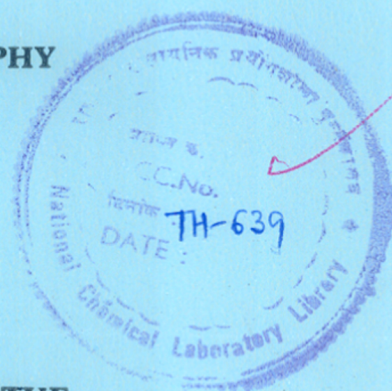


A MOLECULAR APPROACH FOR THE ASSESSMENT OF VARIABILITY IN INDICA RICES

COMPUTERISED

A THESIS
SUBMITTED TO THE
UNIVERSITY OF POONA
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
(IN CHEMISTRY)



BY
VAIJAYANTI VASANT PETHE

M. Phil

577.21:633.18 (043)
PET

DIVISION OF BIOCHEMICAL SCIENCES
NATIONAL CHEMICAL LABORATORY
PUNE-411 008 (INDIA)

1991

C O N T E N T S

	PAGE
ACKNOWLEDGEMENTS	
DECLARATION	
LIST OF ABBREVIATIONS	
LIST OF FIGURES	
LIST OF TABLES	
SYNOPSIS	1-9
CHAPTER 1 : NOVEL FEATURES OF REPEATED DNA SEQUENCES IN RICE GENOME	
1.1 Rice classification	10
1.2 Content and gross arrangement of repeated DNA sequences in rice	13
1.3 Genome specific and species specific repeat DNA sequences	15
1.4 Transcribed repeated sequences	21
1.5 Transposable elements	22
1.6 Repeats showing copy number variation & differential methylation	24
1.7 Ribosomal genes in rice	25
References	30
CHAPTER 2 : MATERIALS & METHODS	
2.1 Materials and Chemicals	
2.1.1 Seed material, growth conditions and harvest of tissue	40
2.1.2 Chemicals, materials & enzymes	43
2.2 Isolation of DNA & criteria of purity	

2.2.1	Isolation of genomic DNA from leaf tissue of rice variety Malkolam and 19 elite varieties of indica rice; shoots, (8-10 days old seedlings) and embryos of variety Malkolam	44
2.2.2	Isolation of plasmid DNA	46
2.3	Restriction endonuclease digestion	48
2.4	Agarose gel electrophoresis	49
2.5	Pre-requirements for cloning	
2.5.1	Vector & host system used for the construction of library	49
2.5.2	Media preparation	51
2.5.3	Growth, culturing & storing of <u>E.coli</u> DH5α	52
2.6	Construction of a partial genomic library of rice	
2.6.1	Preparation of target & vector DNA	53
2.6.2	Ligation of vector & insert DNA	54
2.6.3	Preparation of competent cells & transformation	54
2.6.4	Selection of recombinants & saving the library	57
2.7	Screening the library by colony hybridization	58
2.8	Southern blotting	60
2.9	Labelling probe DNA and its purification	61

2.10	Prehybridization, hybridization, post-hybridization washing of filters and autoradiography	64
2.11	Copy number estimation	66
2.12	RFLP analysis	66
	References	70

CHAPTER 3 : RESULTS

SECTION I

3.1	Construction of a partial genomic library of rice	
3.1.1	Cloning of rice DNA for generation of a partial genomic library and it's screening for repeat clones	73
3.1.2	Determination of insert sizes in a few random genomic clones	75
3.1.3	Identification of a moderately repeated DNA sequence pOSMLC-2 from the library	76

SECTION II

3.2	Characterization of a dispersed repeat sequence pOSMLC-2 in rice genome	
3.2.1	Nature of pOSMLC-2 sequence	81
3.2.2	Organization of pOSMLC-2 sequence	84
3.2.3	Abundance or copy number estimation of pOSMLC-2 sequence	85

3.2.4	Restriction mapping of pOSMlc-2 sequence	88
3.2.5	Homology studies within the pOSMlc-2 repeat sequence	93
3.2.6	Organization of the two sub-repeats namely, LSR and SSR of pOSMlc-2	95
3.2.7	Methylation status of pOSMlc-2 at three different stages in rice (namely embryo, seedling and leaf)	96
3.2.8	Presence and distribution of pOSMlc-2 sequence in 4 genome types of rice, namely AA, CC, CCDD and EE	105

SECTION III

3.3	Assessment of genomic variability among nineteen cultivars of rice using the dispersed repeat sequence pOSMlc-2	106
3.3.1	Agarose gel electrophoresis patterns of DNA of nineteen cultivars with a few restriction endonucleases namely, <u>EcoRV</u> , <u>EcoRI</u> , <u>XbaI</u> , <u>DraI</u> , <u>ScaI</u> , <u>and HindIII</u>	108
3.3.2	Southern blot hybridization patterns of P-32-labelled pOSMlc-2 insert with the restriction enzyme digests of DNAs of 19 cultivars with <u>EcoRV</u> , <u>EcoRI</u> , <u>XbaI</u> <u>DraI</u> , <u>ScaI</u> and <u>HindIII</u>	115

3.3.3	Analysis of the dendrogram generated based on Nei's F values for polymorphic enzyme-probe combinations	126
	References	129

CHAPTER 4 : DISCUSSION

4.1	Present status of repeats in plants	131
4.2	Highlights of work on repeat sequences in rice in my laboratory and my own research	133
4.3	Why construct a partial <u>PstI</u> rice genomic library ? Low level of 'C' methylation in rice repeats	135
4.4	Random organization of pOSMLC-2 sequence in the rice genome	138
4.5	Differential organization of two sub-repeats of pOSMLC-2 : probable role as transposable elements	140
4.6	Predominance and stage specific adenine methylation in rice genome	142
4.7	Role of pOSMLC-2 as a RFLP marker	144
	References	149

CURRICULUM VITAE

ACKNOWLEDGEMENTS

Dr. P.K. Ranjekar has been my guide for my Ph.D. work. It was a great privilege to work under his guidance. I do not wish to be formal. But I feel happy to record and express very sincerely my deep sense of gratitude for the help, encouragement and advice I received from him from time to time.

Dr. (Mrs.) Vidya Gupta was kind enough to go through all the portions of my thesis. Her detailed and incisive comments were extremely useful to me. Discussion with her has been always rewarding.

Dr. (Mrs.) Meena Lagu and Prafulla rendered me generous help in multifarious activities concerning my experimental work. I am indeed pleased to express my indebtedness to them. I am immensely thankful to Mr. Dinesh Lagu and Prafulla for their help in writing the computer program which helped me to analyse my data.

I had an opportunity to visit and work at Perpignan, France under the Rockefeller Foundation Fellowship program, for a period of three months (1989). Dr. M. Delseny, an internationally reputed scientist in the area of molecular biology, is the Head of this laboratory and is assisted by Dr. Françoise

Grellet, a well known scientist in her own right. I was fortunate in having guidance and encouragement from both Dr. Delseny and Dr. Grellet. It was inspiring to have intellectual interaction with them.

The work towards completion of Ph.D. thesis is usually a pleasant experience because it gives one an opportunity to interact intellectually with persons having similar academic interests. In this connection, I will like to make special mention of Madhu Dhar, Madhur, Kalpana, Shubhada, B.G. Patil, Ramakrishna, Vivek, Medha, Radha, Lalitha, Suman, Gagan, Mohini, Vishwas, Urmila, Sujata, Savita, Renuka, and two senior scientists Dr.Nirmala and Dr.Hendre. I am also glad to record my appreciation for the help I received from Shirish during my stay in France. The technical assistance given by Kamala is gratefully acknowledged.

I am deeply indebted to Dr. John Barnabas for giving me encouragement and providing laboratory facilities at National Chemical Laboratory, as the Head of Division of Biochemical Sciences.

My thanks are due to Prof B.S. Shah, Head Department of Agricultural Botany, College of Agriculture, Pune for providing the glass house facilities.

It was nice of Mr. Ajit Prabhakar Ranjekar to have undertaken and completed the typing of my

thesis. I appreciate his sincerity and boundless enthusiasm in preparing the typescript. I thank Miss Usha Deshpande for typing the bibliography of my thesis.

The financial help in the form of a Research Fellowship from the Council of Scientific and Industrial Research (CSIR), New Delhi is gratefully acknowledged.

And one more set of persons to thank: those who are behind the curtain but who nevertheless play significant role, viz the members of my family: Father Mother, Vasanti and Varsha.


Vaijayanti V. Pethe

DECLARATION

Certified that the work incorporated in the thesis "A molecular approach for the assessment of variability in indica rices," submitted by Miss Vaijayanti V. Pethe was carried out by the candidate under my supervision. Such material as has been obtained from other sources has been duly acknowledged in the thesis.



(P.K. Ranjekar)

Research Guide

LIST OF ABBREVIATIONS

bp	:	base pairs
kbp	:	kilo base pairs
gm	:	grams
mg	:	milligrams
ug	:	micrograms
pg	:	picograms
ng	:	nanograms
ml	:	millilitre
ul	:	microlitre
U	:	Units of enzyme
Mol.wt.	:	molecular weight
nm	:	nanometer
Ci	:	Curies
mCi	:	milli Curie
uCi	:	micro Curie
cpm	:	counts per minute
rpm	:	revolutions per minute
O.D.	:	Optical Density
A	:	Adenine
T	:	Thymine
G	:	Guanine
C	:	Cytosine
dCTP	:	deoxycytidine 5'-triphosphate

dATP	:	deoxyadenosine 5'-triphosphate
dGTP	:	deoxyguanosine 5'-triphosphate
dTTP	:	deoxythymidine 5'-triphosphate
Tris	:	Tris-hydroxymethyl amino methane
SDS	:	Sodium dodecyl sulphate
EDTA	:	Ethylene diamine tetra acetic acid
TAE	:	Tris-acetate-EDTA buffer
TBE	:	Tris-borate-EDTA buffer
T _m	:	Melting temperature
DTT	:	Dithio threitol
cv.	:	cultivar
BSA	:	Bovine serum albumin
BPB	:	Bromophenol blue
C	:	degrees Celsius
ssp.	:	sub species

