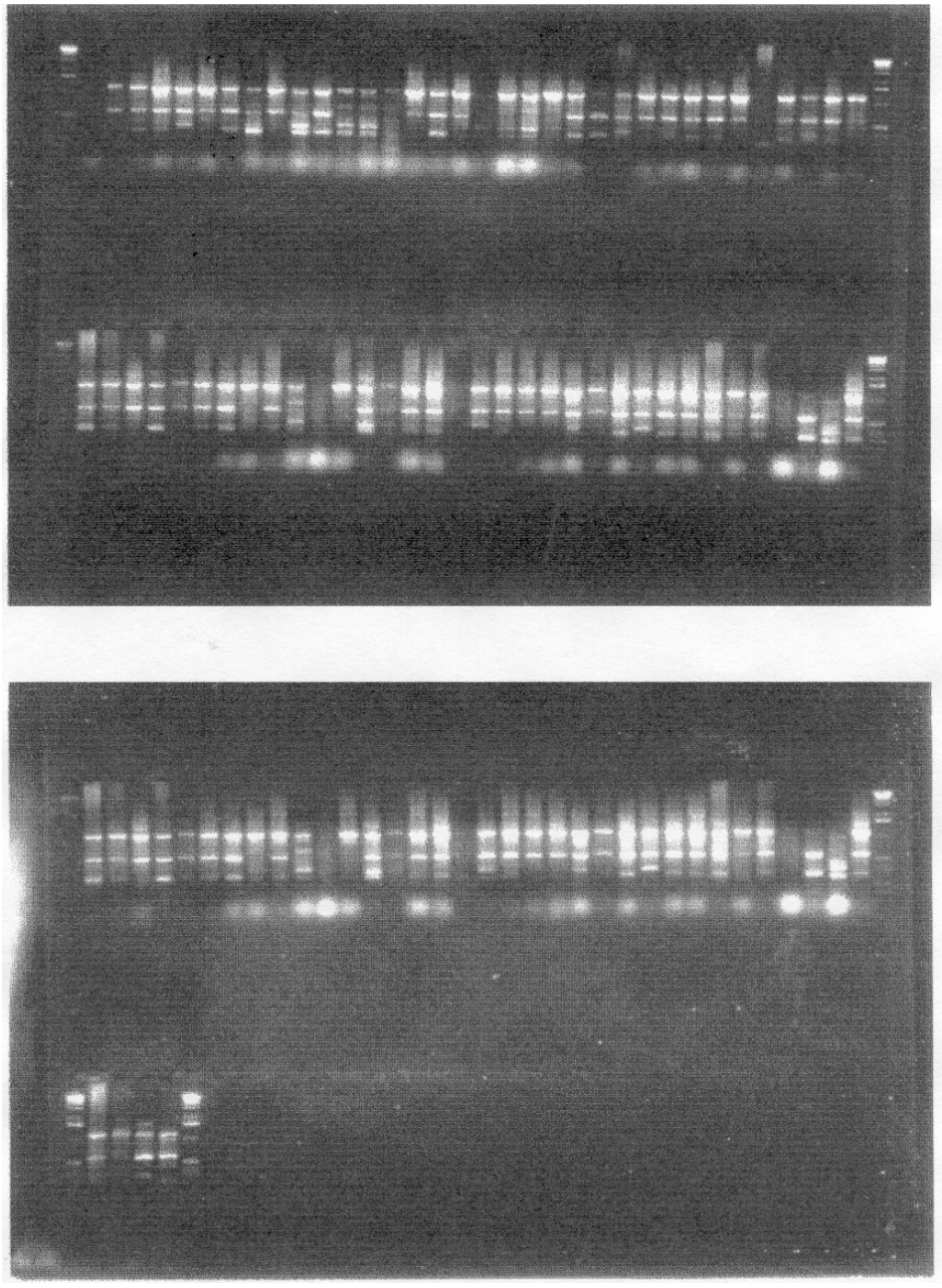

Fig. 3: OPAA11

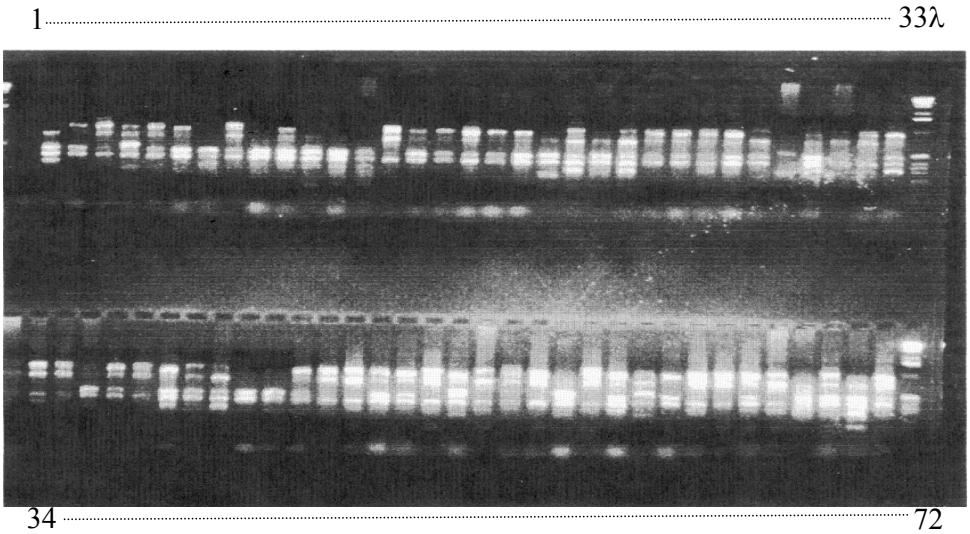

Fig. 4: OPAL08

References

Bhat KV, S Lakhanpaul, KPS Chandel, RL Jarret. 1995. Molecular markers for characterization and identification of gene resources of perennial crops. P.105-116 *in* WG Ayad, T Hodhkin, A Jaradat, VR Rao (eds.). Molecular genetic techniques for plant genetic resources. Report of an IPGRI Workshop, 9-11 October 1995, Rome, Italy

Bretting PK, MP Widrlechner. 1995. Genetic markers and plant genetic resource management. P. 11- 86 *in* Planr Breeding Reviews, Volume 13, Edited by J Janick. John Wiley & Son Inc. Canada.

Brown ADH, BS Weir. 1983. Measuring genetic variability in plant populations p. 219-240 *in* SD Tanksley and TJ Orton (eds.). Isozymes in plant genetics and breeding 1. Vol. 1A. Development in plant genetics and breeding 1. Elsevier, Amsterdam

Chang TT. 1985. Germplasm enhancement and utilization. Iowa State J. Res. 59:399-424

Chunwongse JG, B Martin, SD Tanksley. 1993. Pregermination genotypic screening using PCR amplification of half-seeds. Theor. Appl. Genet. 86: 694-698

Clark AG, CMS Lanigan. 1993. Prospects for estimating nucleotide divergence with RAPDs. Mol. Biol. Evol. 10:1096-1111

Duvick DN. 1990. Genetic enhancement and plant breeding. P. 90-96 *in* J Janick and JE Simon (eds.). Advances in New Crops. Pros. First National Symposium on New

Crops: Research, Development, Economics. Timber Press, Portland, OR

Edward AWF. 1992. Likelihood (expanded ed.) John Hopskins Univ. Press, Baltimore.