LIST OF FIGURES

FIG.NO.	TITLE	PAGE NO.
3.1 A	Agarose gel (1%) electrophoresis of <u>PstI</u> restricted Malkolam Rice DNA	77
3.1 B	Agarose gel (1%) electrophoresis of <u>PstI</u> restricted pUC18 DNA	77
3.2	Screening of partial genomic library for low copy & repeat sequences	78
3.3 A	Agarose gel electrophoresis of random genomic clones digested with <u>PstI</u>	79
3.3 B	Agarose gel electrophoresis of random genomic clones digested with <u>EcoRI</u> + <u>HindIII</u>	80
3.3 C	Agarose gel electrophoresis of pOSM1C-2 DNA digested with <u>PstI</u>	80
3.4 A	Agarose gel (1%) electrophoresis of Malkolam DNA digested with different units of <u>PstI</u>	83
3.4 B	Autoradiogram showing Southern hybridization of P-32-labelled pOSM1C-2 insert with digests as shown in Fig. 3.4 A	83

3.5 A	Agarose gel (1%) electrophoresis of Malkolam DNA digested with different restriction enzymes	86
3.5 B	Autoradiogram showing Southern hybridization of P-32 -labelled pOSM1C-2 insert with DNA digests as described in 3.5 A	86
3.6	Copy number estimation of pOSM1C-2 repeat sequence in rice genome	87
3.7 A	Agarose gel (1%) electrophoresis of pOSM1C-2 DNA digested with enzymes like <u>EcoRI</u> , <u>HindIII</u> , <u>AccI</u> , <u>XbaI</u> , <u>BamHI</u> , <u>SmaI</u> and <u>KpnI</u>	91
3.7 B	Agarose gel (1%) electrophoresis of pOSM1C-2 DNA digested with <u>PstI</u> , <u>PstI</u> + <u>EcoRV</u> , <u>PstI</u> + <u>ClaI</u> , <u>PstI</u> + <u>BglI</u> and <u>PstI</u> + <u>HindII</u>	91
3.7 C	Agarose gel (1%) electrophoresis of pOSM1C-2 DNA digested with <u>PvuII</u> and <u>ScaI</u>	91
3.7 D	Restrictase map of pOSM1C-2 DNA	92
3.8 A	Agarose gel (1%) electrophoresis of pOSM1C-2 DNA digested with double enzymes, namely, <u>PstI</u> + <u>AccI</u> , <u>PstI</u> + <u>EcoRI</u> and <u>PstI</u> + <u>SmaI</u>	94
3.8 B	Autoradiogram showing Southern hybridization of P-32-labelled <u>AccI</u> - <u>PstI</u> fragment	

	(0.450 kbp) with the DNA digests as described in Fig. 3.8 A	94
3.9 A	Autoradiogram showing Southern hybridization of P-32-labelled <u>PstI-AccI</u> (1.1 kbp) fragment and <u>AccI-PstI</u> (0.450 kbp) fragment to <u>EcoRI</u> digest of Malkolam DNA	97
3.9 B	Autoradiogram showing Southern hybridization of P-32-labelled <u>PstI-AccI</u> (1.1 kbp) fragment and <u>AccI-PstI</u> (0.450 kbp) fragment to <u>PstI</u> digest of Malkolam DNA	97
3.10 A	Agarose gel (1%) electrophoresis of Malkolam embryo DNA digested with <u>MboI</u> , <u>Sau3AI</u> , <u>DpnI</u> , <u>MspI</u> and <u>HpaII</u>	100
3.10 B	Agarose gel (1%) electrophoresis of Malkolam shoot DNA (8-10 days old seedlings) digested with <u>MboI</u> , <u>Sau3AI</u> , <u>DpnI</u> , <u>MspI</u> and <u>HpaII</u>	100
3.10 C	Agarose gel (1%) electrophoresis of Malkolam leaf DNA (1-2 months old) digested with <u>MboI</u> , <u>Sau3AI</u> , <u>DpnI</u> , <u>MspI</u> & <u>HpaII</u>	100
3.11 A	Autoradiogram showing Southern blot hybridization of P-32-labelled pOSM1C-2 insert with DNA digests as described in Fig. 3.10 A	102
3.11 B	Autoradiogram showing Southern blot	

	hybridization of P-32-labelled pOSM1C-2 insert with DNA digests as described in Fig.3.10 B	102
3.11 C	Autoradiogram showing Southern blot hybridization of P-32-labelled pOSM1C-2 insert with DNA digests as described in Fig. 3.10 C	102
3.12 A	Agarose gel (1%) electrophoresis of rice entries belonging to AA,CC,CCDD and EE genome types digested with <u>ScaI</u>	104
3.12 B	Autoradiogram showing Southern blot hybridization of P-32-labelled pOSM1C-2 insert with the DNA digests as described in Fig. 3.12 A	104
3.13 A	Photograph of rice germplasm grown in pots in the glass house of Agriculture College, Pune	109
3.13 B	Photograph showing the leaf tissue used for DNA extractions	109
3.14	Agarose gel (1%) electrophoresis pattern of high mol. wt. genomic DNAs isolated from leaf tissue (1.5-2 months old) of rice cultivars	110
3.15 A	Agarose gel (1%) electrophoresis pattern of DNAs from 19 rice cultivars digested with <u>EcoRV</u>	112

3.15 B	Agarose gel (1%) electrophoresis pattern of DNAs from 19 rice cultivars digested with <u>EcoRI</u>	113
3.15 C	Agarose gel (1%) electrophoresis pattern of DNAs from 19 rice cultivars digested with <u>XbaI</u>	114
3.15 D	Agarose gel (1%) electrophoresis pattern of DNAs from 19 rice cultivars digested with <u>DraI</u>	114
3.15 E	Agarose gel (1%) electrophoresis pattern of DNAs from 19 rice cultivars digested with <u>ScaI</u>	116
3.15 F	Agarose gel (1%) electrophoresis pattern of DNAs from 19 rice cultivars digested with <u>HindIII</u>	116
3.16 A	Southern blot hybridization of P-32-labelled pOSM1C-2 insert with DNA digests as described in Fig. 3.15 A	118
3.16 B	Southern blot hybridization of P-32-labelled pOSM1C-2 insert with DNA digests as described in Fig. 3.15 B	120
3.16 C	Southern blot hybridization of P-32-labelled pOSM1C-2 insert with DNA digests	

	as described in Fig. 3.15 C	122
3.16 D	Southern blot hybridization of P-32- labelled pOSM1C-2 insert with DNA digests as described in Fig. 3.15 D	123
3.16 E	Southern blot hybridization of P-32- labelled pOSM1C-2 insert with DNA digests as described in Fig. 3.15 E	125
3.17	A computer-generated dendrogram of 18 varieties of <u>O. sativa</u> subspecies indica by RFLP analysis	127

LIST OF TABLES

Table No.	Title	Page No.
1.1	Characteristic features of sativa subspecies	12
1.2	Rice species with their genome type and characteristics	17
2.1	Rice germplasm used for the assess ment of genetic variability	41-42

SYNOPSIS

A brief summary of work on repeated DNA sequences in rice

Rice is the most important food crop of the developing world and a staple food for more than half the world's population. It belongs to the family Gramineae, and genus Oryza, which includes 20 wild species and two cultigens O. sativa and O. glaberrima which have the same genome AA, but were independently domesticated (1). While O. sativa is the major cultivated rice on the Asian continent; O. glaberrima is cultivated in Western Africa. O. sativa has three dominant sub species namely indica (origin:India), japonica/sinica (origin:China), javanaica/bulu (origin:Indonesia). Rice genome has been classified into six types : AA, BB, CC, DD, EE and FF (2). It has twelve chromosomes ($n=12$), containing only 5.8×10^5 kbp per haploid genome, and has 52% of repetitive DNA as determined by Cot analysis (3,4).

The organization of repeated DNA sequences in rice O. sativa has been studied by Gupta et al. (5) and Dhar et al. (6). Recently, a satellite and a moderately repeated DNA sequence from O. sativa have been used as hybridization probes to study interrelationship of cultivated rice species with the O. perennis complex (7). In O. sativa, T.Wu and R.Wu (8) have identified a novel transcribed repetitive DNA sequence that shows

homology to both 5S and tRNA genes and also resembles 5S RNA gene in organization . DNA sequences specific to 4 genome types of rice namely AA, CC, EE and FF have been isolated and characterized (9). Attempts are also being made to use these genome specific sequences as molecular markers to study the origin and evolutionary relationships among various genomes and species of rice. Higher eukaryotic ribosomal genes are organized in tandem repeats. Variation of the intergenic spacer region of the ribosomal DNA is observed in cultivated and wild rice species (10). Recently, restriction fragment length polymorphism of rDNA spacer has been studied in Orzya using a cloned rice rDNA probe. While African rice of the species O. glaberrima do not display any rDNA size variation, wild rice species show extensive variation (11).

In the present work, an in-depth analysis of a rice dispersed repeat has been carried out with an aim of gaining more information about its organization, structure, abundance, methylation status at three different stages in rice and its potential use as a molecular marker for studying genomic/genetic variability among rice germplasm.

Thesis plan

The thesis is organized as follows

Chapter 1: Introduction

Here, I have compiled most of the available recent literature on rice repeated DNA sequences under the title: Novel features of repeated DNA sequences in the rice genome.

Chapter 2: Materials and Methods

During my Ph.D. work, I have handled an array of molecular techniques starting from DNA extractions up to DNA sequencing. The detailed protocols of all these experimental procedures are included in this section.

Chapter 3: Results

This section is divided into following three parts.

SECTION I : Construction of a partial genomic library of rice.

SECTION II : Characterization of a dispersed repeat sequence pOSM1C-2 in rice genome.

SECTION III : Assessment of genomic/genetic variability among 18 cultivars of rice using dispersed repeat sequence pOSM1C-2.

Chapter 4: Discussion

Here, I have made an attempt to discuss the results which I have described in three sections of chapter 3. Mainly I have discussed the role of repeat sequence

posM1C-2 in genome organization and its potential role in analysis of genomic variability among selected rice varieties.

Summary of work

Construction of a partial genomic library of rice

To start with, I constructed a partial genomic library of rice O. sativa cv. Malkolam in a plasmid vector pUC18. Malkolam (an indica rice) was selected because it is locally cultivated and is a land race. It has a conserved gene pool which is quite different from the modern cultivars and is likely to contain genes of commercial value. It has a few elite characters like early maturity, medium slender (MS) type of grain and good cooking qualities. The vector pUC18 was of choice because of its easy screening procedures and its direct use in sequencing and expression studies. After selecting the recombinants, I screened the library for repetitive/low copy clones by colony hybridization using P-32-labelled genomic DNA against the colony filters. Since my interest was mainly in repeated DNA sequences, I specifically selected those colonies which showed intense signals on the autoradiogram. I checked the insert sizes in a few of these random genomic clones by digestion with PstI and the size ranged from 0.2 to 2 kbp. The library

represented 55% low copy clones and 45% repeat sequences and served as a source of required types of probes for further studies.

Characterization of a dispersed repeat sequence pOSM1C-2

In my initial analysis of the PstI probes, I identified one dispersed repeat sequence in Malkolam rice genome. During the characterization of repeat, I digested the genomic DNA with different units of PstI (ie. setting up partial and complete digestions of genomic DNA), and challenged these digests with P-32-labelled pOSM1C-2. I observed that the repeat hybridized giving prominent bands superimposed on a background smear. This pattern is expected for a repeat sequence that is interspersed with either single copy or repeated sequences. Interestingly thus, the repeat appeared to be dispersed or non-tandem in nature. Since very little was known with respect to dispersed repeats in rice, I decided to concentrate on this particular repeat sequence and study its nature, organization, abundance and distribution in the rice genome. This repeat, as I named it pOSM1C-2, gave a moderately intense signal on the colony autoradiogram and had an inserted DNA fragment of 1.55 kbp.

After confirming that pOSM1C-2 is a dispersed repeat, its organization was studied in the rice

genome. For this purpose, genomic DNA was digested with 15 different restriction endonucleases and probed with pOSM1C-2 repeat sequence. In the case of BamHI, SacI, SalI and BstEII digests, the repeat hybridized to high molecular weight relic regions in the genomic digests, while distinct hybridizing bands of different molecular weights with a background smear were observed in case of EcoRI, BglII, HaeIII, BglI, PvuII, ScaI and HindIII digests indicating that the former set of enzymes has rare sites close to pOSM1C-2 repeat while the latter set has sites in the vicinity of region homologous to pOSM1C-2 sequence in the rice genome.

The next criterion of characterisation was to quantitate the abundance of the repeat in the genome. A filter containing dilution series of this repeat and three different concentrations of genomic DNAs was hybridized with the P-32-labelled pOSM1C-2 insert. When the autoradiogram was scanned with a densitometric scanner, the repeat showed the presence of approximately 150 copies in the rice haploid genome.

The restriction mapping work revealed the occurrence of a single site for enzymes like EcoRI, ScaI, HincII, AccI and SmaI in the repeat while absence of sites for BamHI, HindIII, ClaI, EcoRV, PvuII, BglI, BglII and SacI.

To investigate homologies within the repeat

sequence, smaller AccI-PstI fragment (0.450 kbp) was used as probe against the double digests of pOSM1C-2 with PstI-AccI, PstI-EcoRI and PstI-SmaI. The probe showed intense hybridization to itself and very low homology in the other part of the repeat i.e. PstI-AccI larger fragment (1.1 kbp). Further studies revealed that these two subdivisions of the repeat namely AccI-PstI (0.450 kbp) and PstI-AccI (1.1 kbp) showed differential organization and abundance in the genome.

Assessment of genomic variability among 18 cultivars of rice using the dispersed repeat sequence pOSM1C-2

The main aim of this study was to explore the potential role of pOSM1C-2 as a RFLP marker for distinguishing rice cultivars. It is now well known that RFLP technology has opened a door to detect and measure genetic variation in plants, in a manner not previously possible (12). Genomic DNAs from 19 elite cultivars were digested with 6 restriction enzymes namely EcoRI, EcoRV, DraI, ScaI, XbaI and HindIII and probed with P-32-labelled pOSM1C-2 insert. The probe was informative in revealing polymorphism with three enzymes namely EcoRV, DraI, and ScaI and deciphered different number of alleles with each enzyme. The similarity indices between the cultivars were calculated according to Nei's F statistic (13) using computer software developed in our laboratory. A dendro-

gram was then constructed by the same software using the unweighted pair group method with arithmetic mean (UPGMA) for the variability analysis. For the construction of the dendrogram, polymorphic probe-enzyme combinations of only 18 cultivars were selected. One cultivar namely Pusa-33 was excluded all the time since the autoradiography picture with this cultivar was inconclusive. From the dendrogram, it was possible to group the cultivars into two major clusters, with smaller sub-clusters within them. Interestingly enough, this grouping agreed with the pedigrees of these cultivars. These interrelationships deciphered amongst closely related indica cultivars will prove useful in the planning of future rice breeding programs in Maharashtra state and India.

REFERENCES

1. Chang, T.T. (1984). *Science*, 224: 251-256.
2. Hiroka, M. (1984). In: Tsunoda S. Takahashi N (eds) *Biology of rice*. Jap Sci Press, Tokyo, 3-30.
3. Deshpande, V.G., Ranjekar, P.K. (1980). *Hoppe-Seyler's Z. Physiol. Chem.*, 361: 1223-1233.
4. Zhou, G.Y. (1986). In: *Rice Genetic Proc. Int. Rice Genet. Symp.* Island, Manila, 867-876.
5. Gupta, V.S., Gadre S.R., and Ranjekar, P.K. (1981). *Biochim. Biophys. Acta.*, 656: 147-154.
6. Dhar, M.S., Dabak, M.M., Gupta, V.S. and Ranjekar,

- P.K. (1988). Plant Science, 55: 43-52.
7. Pental, D. and Barnes, S.R. (1985). Theor. Appl. Genet., 70: 185-191.
8. Wu, T. and Wu, R. (1987). Nucleic Acids Res., 15: 5913- 23.
9. Zhao, X., Wu, T., Xie, Y. and Wu, R. (1989). Theor. Appl. Genet., 78: 201-209.
10. Sano, Y. and Sano, R. (1990) Genome, 33: 209-218.
11. Cordes, F., Second, G. and Delseny. M (1990). Theor. Appl. Genet., 79: 81-88.
12. Wang, Z.Y. and Tanksley, S.D. (1989). Genome, 32: 1113- 1118.
13. Nei, M. (1987). In: Molecular evolutionary genetics. University Press, New York. 106-107.

CHAPTER ONE

NOVEL FEATURES OF REPEATED DNA SEQUENCES IN RICE GENOME

1.1 Classification of rice

Rice, Oryza sativa L., belongs to the genus Oryza and has about 20 well defined species and two cultigens, namely Oryza sativa L. and Oryza glaberrima Steud (1,2,3). While Oryza sativa L. is the major cultivated rice on the Asian continent, Oryza glaberri ma is cultivated in Western Africa. Both are diploid species with somatic chromosome number (2n) equal to 24, morphologically distinguishable from one another and reproductively isolated (4,5). The progenitors of cultivated species are assumed to be O. rufipogon for O. sativa and O. barthii for O. glaberrima, since only these wild relatives produced fertile progeny when crossed with respective cultivated forms (6).

The Asiatic cultivated rice, belonging to the sativa complex, is classified on the basis of morphophysiological characters into 3 groups or ecotypes, namely, indica, japonica and javanica (7,8). Indica rices are distributed in the tropics and subtropics of the Indian subcontinent, south east Asia; and South and Central China. Japonica (or sinica) rices have originated from China; while javanica (bulu) rices are grown in Indonesia. The distinguishing characters of all the three sub-species are depicted in Table 1.1. G. Second (9,10) has reported four major groups namely japonica, indica, intermediate japonica and intermedi-

ate indica and studied the molecular similarities between these different species using isozyme polymorphisms at 24 loci for 181 plants. Glaszmann (11,12) classified rice varieties into six groups based on genetic affinities within groups and genetic divergence between groups. He investigated the allelic variation at fifteen isozyme loci in about 1700 rice varieties from all the rice growing countries of Asia. According to his classification, group I includes indicas, while group VI corresponds to temperate japonicas. Groups II and V rices are distributed from Burma to Iran and include intermediates between japonica and indica. Groups III and IV are represented by some deep water rices of Bangladesh and Northeast India.

Rice genetic resources can be classified as primary, secondary and tertiary gene pools (13). Primary gene pool consists of cultivated varieties, land races and weedy populations. Closely related wild species comprise secondary gene pool. Wild rice species with AA genome such as O. nivara, O. rufipogon, O. longistaminata, and O. barthii and the African cultigen O. glaberrima fall in this category. These closely related wild species prove to be very important source of genes for rice breeding and improvement, since they have homologous genomes and there is normal chromosome

TABLE No 1.1

Characteristic features of sativa subspecies (13,14)

Character	Indica	Javanica	Japonica

Grain:			
Shape	Narrow	Large	Short
Awns	usually absent	usually present	Short
Degree of grain shedding/scattering:	Easy	Difficult	Difficult
Plant status:			
Height	Tall	Tall	Short
Colour	Light green	Light green	Deep green
Texture	Hard	Hard	Soft
Panicle:			
Length	Medium	Long	Short
Branching	Intermediate	Many	Few
Density	Moderate	Moderate	High
Weight	Light	Heavy	Heavy
Tillers:			
Number	Large	Small	Large
Habit	Spreading	Erect	Erect
Flag leaf	Long,narrow	Long,wide	Short,narrow

pairing in the F1 hybrids, thereby reducing sterility problems. Thus the utilization of these wild species is becoming an important part of current rice breeding programs (15,16,17). The more distantly related wild species such as O. punctata, O. officinalis, O. australiensis which have nonhomologous genomes form the rice tertiary gene pool. These species are difficult to cross with the AA genome species and the F1 hybrids are completely sterile as there is no or less chromosome pairing. Recent investigations have shown that some of these distantly related species are highly resistant to insects, pests as well as many diseases and have several other genes of agronomically important traits useful for rice improvement (18,19).

1.2 Content and gross arrangement of repeated sequences in rice genome

The early molecular studies have revealed several interesting features of rice genome. First and foremost, Bennett and Smith (20,21) have shown that the haploid DNA content of rice ranges from 0.6-1 pg. Such low DNA content of rice, which falls just next to Arabidopsis thaliana (C = 0.06 pg), is an attractive attribute in a higher plant species like rice. It is well established that majority of agronomically important plant species have large genomes and contain very high percentage of repetitive sequences

(20-26).