Jeffrey AJ, V Wilson, SL Thein. 1985. Hypervariable "minisatellite" regions in human DNA. Nature 314:67-73

Kimura M. 1983. The neutral theory of molecular evolution. Cambridge Univ. Press, Cambridge

Kresovich S., JR McFerson. 1992. Assessment and management of plant genetic diversity: consideration of intra- and interspecific variation. Field Crop Res. 29:185-204

Murray MG, J Ma, J Romero-Severson, DP West, JH Cramer. 1988. Restriction fragment length polymorphisms: what are they and how can breeders use them? Proc. Annu. Corn Sorghum Res. Conf. 43:72-87

Nakamura Y, M Leppert, P O'Connell, R Wolff, T Holm, M Culver, C Martin, E Fujimoto, M Hoff, E Kumlin, R White. 1987. Variable number of tandem repeat (VNTR) markers for human gene mapping. Science 235:1616-1622

Nei M. 1973. Analysis of gene diversity in subdivided population. Proc. Nat. Acad. Sci. USA 70: 3321-3323

Nei M. 1987. Molecular evolutionary genetics. Columbia Univ. Press, New York

Smith JSC, OS Smith. 1992. Fingerprinting crop varieties. Adv. Agron. 47: 85-140

Smith JSC. 1989. Gene markers and their uses in the conservation, evaluation, and utilization of genetic resources of maize (*Zea maydis* L.). P. 125-138 *in* HT Stalker and C Chapman (eds.). Scientific management of germplasm: characterization, evaluation, and enhancement. IBPGR training courses: lecture series, 2. Dept. of Crop Science, North Carolina State, Univ. Raleigh, NC and IBPGR, Rome.

William CE, DA St. Clair. 1993. Phenetic relationships and levels of variability detected by restriction fragment length polymorphism and random amplified polymorphic DNA analysis of cultivated and wild accession of *Lycopersicum esculentum*. Genome 36: 619-630

TÓM TẮT

Sử dụng marker phân tử trong nghiên cứu tính đa dạng di truyền cây lúa

Chúng tôi sử dụng marker phân tử RAPD như một công cụ trong kỹ thuật "DNA fingerpringting" để đánh giá qũy gen cây lúa. Sự đa dạng của 72 mẫu giống lúa địa phương ở Việt Nam đã được phân tích bằng RAPD. Với sự trợ giúp của PCR, các mẫu DNA khuếch đại đã được xem xét về độ lớn và độ đa hình. Ma trận tương quan được thiết lập theo chương trình Genstat. Phân tích nhóm di truyền được thực hiện theo phương pháp UPGMA trên Genstat, dựa trên khoảng cách di truyền của Nei đề xuất. Trong 20 primer thử nghiệm thuộc OPA kit, có 10 primer cho kết qủa khuếch đại các băng tốt nhất, với 59 loci, đó là *OPAA*11, *OPAJ*01, *OPAA*13, *OPAB*17, *OPAC*14, *OPAG*08, *OPB*06, *OPAL*09, *OPAL*08, *OPAK*12. Kết qủa có hai cluster chính trong 72 mẫu giống phân tích, bao gồm các giống lúa mùa, lúa chiêm ở đồng bằng sông Hồng, lúa nước sâu của đồng bằng sông Cửu Long (ĐBSCL), giống lúa nếp của Tây Nguyên, giống lúa nước trời của Duyên hải Trung Bộ. Kết qủa cho thấy có nhiều subcluster được phân lập. Thông tin này có lợi cho nhà chọn giống trong sử dụng vật liệu lai giữa những subcluster trong chương trình lai tạo. Những marker có khả năng giúp chúng ta phân biệt từng mẫu giống là *OPAJ* 01 (19 mẫu), *OPAL* 08 (20 mẫu), và *OPAA*11 (37 mẫu). Giống lúa rẩy Jo ahn, Koi me được xếp chung một subcluster. Nhóm giống lúa nước trời ở Duyên hải Trung bộ: Lúa Côn, Lúa Se, Ven Nghệ An, có cùng subcluster với lúa hoang *O. officinalis*. Nhóm lúa nếp như Nếp Sớm, Nếp Ốc được xếp chung với nhóm lúa nước sâu ở ĐBSCL (Nam Vang, Lúa Lem Lùn). Lúa nổi Nếp Cô Ba có cùng nhóm với Nếp cái Hải Dương. Marker phân tử RAPD có thể sử dụng trong phân tích tính đa dạng di truyền tập đoàn giống lúa địa phương, nhưng nó những nhược điểm riêng, cần phải thận trọng trong diễn giải kết qủa số liệu của RAPD.