It was as early as 1980, when Deshpande and Ranjekar (27) showed that the repetitive DNA content accounted for 52% of the total rice genome. This observation has been recently confirmed by McCouch et al. (28). During 1975-1985, study of DNA sequence organization was one of the most important aspects of genomic analysis in eukaryotes and these data have been reviewed extensively by Thompson and Murray (29), Walbot and Goldberg (30) and Flavell (22). These studies have revealed two general patterns of genome organization namely i) "short period" interspersion pattern in which 100-300 bp long repeated sequences are interspersed with longer single copy sequences of approximately 1000-2000 bp (31) and ii) "long period" interspersion pattern where approximately 5000 bp long repeated sequences are interspersed with single copy sequences that may be as long as 35 kbp (32). During the study of genome structure and genome organisation in higher plants from our laboratory, it has been shown that plants exhibit a great diversity in genome organisation patterns, mainly due to variations in the lengths of interspersed repeat DNA sequences (33-37). Interestingly, Gupta et al. (33) have reported that rice is a unique genome in consisting of long stretches of

repeat and single copy sequences which are not interrupted by each other. Intrigued by this novel genome organization, Dhar et al. (38) have carried out further studies to show that repeated DNA sequences are not interrupted by single copy DNA even at a much higher fragment length of 20 kbp. Their analysis has also shown that majority of the repetitive sequences are unusually G+C rich and have high T_m values. Based on their detailed analysis, they have suggested that long and short repeats of the highly repetitive sequences are independent and belong to different repeat families (38).

1.3 Genome specific and species - specific repetitive DNA sequences in rice

Based on the cytogenetic analysis of the chromosome pairing behaviour, rice genome has been classified into six types, namely A-F, including diploid as well as tetraploid species (39,40). A few rice species with their genome types are given in Table No.1.2.

O. sativa, O. glaberrima and their closely related wild species such as O. nivara, O. rufipogon and O. longistaminata share a common genome designated as AA (recognized as A genome complex) while the O. officinalis is a C genome complex (14). The species

with AA genome show regular pairing and recombination (17) and gene transfer can be achieved between these species without much difficulty. For instance, a very useful gene for grassy stunt virus resistance was transferred from O. nivara to O. sativa (41). Molecular similarities of various types of genomes have been investigated using isozyme polymorphisms, and chloroplast and nuclear DNA restriction fragment length polymorphisms (9-12, 42). To date, the occurrence of such genome or species-specific sequences has been reported in several plant species (43-47). The uses of such sequences as chromosome-specific markers, in the detection of chromosome additions or substitutions, as molecular markers for tracing alien chromatin introgression and study of phylogenetics have been demonstrated (48-53).

The repetitive DNA sequences which occur only in a given genome-type of rice and are absent in all the other genome types are referred to as genome-specific sequences. Zhao et al. (54) have cloned and characterized repeated DNA sequences specific to AA, CC, EE and FF genomes. To begin the analysis, four repetitive DNAs, designated as pOa4 (from O. australiensis (EE), 511 bp), pOo2 (from O. officinalis (CC), 366 bp), pOs48 (from O. sativa cv. Labelle (AA), 355 bp) and pOb1 (from O. brachyantha (FF), 155 bp) were used as probes for genomic blot hybridization on 37

TABLE No. 1.2 (13,55,56)

Rice species with their genome types and characteristics

S.No.	Species	2n	Genome	Distribution	Characteristics
1.	<u>O. sativa</u>	24	AA	Worldwide	Cultivated rice everywhere except Africa
2.	<u>O. glaberrima</u>	24	AA	Tropical west Africa	Resistant to GLH, early vegetative vigour.
3.	<u>O. punctata</u>	24 48	BB BBCC	Tropical Northeast Africa	Resistant to WBPH, BPH.
4.	<u>O. minuta</u>	48	BBCC	Philippines	Resistant to Blast and BB
5.	<u>O. eichingeri</u>	24	CC	Tanzania, Uganda	Resistant to BPH, WBPH & GLH
6.	<u>O. latifolia</u>	48	CCDD	Central and South America, West Indies	Tall and perennial
7.	<u>O. australiensis</u>	24	EE	Western & Northern Australia	Resistant to BPH
8.	<u>O. brachyantha</u>	24	FF	West Tropical & Central America	Resistant to rice whorl maggot and stem borer.

GLH : Green leaf hopper
WBPH : White backed plant hopper

BPH : Brown plant hopper
BB : Bacterial blight

577.21:633.18(043)
PET

TH-639

rice entries, including domestic and wild varieties of representative genome types. The repeats pOa4, pOo2, pOs48 and pOb1 were specific to EE, CC, AA, and FF genomes respectively and were largely present as tandem repeats of upto 16 kbp in the genome. The copy numbers of these four repetitive sequences in the rice genome were determined by quantitative slot blot hybridization and were found to be 2,000 (for Os48), 80,000 (for Oa4), 170,000 (for Oo2) and 184,000 (for Ob1). In addition, the copy number of Os48 was found to vary among several related species of AA genome. An interesting observation made by these authors reveals that most traditional cultivars tend to have higher copy numbers than the modern improved cultivars. When AA genome-specific repeat sequence namely Os48 was used as probe, it gave strong hybridization with EE genome, much weaker hybridization with CCDD genome, while no hybridization with BB and BBCC genomes. Weak hybridization to CCDD genome suggests that the CCDD and EE genomes are more closely related to each other than other genome types (namely BB and BBCC). The authors have thus explored the usefulness of these sequences in classifying unknown species of wild or domestic rice and as molecular markers for studying plant genome evolution and species divergence (54).

A CCDD genome specific sequence and a clone

from BB genome type which showed hybridization to BB, CC, BBCC, and CCDD genomes but no hybridization to AA, EE, and FF genome types of rice have been reported (55). Characterization of these sequences in terms of their copy number, length of the repeat unit, genomic organization and sequence analysis are in progress. T. Wu (57) also proposes to isolate genome-type-general repetitive DNA sequences from rice. Simultaneously, a repetitive DNA sequence, CC-1 (insert: 190 bp) has been isolated from O. officinalis and is found to be present as tandem repeats in CC diploid and CC tetraploid genomes. This family of sequences is present in AA and BB, BBCC or CCDD genomes but absent in EE and FF genomes. The relationship drawn by the authors is that CC genome type is most closely related to AA and most distantly related to EE and FF genome types. This analysis thus provides an amplified view of evolutionary relationships among species and genome types of the rice genus. Moreover, these sequences could be used as molecular markers for monitoring the chromosome transfer from particular genome type to others. The methods of using genome-type-specific repetitive DNA sequence in classifying unknown rice species would be simpler and more accurate than other methods which involve comparison of morphological traits, determination of fertility of

crosses or cytological analysis and pairing at meiosis.

Repeated sequences from O. officinalis have been cloned and sequenced (58). Three of the sequences are organized as tandem repeats of units 374 bp, 367 bp and 193 bp. Another one is a dispersed repeat 520 bp long. Two sequences (approximately 370 bp and 500 bp) clearly show CC genome specificity. In addition, 500 bp sequence discriminates between O. officinalis (CC) and O. eichingeri (CC) in which it is much less represented. Although the 370 bp sequence is represented in different species of BBCC and absent in CCDD genome, the 500 bp sequence is apparently absent in both the BBCC and CCDD genomes. Repeats from the rDNA spacer region from rice O. sativa (Japonica) have been found to exhibit genome specificity. Two types of repeats are reported; 250 bp repeat which hybridizes to AA, BB, CC, and BBCC but not to CCDD or EE genomes and a 80 bp sequence which specifies the AA genome.

Species-specific repetitive DNA sequences which serve as markers in rice backcross breeding programs have been isolated from the genomes of O. minuta (BBCC), O. punctata (BB) and O. officinalis (CC) (59). Three sequences from O. minuta and one sequence from O. punctata show specificity at species or genome level, and do not cross-hybridize with genome-specific sequences that have been isolated by other rice researchers. The genome

organization, chromosome distribution and taxonomic occurrence of these sequences are under investigation. Recently, genomic DNA sequences specific to wild species of rice O. minuta have been isolated and their role in analysis of alien chromatin introgression in the progeny of interspecific cross of O. minuta x O. sativa is investigated (60).

1.4 Transcribed repetitive sequences in rice

A novel, transcribed repetitive DNA sequence from rice O. sativa cv. Labelle has been isolated & sequenced (61). The sequence RC48 (cDNA clone of 355 bp) gives hybridization patterns resembling that of highly repeated satellite DNA. In the genomic blot of PstI digests of four rice cultivars, namely, Labelle, IR36, Calrose 76 and O. australiensis, the probe shows ladders of 355 bp repeats upto 20 units (7,000 bp). Northern blot analysis, using RC48 sequence as probe, gives at least 4 distinct hybridizing bands with sizes ranging from 1-4 kbp. The sequence of RC48 DNA is interesting in that (i) two segments of RC48 (positions 1-124, 182-302) show 45% and 42% sequence homology, respectively, to wheat 5S RNA (62); (ii) four segments of RC48 show 41%, 45%, 53%, and 43% sequence homology, respectively, with wheat tRNA. Moreover, RC48 includes the polymerase III-type promoter sequence (63,64,65). In addition to the

above sequence homology, the rice repetitive DNA also resembles the 5S RNA gene in organization. This observation raises the possibility that RC48 may be evolved from the common ancestor of 5S RNA and tRNA. This is the first report showing that certain repetitive sequences in plants are transcriptionally active. A detailed analysis of the three highly transcribed sequences (namely, Os48 and the sequences immediately flanking the Os48 repeat) involving the structural organization, relative levels of transcription, tissue and developmental patterns; and sequence elements that regulate the transcription of these sequences is proposed (57).

1.5 Transposable elements

Transposable elements (mobile genetic elements or transposons or insertion elements) are specific DNA fragments which are able to move from one site to another, either on the same chromosome or different one. They are responsible for generation of mutations, or chromosomal rearrangements and thus affect gene expression (66,67). They are valuable tools in several species for the isolation of genes via tagging strategies. A characteristic feature of transposons is presence of flanking long direct repeats (long terminal repeats(LTRs) 200-1000bp) which are sup-

posed to be the substrates for transposase activity (68).

Although, molecular cloning of transposable elements in plants has been described in a number of species such as maize, wheat, Arabidopsis, and soy bean, they have been well characterized at both genetic and molecular levels in only two species namely, Antirrhinum majus and Zea mays (67). The first, genetically defined by Mc Clintock (69) as the AcDs system in maize, has been studied in detail at molecular level. The Ac element (Activator, 4.5kbp) is named for its ability to activate chromosome breakage at another locus, the Ds locus (Dissociator, size variable). Ac is able to transpose itself from one location to another location by itself, and hence it is also defined as autonomous transposable element, while Ds is a non-autonomous element because it requires Ac for its movement.

Wu et al. (70) have embarked on a project to isolate transposable elements from different rice varieties including wild. They have recently reported first transposable element in rice (O. sativa) which they refer to as Mu element. Nucleotide sequence of this element reveals that it contains a pair of 217 bp terminal inverted repeats and it is flanked

by two 9 bp direct repeats. The GC content of the internal portion of this element is unusually high. The authors have reported its presence in all eighteen lines of rice including wild varieties, though copy number shows variation in different lines.

1.6 Rice repeats showing copy number variation and differential methylation

Kikuchi et al. (71) and Zheng et al. (72) have reported that rice repeated sequences show a copy number variation during dedifferentiation and redifferentiation processes. Dedifferentiation is the process when callus is induced to form from explants by addition of auxins, while in redifferentiation, plants are regenerated from callus by changes in the hormonal composition in the medium. Here one repeat sequence (pRB 301) has shown 50 times increase in copy number during callus formation, whereas the other repeat sequence (pBR 401) decreases by 10,000 fold in the same process. Such variations during dedifferentiation and redifferentiation are also referred to as modulation.

Gupta et al. (73) and Dhar et al. (74)

have reported molecular cloning, restriction enzyme analysis; and differential methylation of a cloned repeat element (pRL7) during the transition from embryo to shoot stage. This repeat unit is present in a well-organised manner in the embryo DNA, whereas in case of shoot DNA, many copies of this sequence are present in a varied environment with random distribution, while some are organized to give a particular band pattern. Some form of genomic rearrangement of this sequence may be partly due to the changes in methylation of this repeat unit. In plants, this is probably the first report of differential methylation of a repeat sequence without any external inducing factor.

1.7 Ribosomal genes in rice

Initial experiments on rice ribosomal genes have indicated that rDNA is heterogenous i.e. the length of the major repeat unit varies from 7.6-8.8 kbp (75,76). One group, however, has reported 8.8 kbp rDNA repeat unit as the major one, occurring about 1060 times in the diploid genome of rice variety IR20 (77). The heterogeneity of the repeat unit is due to the presence of differing number of tandem repeats within the non-transcribed spacer. Such variations in the length of the rDNA unit is a common feature in plants (78). Takaiwa et al. (79) have cloned a major class

(7.8 kbp) of rice repeating units and determined the entire nucleotide sequence of the spacer region between 17S and 25S rDNA. In fact, the same group of scientists at Nagoya, Japan is the first to determine the complete nucleotide sequence of 17S and 25S rDNA from a higher plant like rice (78,80). Their results reveal that 17S rDNA is 1812 bp, 25S rDNA is 3777 bp and 5.8S rDNA is 163-164 bp long. The length of the internal transcribed spacer (ITS-1) between 17S and 5.8S rDNA is 194-195 bp while that of the second internal transcribed spacer (ITS-2) between 5.8S and 25S rDNA is 233 bp.

Fukui et al. (81) have shown by in situ hybridisation, that rDNA repeats in rice (a cultivar of Japonica type) are present on chromosome 10. In addition, McCouch et al. (28) have shown rDNA repeats to be located on chromosome 9 by trisomic dosage analysis. Studies by Sano et al. (6) have revealed presence of rDNA (in 2 accessions C 8669 & 221) on two unlinked loci suggesting that number of chromosomes carrying NOR regions might differ among rice strains.

Putative initiation and termination sites for rDNA transcription in rice IR-20 have been reported (82,83). Three direct repeats of 320, 340 and 238 nucleotides have been detected upstream to the 5' end of 18S rRNA gene of rDNA unit present on 9.8 kbp EcoRI

fragment of rice DNA. The primer extension analysis has shown that the site of initiation of transcription is in the first repeat at an adenine residue, 623rd nucleotide upstream to the 5' end of 18S rRNA gene (82). The promoters for the rDNA transcription are located within +20 to -150 nucleotides, +1 being the start point of the precursor rRNA (84). These authors also report presence of multiple spacer promoters for initiation of transcription. The spacer promoters may serve as a sink for transcription factors and thus enhance transcription of rDNAs. The termination of rDNA transcription in plants has been shown to be very near to the 25S rDNA gene (85,86). Nandabalan et al. (83) have reported that there are two transcription termination sites for the precursor rRNA at 616 and 620 nucleotides downstream to the end of 25S rRNA gene in the rice IR20.

An extensive survey of rDNA spacer length variation (or RFLP of the rDNA spacer) is made in the genus Oryza (including 47 wild and 58 cultivated species) using a cloned rice rDNA probe (87). The overall observations are as follows:(1) The accessions corresponding to the sativa complex (cultivated varieties) show seven classes for the rDNA spacer. Each class differs from the next by an increment of 300 bp. Their regular increase in rDNA spacer size by units of ap-

proximately 300 bp strongly indicates that subrepeats similar to those described in other species exist in rice and are responsible for size variation. This has been recently confirmed by restriction mapping and sequence analysis of subclones of RR217 (88,6). Studies on the sub-species of the sativa complex indicate that japonica samples have spacer fragment falling preferentially into classes 1-3 and that typical indica have longer spacers, ranging from classes 4-8 (87). The wild species of the sativa complex (mainly 2 major species O. rufipogon and O. longistaminata) show greater spacer length variability than cultivated O. sativa, with as many as 13 different sizes of external spacer. Classification of japonica or indica based on rDNA pattern generally agrees with classification on the basis of isozyme patterns. In contrast, the glaberrima complex does not display any rDNA size variation, and exhibits a single class of spacer fragment. Extensive variation in rDNA spacer is observed in wild species indicating that variation in spacer length among all rice species cannot be solely due to variation in the number of small subrepeats. Based on these data, it is felt that rDNA spacer length polymorphism can be used as a reliable tool in analysing the evolution of rice species.

Spacer-length (sl) variation in the O. sativa-O.rufipogon and O. glaberrima-O. barthi com

plexes has been studied by Sano et al. (6). Their studies have revealed that the former complex is highly polymorphic for rDNA repeats and seems to share various sl variants of the repeat. The latter complex, however, is monomorphic for rDNA sl variation and carries only one sl variant. Their studies depict that different taxa carry different types of rDNA repeat in terms of IGS length and the subrepeat structure and that cultivated species are not clearly distinct from their progenitors with regard to rDNA variation.

REFERENCES

1. Chang, T.T. (1984)
Science, 224:251-256.
2. Oka, H.I. (1974)
Genetics, 78:475-486.
3. Chang, T.T. (1976)
Euphytica, 25:425-445.
4. Tateoka, T. (1963)
Bot. Mag., 76:165-173.
5. Sano, Y. (1986)
Sterility barriers between Oryza sativa and
O. glaberrima. In Rice Genetics. Edited by
the International Rice Research Institute,
Los Banos. pp 109-118.
6. Sano, Y. and Sano, R. (1990)
Genome, 33:209-218.
7. Oka, H.I. (1988)
Origin of cultivated rice. Developments in Crop
Science, 14. Japan Scientific Societies Press,
Tokyo.
8. Tsunoda, S. and Takahashi, N. (1984)
Biology of Rice. Developments in Crop Science, 7.
Japan Scientific Societies Press, Tokyo.
9. Second, G. (1982)
Jpn. J. Genet., 57:257.

10. Second, G. (1985)
Genet. Sel. Evol., 17:89-114.
11. Glaszmann, J.C. (1987)
Theor. Appl. Genet., 74:21-30.
12. Glaszmann, J.C. (1988)
Genome, 30:782-792.
13. Khush, G.S.(1991).
Variability of germplasm in rice. In : RFLP in
plant breeding , at ICGEB, New Delhi, Feb. 4-22
pp 50-55.
14. Nayar, N.M. (1973)
Origin and Cytogenetics of Rice. In: Advances in
Genetics. pp 153-292. E.W. Caspary (Ed.),
Academic Press, London.
15. Sitch, L.A. (1990)
Incompatibility barriers operating in crosses of
Oryza sativa with related species and genera. pp
77-93. In: Gene Manipulation and Plant Improvement.
II. J.P. Gustafson (Ed.) Plenum Press, New York.
16. Sitch, L.A., Brar, D.S., Guiterdoni, E., Hibino,
H., Hilleris-Lambers, D., Jena, K.K., Khush, G.S.,
Leung, H., Mew, T.W., Saxena, R.G. and Zapata,
F.J. (1988)
Germplasm enhancement - Wide hybridization. IRRI,
International Programme Review, IRRI, Philippines.

17. Jena, K.K. and Khush, G.S. (1989)
Genome, 32:449-455.
18. Heinrichs, E.A., Medrano, F.G. and Rapusas, H.R.
(1985)
Genetic evaluation for insect resistance in rice.
International Rice Research Institute, P.O.Box
933, Manila, Philippines.
19. Swaminathan, M.S. (1986)
Plant Mol. Biol. Rep., 4:1-17.
20. Bennett, M.D. and Smith, J.B. (1976)
Philos. Trans. Roy. Soc. Lon. B 274:227-274.
21. Bennett, M.D., Smith, J.B. and Heslop-Harrison,
J.S. (1982)
Proc. R. Soc. Lon., B216:179-199.
22. Flavell, R. (1980)
Annu. Rev. Plant Physiol., 31:569-596.
23. Flavell, R.B. (1982)
In Genome evolution, G.A. Drover and R.B.Flavell (Eds.)
Academic Press, London, 301-323.
24. Iyengar, G.A.S. and Ken, S.K. (1978)
Environ. Exp. Bot., 18:219-224.
25. Evans, I.J., James, A.M. and Barnes, S.R. (1983)
J. Mol. Biol., 170:803-826.
26. Flavell, R.B. (1986)
Phil. Trans. Roy. Soc. Lon., B312:227-247.

27. Deshpande, V.G. and Ranjekar, P.K. (1980)
Hoppe-Seylers Z. Physiol. Chem., 361:1223-1233.
28. McCouch, S.R., Kochert, G., Yu, Z.H., Wang, Z.Y.,
Khush, G.S., Koffman, W.R. and Tanksley, S.D.
(1988)
Theor. Appl. Genet., 76:815-829.
29. Thompson, W.F. and Murray, M.G. (1981)
The nuclear genome: Structure and Function, in A.
Marcus (Ed.), The Biochemistry of Plants: A
comprehensive treatise, Vol. 6, Proteins and
Nucleic Acids, Academic Press, pp 1-82.
30. Walbot, V. and Goldberg, R. (1979)
Plant genome organization and its relationship to
classical plant genetics, In: T.C. Hall and J.W.
Davies (Eds.), Nucleic Acids in Plants, Vol. I,
CRC Press, Florida, 1979, pp 3-41.
31. Davidson, E.H., Hough, B.R., Amenson, C.S. and
Britten, R.J. (1973)
J. Mol. Biol., 77:1-23.
32. Crain, W.R., Eden, F.C., Pearson, W.R., Davidson,
E.H. and Britten, R.J. (1976)
Chromosoma (Berl.), 56:309-326.
33. Gupta, V.S., Gadre, S.R. and Ranjekar, P.K. (1981)
Biochim. Biophys. Acta., 656:147-154.

34. Gupta, V.S. and Ranjekar, P.K. (1981)
J. Biosci., 3:417-430.
35. Bhawe, M.R., Gupta, V.S. and Ranjekar, P.K. (1986)
Plant Syst. Evol., 152:133-151.
36. Sivaraman, L., Gupta, V.S. and Ranjekar, P. K.
(1986)
Plant Mol. Biol., 6:375-388.
37. Lagu, M., Ranjekar, P.K. and Pillay, D.T.N. (1986)
Cell Biol. Int. Rep., 10(11):869-874.
38. Dhar, M.S., Dabak, M.M., Gupta, V.S. and Ranjekar,
P.K. (1988)
Plant Sci., 55:43-52.
39. Hiroka, M. (1984)
Wild plants and domestication. In: Tsunoda, S.,
Takahashi, N. (Eds.) Biology of rice. Jap. Sci.
Press, Tokyo, pp 3-30.
40. Vaughan, D.A. (1989)
The genus Oryza L., Current status of taxonomy.
IRRI Research Paper series, No. 138, p. 21.
41. Khush, G.S., Ling, K.C., Aquino, R.C. and Aguiro,
V.M. (1977)
Breeding for resistance to grassy stunt virus in
rice. Proc. Int. Cong., Society for Breeding
Researches in Asia and Oceania (SABRAO), 3rd,
Canberra, Australia No. 1-4, 6:3-9.

42. DeKochko, A. (1987)
Theor. Appl. Genet., 73:675-682.
43. Metzlauff, M., Troebner, W., Baldauf, F., Schlegel, R. and Cullum, J. (1986)
Theor. Appl. Genet., 72:207-210.
44. Sonina, N.V., Lushnikova, A.A., Tihanov, A.P. and Ananiev, E.V. (1989)
Theor. Appl. Genet., 78:589-593.
45. Hueros, G., Monte, J.V. and Ferrer, E. (1990)
Theor. Appl. Genet., 80:24-32.
46. Schmidt, T., Junghans, H. and Metzlauff, M. (1990)
Theor. Appl. Genet., 79:177-181.
47. Lavesque, H., Mathieu, C. and A.G.L. de Courcel (1990)
Theor. Appl. Genet., 80:602-608.
48. Rayburn, A.L. and Gill, B.S. (1987)
Amer. J. Bot., 74:574-580.
49. McIntyre, C.L., Clarke, B.C. and Appels, R. (1988)
Plant Syst. Evol., 160:39-59.
50. Appels, R., Reddy, P., McIntyre, C.L., Moran, L.B., Frankel, O.H. and Clarke, B.C. (1989)
Genome, 31:122-133.
51. McIntyre, C.L., Pereira, S., Moran, L.B. and Appels, R. (1990)
Genome, 33:635-640.

52. Fabijanski, S., Fedak, G., Armstrong, K. and Altoosar, I. (1990)
Theor. Appl. Genet., 79:1-7.
53. Pental, D. and Barnes, S.R. (1985)
Theor. Appl. Genet., 70:185-191.
54. Zhao, X., Wu, T., Xie, Y. and Wu, R. (1989)
Theor. Appl. Genet., 77:353-359.
55. Wu, T. and Wu, R. (1990)
Repetitive DNA sequence in CC genome types in the genus Oryza. Personal communication.
56. Toenniessen, G.H. (1990)
Rice Biotechnology: Progress and Prospects. SCI Conference on Pest Management in Rice, London, June 4-7
57. Wu, T. (1990)
Identification and characterization of repetitive DNA sequences in rice. An "A- Exam" Proposal.
58. Reddy, A.S., Cordesse, F., Kiefer, M.C. and Delseny, M. (1990)
Abstr. Fourth Annual Meeting of the Rockefeller Foundation's International Program on Rice Biotechnology, May 9-12, 1990.
59. Aswidinnoor, H., Dallas, J.F., McIntyre, C.L. and Gustafson, J.P. (1990)

- Abstr. Second International Rice Genetics Symposium, May 14-18, Manila, Philippines.
60. Aswidinnoor, H. (1991)
Genome specific repetitive DNA probes for monitoring introgression in rice and its potentials for plant breeding'. A Dissertation in partial fulfillment of Ph.D., Aug. 1991.
61. Wu, T. and Wu, R. (1987)
Nucl. Acids Res., 15:5913-5923.
62. Erdmann, V.A. (1982)
Nucl. Acids Res., 10:r93-r115.
63. Daniels, G. R., Fox, G. M., Loewensteiner, D., Schmid, C.W. and Deininger, L. (1983)
Nucl. Acids Res., 11: 7579-7593.
64. Daniels, G.R. and Deininger, P.L. (1985)
Nucl. Acids Res., 11:7595-7610.
65. Gabriella, G., Hofstetter, H. and Brinstiel, M.L. (1981)
Nature, 294:626-631.
66. Vedel, F. and Delseny, M. (1987)
Plant Physiol. Biochem., 25:191-210.
67. Wessler, S.R. (1990)
The Plant Cell, 2:685-686.
68. Mitrick, A.J., Babcock, M.S., Fuerstenberg, S.M., Fuerstenberg, S.I., Freeling, M. and Simpson, R.B.

- (1989)
Plant Mol. Biol., 12:633-642.
69. McClintock, B. (1948)
Mutable loci in maize. Carnegie Inst. Washington
Yearbook, 47:155-169.
70. Wu, R. and Xie, Y. (1990)
Transposable elements in rice. Abstr. Second
International Rice Genetics Symposium, May 14-18,
at IRRI, Manila, Philippines.
71. Kikuchi, S., Takaiwa, F. and Oono, K. (1987)
Theor. Appl. Genet., 210:373-380.
72. Zheng, K.L., Castiglione, S., Biasini, M.G.,
Biroli, A., Morandi, C. and Sala, F. (1987)
Theor. Appl. Genet., 74:65-70.
73. Gupta, V.S., Dhar, M.S., Patil, B.G., Narvekar,
G.S., Rawat, S.R. and Ranjekar, P.K. (1990)
J. Biosci., 15:261-269.
74. Dhar, M.S., Pethe, V.V., Gupta, V.S. and Ranjekar,
P.K. (1990)
Theor. Appl. Genet., 80:402-408.
75. Oono, K. and Sugiura, M. (1980)
Chromosoma, 76:85-89.
76. Olmedilla, A., Delcasso, D. and Delsemy, M. (1984)
Plant Sci. Lett., 37:123-127.

77. Reddy, P.S. and Padayatty, J.D. (1987)
Ind. J. Biochem. Biophys., 24:293-295.
78. Takaiwa, F., Oono, K. and Sugiura, M. (1984)
Nucl. Acids Res., 12:5441-5447.
79. Takaiwa, F., Oono, K. and Sugiura, M. (1985)
Plant Mol. Biol., 4:355-364.
80. Takaiwa, F., Oono, K., Iida, Y. and Sugiura, M.
(1985)
Gene, 37 :255-259.
81. Fukui, K., Kakeda, K., Hashimoto, J. and Matsuoka,
S. (1987)
Rice Genet. Newsl., 4:114-116.
82. Nandabalan, K. and Padayatty, J.D. (1989)
Biochem. Biophys. Res. Commu., 160:1117-1123.
83. Nandabalan, K. and Padayatty, J.D. (1988)
Biochem. Biophys. Res. Commu., 157:68-74.
84. Sollner-Webb, B. and Tower, J. (1986)
Ann. Rev. Biochem., 55:801-830.
85. Toloczkyi, C. and Feix, G. (1986)
Nucleic Acids Res., 14:4969-4985.
86. Delcasso-Tremousaygue, D., Grellet, F., Panabieres,
F., Ananiev, E.D. and Delseny, M. (1988)
Eur. J. Biochem., 172:767-776.
87. Cordesse, F., Second, G. and Delseny, M. (1990)
Theor. Appl. Genet., 79:81-88.
88. Cordesse, F. Unpublished results.

CHAPTER TWO

MATERIALS AND METHODS

2.1 MATERIALS AND CHEMICALS

2.1.1 Seed material, growth conditions and harvest of tissue

Seeds of rice O. sativa cv. Malkolam used for the construction of genomic library, were obtained from Agriculture Research Station, Palghar. Dist Thane, India.

Nineteen accessions of indica rice were used as germplasm for the survey of genetic variability at the molecular level. The names of these varieties with elite characters are given in Table 2.1. These seeds were obtained from Prof R.B. Shinde, College of Agriculture, Kolhapur, Maharashtra State, India.

The wild rice germplasm was procured from Dr. R.N. Mishra, Central Rice Research Institute, Cuttack, India. The accessions belonging to following genome types were used for the study 1) Basmati-370 (AA), 2) IR-64 (AA), 3) O. officinalis (CC), 4) O. eichingeri (CC), 5) O. alta (CCDD), 6) O. latifolia (CCDD), and 7) O. australiensis (EE). All rice seeds were grown in pots (containing 1:1 compost soil mixture) in the glass house at the College of Agriculture, Pune. Plants were watered for 6-8 weeks. Leaf material was harvested when plants were approximately two months old and was immediately frozen in liquid nitrogen and stored at -70°C till further use.

TABLE 2.1

Rice germplasm used for the assessment of genetic variability

Name of the variety	Pedigree	Grain Type	Special features or Remarks
1. RP4-14	T90/IR8	LS	High yielder with good grain quality, susceptible to BPH
2. Rasi	T(N)1/Co29	MS	High seedling vigour, resistant to BL, drought resistant, suitable for uplands
3. VDN 7850 (Kundalika)	RTN-24/IET 3228	SS	Fine variety, recommended for sub-mountain zone in Kharif season, good cooking quality
4. Karjat 184	T(N)1/Kolamba 540	LS	Recommended for uplands & hill slopes, variety with fine grain & good cooking quality
5. Jaya	T(N)1/T141	LB	Maturess a few days earlier than IR8. Has high degree of stability for yield
6. Radhanagari 185-2	Hawalsall7/T(N)1	LB	Early maturing, suitable for direct seeding under upland conditions
7. BG 380-2	-ND-	LS	Introduced into India during green revolution, widely adaptable
8. Jalgaon-5	Local selection from Batu	LS	Recommended for rainfed & irrigated uplands, plant parts pigmented, fine scent
9. VDN 3-51-18 (Indrayani)	IR8/Ambemohar 157	LS	Scented, Resistant to BL
10. Ratnagiri 73	RTN-23/KJT 87-2	LB	Resistant to BL, early maturing

Name of the variety	Pedigree	Grain type	Special features or Remarks
11. Ratna (CR 44-11)	TKM6/IR8	LS	good grain quality, wide adaptability
12. VDN 4288 (Pawana)	Pusa 33/ IR28	LB	Scented, moderate yielder Kharif crop
13. Ratnagiri 24(RTN 24)	Ziniya 63/ T(N)1	LS	Short duration variety, with good grain quality, suitable for uplands, susceptible to BL & BB
14. Ratnagiri 711	IR8/ Ratnagiri 24	LS	Short duration variety with good grain quality, suitable for uplands
15. Ratnagiri 1	IR8/RTN 24	LB	Resistant to BL & tolera- nt to GM
16. Palghar 39-4-1	-ND-	LB	ND
17. Pusa-33	Sabarmati/ Ratna	LS	Early maturing variety, Comparable to Basmati- 370 as regards to amylos- e content
18. ACK 5	Dodga 6-2-2/ IR8	SB	Recommended for direct seeding, tolerant to drought & iron chlorosis
19. Karjat 35-3	HR19/IR8	MS	Early maturing, fine with good cooking quality

ND - No Data available
 LS - Long Slender
 MS - Medium Slender
 SS - Short Slender
 LB - Long Bold
 SB - Short Bold

BPH - Brown Plant Hopper
 BL - Blast
 BB - Bacterial Blight
 GM - Gall Midge

2.1.2 Chemicals, materials and enzymes

All routine chemicals used throughout the work were of Analytical Reagent (AR) or Guaranteed Reagent (GR) grade and were obtained from Qualigens, British Drug House (BDH), Sarabhai chemicals or E. Merck, India.

Solvents such as chloroform, ethanol (Swastik Laboratory Pvt. Ltd., Pune) and Phenol (SD's Laboratory Chemical Industry, Bombay) were freshly distilled prior to use. Fine chemicals including Tris [tris (hydroxymethyl amino methane)], sarkosyl, SDS [Sodium dodecyl sulphate], agarose, ampicillin and Bovine Serum Albumin (free from DNase activity) were obtained from Sigma Chemical Co., USA. RNase A (free of DNase activity and X-gal (5-bromo-4-chloro-3-indolyl β -D-galactoside) were also from Sigma Chemical Co., USA. Molecular weight markers (Lambda phage DNA digested with Hind III) or lambda DNA (one kb ladder) were from Bethesda Research Laboratories, USA.

All the restriction endonucleases used for digestions of DNAs were from Amersham, U.K; Boehringer Mannheim (W. Germany); New England Biolabs (NEB), USA or Bethesda Research Laboratories (BRL), USA.

Transfer membranes used for Southern blotting of DNA were either Hybond-C or Hybond N (Amersham, U.K,) or Gene Screen Plus (NEN Research Products).

Radio-labelled alpha P-32-dCTP and labelling kits for preparation of probe were obtained from Radio chemical centre, Amersham, U.K.

Black and white photographic film ORWO (made in German Democratic Republic) (135/36 mm, 125 ASA) and X-ray films : Fuji (Fuji Photo film Co Ltd, Japan) were used for gel photography and autoradiography respectively.

Commercially available developer (IPC 163) and fixer were used for developing the X-ray films (India Photographic Company Ltd., Bombay).

2.2 Isolation of DNA and criteria of purity

2.2.1 Isolation of genomic DNA

Rice DNA from leaf tissue, shoot (8-10 days old seedling) and embryo was extracted by using the protocol of Shure et al.(1).

The frozen tissue was ground to a fine powder with 0.5 mm glass beads using liquid nitrogen in a mortar. The powder was then suspended in urea-phenol extraction buffer (8.0M urea, 0.35 M NaCl, 0.05 M Tris-HCL , pH 7.5, and 0.02 M EDTA). 2% Sarkosyl and 5% Tris saturated phenol were freshly added to the buffer. The

suspension was incubated at 65°C for 15-20 minutes and was then extracted with equal volume of 3:1 phenol chloroform mixture, (chloroform with 5% isoamyl alcohol) to remove proteins and pigments. The upper aqueous phase containing nucleic acids was collected after centrifugation at 10,000 rpm in Sorvall RC 5B centrifuge (rotor : SS34) for 10 minutes at 4°C and was dephenylated by extracting twice with chloroform-isoamyl alcohol (24:1 v/v) mixture. 1/10th volume of 3M Na acetate was added and then nucleic acids were precipitated from the aqueous layer with two volumes of chilled ethanol. The pellet was then washed with 70% ethanol, air dried and suspended in TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0). RNA contaminants were removed by addition of RNase A (made DNAase free by boiling for 10 minutes) to the final concentration of 100 ug/ml, and incubating the mixture for 1 hour at 37°C. RNAase reaction was terminated by deproteinization as described above and DNA was reprecipitated from the aqueous layer by adding two volumes of chilled ethanol. DNA was immediately spooled with a glass rod, washed twice with 70% alcohol, air dried and dissolved in TE (Tris-HCl 10 mM, EDTA 1 mM pH 8.0). The yield of DNA was about 2 mg from 10 gm of fresh weight leaf tissue.

All DNA preparations were checked for protein and polysaccharide contamination spectrophotometrically. DNA was scanned in the range of 220-320 nm on a Shimadzu double beam spectrophotometer, model UV 210-A. Only those preparations which exhibited an optical density ratio $A_{280}/A_{260} = 0.45-0.55$ and $A_{230}/A_{260} = 0.45$ were used (2). The molecular weights of all the DNAs were more than 23 kbp as determined from agarose gel electrophoresis.

2.2.2 Isolation of plasmid DNA

Several methods have been reported to isolate plasmid DNA from small volume cultures inoculated with single colony of bacteria, eg. alkaline lysis method (3,4), modified boiling Triton lysis method (5,6) and a two step miniprep procedure recently described by Serghini et al. (7).

All plasmid DNA extractions were done by the alkaline lysis protocol described by Birnboim et al., 1979 (3). A single recombinant colony from the master plate was inoculated in 1ml LB medium containing ampicillin (50 ug/ml). This was grown to saturation, at 37°C, overnight with shaking at 250 rpm. The culture was spun in 1.5 ml eppendorf for 1 minute to pellet cells and supernatant was discarded. The cell pellet was suspended in 100 ul solution I [GTE buffer : 50 mM

Glucose, 25 mM Tris-HCl, (pH 8.0), and 10 mM EDTA (pH 8.0)], vortexed and rested at room temperature for 5 minutes. Then 200 ul of freshly prepared solution II [1% SDS, 0.2N NaOH] was added to the above cell suspension, mixed well by tapping the tube and rested on ice for 5 minutes. SDS essentially denatures the bacterial proteins and chromosomal DNA, while NaOH denatures the plasmid DNA. The suspension was neutralized by adding 150 ul potassium acetate buffer [To 60 ml of 5 M potassium acetate, add 11.5 ml glacial acetic acid and 28.5 ml water] and it was again rested on ice for 10 minutes, to enable the plasmid DNA to reanneal rapidly. The suspension was microfuged for 5 minutes to pellet out the cell debris and chromosomal DNA. Supernatant containing the plasmid DNA was transferred to a fresh tube to isolate plasmid DNA as described earlier by ethanol precipitation.

This procedure can be easily adapted to accommodate bacterial cultures that range in size from 1 ml to 1 litre. It can be used successfully to isolate plasmid DNA simultaneously from many small cultures (minipreps) as well as from large scale cultures (maxipreps) and cultures of intermediate size (midi preps) (2).

2.3 Restriction endonuclease digestions of DNA

In a typical digestion experiment, 1-2 ug of DNA was incubated with a suitable restriction enzyme (8-10 U/ug DNA) in a reaction volume of 35-40 ul at 37°C. The salt concentrations of buffers were according to supplier's recommendations. The enzyme to DNA and enzyme volume to reaction volume ratios were carefully adjusted to avoid non-specific (star) activity of restriction enzymes (2). Following incubation, the enzyme digestion was stopped by addition of 10 X reaction terminating buffer (50% glycerol, 100 mM EDTA, 0.25% bromophenol blue) to a final concentration of 1 X. Control experiments were performed using commercial lambda DNA digested with different restriction enzymes to check the reaction conditions.

For Southern analysis, 5-7 ug of genomic DNA was incubated with 50-70 U of suitable restriction enzyme, in a volume of 40 ul at 37°C using assay conditions as recommended by supplier. Double digestions were carried out for mapping the restriction endonuclease sites on the target DNA. For this purpose, plasmid DNA was first digested with an enzyme requiring low salt assay buffer, followed by incubation of plasmid with high salt buffer enzyme. Digestions were allowed to proceed at 37°C. The reaction was terminated in the same way using the 10 X termination buffer as

described above.

2.4 Agarose gel electrophoresis

The DNA digests were analysed on 1% neutral agarose horizontal slab gels in TAE 1X [Tris Acetate (40mM), pH 8.0, Sodium acetate (5mM) and EDTA (2mM), pH 8.0] at a constant current of 20 mA, for 10-12 hrs in case of genomic DNAs and 3-5 hrs for plasmid DNAs. Alternatively, TBE buffer was also used [Tris (0.089 M), Boric acid (0.089 M) and EDTA (0.002 M)].

After gel electrophoresis, the gels were stained in dark in ethidium bromide (100ng/ml) and were visualized on a long wavelength (302 nm) UV transilluminator (UV Products, San Gabriel, California, USA.) and photographed with a 35mm SLR camera (Minolta X 700 with macrophotography and zoom lens system) using a red filter on 135/36 B/W ORWO 125 ASA film.

2.5 Pre-requirements for cloning

2.5.1 Vector and host system used for the construction of partial genomic library

The plasmid vector pUC18 and *E. coli* host DH5 α were used for the construction of the partial rice genomic library. The details regarding the genotypes, selection markers, size and restriction enzyme sites are given below.

pUC 18 (2.695 kbp)

The pUC plasmids (eg. pUC 18, pUC 19, pUC 118, pUC 119) are extremely versatile vectors consisting of an array of unique cloning sites, a modified ampicillin resistance gene and an origin of replication derived from pBR322 ligated to a portion of the lac Z gene of E. coli (8,9,10). The pUC plasmids are small and can be easily introduced into E. coli cells by transformation, even when they contain large DNA inserts. Restriction mapping of recombinant molecules is facilitated by the small size of the vector. Being high copy number plasmids, yields of DNA preparations are high. Moreover, the screening of the recombinants is easy, wherein the non-recombinants produce blue colonies and recombinants produce white colonies on ampicillin plates containing X-gal or Blueo-gal (5-Bromo-4-chloro-3-indolyl- β -D-galactoside).

E. coli DH5 α : (11,12)

E.coli DH5 α is a recombination-deficient suppressing strain used for plating and growth of plasmids and cosmids. The 80 Δ lacZ Δ M15 permits alpha complementation with the amino-terminus of β -gal encoded in pUC vectors (2).

2.5.2 Media preparation

Media were prepared in glass distilled water and were autoclaved at 15 psi (pounds per square inch) for 20 minutes. Heat labile components like antibiotics were filter sterilized through autoclaved milipore HAWP membranes and were added freshly to the cooled media (45°C).

For the preparation of solid media, agar was included before autoclaving. When the medium was removed from the autoclave, it was swirled gently to distribute the melted agar evenly through the solution. The medium was then cooled to 50°C, before adding thermolabile substances (eg. antibiotics) (2). The concentration of the agar used was 1.5% throughout.

Composition of the medium and concentration of antibiotics used are given below:

LB medium (Luria-Bertani medium)

Per liter

Tryptone	: 10 gms
Sodium chloride	: 5 gms
Yeast extract	: 5 gms
pH	: 7.2-7.5, adjusted with 1N NaOH

Plating agar for plates

Per litre

Tryptone	: 10 gms
Sodium chloride	: 5 gms
Yeast extract	: 5 gms
Agar (bacto-agar)	: 15 gms
pH	: 7.2 - 7.5, adjusted with 1N NaOH

Antibiotics

Ampicillin trihydrate was dissolved in a stock concentration of 100 mg/ml in water. Effective concentration of 100 µg/ml was used for both liquid and solid media.

2.5.3 Growth, culturing & storing of the host

E.coli DH5 alpha

E. coli DH5α cells were inoculated from a solid medium (single colony on a plate) into 10 ml liquid LB medium and were allowed to grow at 37°C overnight with shaking (150-200 rpm). This freshly prepared culture was used each time for further inoculation whenever necessary.

Cells were stored in stab vials at 0-4°C or as frozen glycerol suspension (15%) at -70°C. Stabs were subcultured every month, while glycerol cultures were stored for at least 6 months.

2.6 Construction of a partial genomic library of rice

A partial genomic library of rice O. sativa cv. Malkolam was prepared in the plasmid vector pUC18. The important steps in this procedure were as follows:

2.6.1 Preparation of vector and target DNA

Commercially prepared PstI restricted pUC18 vector DNA (Appligene 50 ng/ul) was used for ligation with the target DNA. For the preparation of target DNA, total rice DNA (cv. Malkolam) was digested to completion with PstI (Appligene 10 U/ul) as described by Young et al. (13). Digested DNA was size fractionated on horizontal gels which were run in TBE 1X buffer [Tris (0.089 M), Boric acid (0.089 M) and EDTA (0.002 M)]. Size selected DNA (200 bp - 3 kbp) was eluted from agarose gel onto Whatman DE81 paper (2.3 cm, Whatman Ltd, Maidstone, England) (14). DNA was released from the paper by incubating the tube with high salt elution buffer (1.5 M NaCl in TE, pH 8.0) at 42°C for 15 minutes. This step is repeated twice for good recovery of DNA. DNA from the aqueous phase was precipitated using two volumes of 95% ethanol, dissolved in TE (Tris.HCl 10 mM, pH 8.0, and EDTA 1mM, pH 8.0) and quantified by band-intensity on ethidium bromide stained gels using a

defined amount of standard DNA (pCA8/ HindII fragments). The size fractionated target DNA was used for ligation with PstI restricted vector DNA.

2.6.2 Ligation of target DNA with pUC18 DNA to generate chimeric plasmids

PstI digested, and size selected (0.2 kbp-3 kbp) rice DNA (300 ng) was mixed with PstI linearized pUC18 DNA (100 ng) (Appligene, 50 ng/ μ l) in a reaction volume of 50 μ l containing 5 μ l 10 X bacteriophage T₄ DNA ligase buffer (500 mM Tris- HCl, pH 7.4, 100 mM DTT, 10 mM spermidine, 1mg/ml BSA), 5 μ l 0.1 M ATP, 5 μ l 0.1 M MgCl₂ and 2.5U T₄ DNA ligase (Amersham 50 U/20 μ l). Ligation was carried out at 4°C for 16 hours. This ligation product was directly used to transform E. coli DH5 α competent cells.

2.6.3 Preparation of competent cells and transformation of E. coli DH5 alpha with chimeric plasmid

Preparation of highly 'competent' host cells is a key step towards a good efficiency of transformation. Bacterial transformation methods are based on the observations of Mandel and Higa (15) who showed that bacteria treated with ice-cold solutions of CaCl₂ and then briefly heated could be used to transfect with

phage DNA . In 1973, this method was also shown to work for plasmid DNA by Cohen, Chang and Hsu (16). Apparently, the treatment includes a transient state of "competence" in the recipient bacteria, during which they are able to take up DNAs from a variety of sources.

Although different methods are now available for bacterial transformations, the procedure developed by Hanahan (12) is, however, routinely used . This procedure can yield competent cultures of E. coli strains DH1, DH5, DH5 alpha and MM 294 that can be transformed at frequencies more than 5×10^8 transformed colonies per microgram of plasmid DNA. Transformation of E. coli was carried out with 50-100 ng of ligated DNA. Along with the ligated samples, 25 ng of uncut plasmid DNA (pUC18) and host DH5 alpha cells were used as positive and negative controls respectively. The transformation procedure was essentially as described by Hanahan 1983 (12). A single colony of E. coli DH5 alpha was inoculated in 10 ml LB and was grown over night at 37°C with shaking (250 rpm). One ml of this saturated overnight grown culture was inoculated in 50 ml LB and incubated at 37 C (approximately 1.5 hrs) till the culture obtained an OD of 0.45-0.55 (cell density of 5×10^7 cells/ml) at 550 nm. Subsequent

ly, the culture was immediately chilled on ice for 20 minutes to stop further cell growth. Cells were then pelleted by centrifugation at 3000 rpm for 7 minutes at 4°C in a Sorvall RC 5B centrifuge. The supernatant was poured off and the cell pellet was gently suspended in 25ml ice-cold 100 mM CaCl₂ sterile solution and rested on ice. Exposure to calcium ions enables the cells to take up DNA or to become competent. Cell suspension was again centrifuged for 5 minutes at 2,500 rpm, 4°C. The supernatant was discarded and cell pellet then suspended in 2 ml ice-cold 100 mM CaCl₂ solution. 200 ul of competent cells were then dispensed into prechilled sterile eppendorf tubes. The ligated and control plasmid DNAs were added to the cell aliquots and rested on ice for 40 minutes. Cells were then heat shocked at 42°C for 2.5 minutes. Alternatively, cells were incubated at 37°C for 5 minutes. The heat shock treatment enlarges the cell pore size, enabling the DNA in the suspension to enter the cells. Immediately after this step, 800 ul of sterile LB medium was added to each tube, and tubes were placed on a shaker at 250 rpm for 1 hour at 37°C. This helps the host cells to recover and express antibiotic resistance. To this culture, 120 ul of X-gal (20 mg/ml) was added, the solution mixed well and 100 ul aliquots of this transformation culture were spread on hard Luria agar plates (Agar 1.5%) containing ampicillin (100 ug/ml) (2). When the

plates were dry, they were incubated at 37°C for 12-16 hours.

2.6.4 Selection of recombinants, saving the library, preparation of library master plates and transfer and replication of colonies onto filters

Selection

Colonies of E. coli cells harboring plasmids with rice inserts (chimeric plasmids) were selected on the basis of X-gal-ampicillin screening procedures. White colonies generally represent the recombinants, while blue are the non-recombinants.

Saving the library

Individual white colony (clone) was inoculated with sterile tooth picks into 120 ul LB (containing 100 ug/ml ampicillin) in 96 wells Elisa microtitre plate (96 'U' shaped cups). and incubated with shaking at 37°C, for 16 hours. 15% sterile glycerol was added and the culture was again incubated at 37 C with gentle shaking for 2 hours. Six such plates, each having 96 recombinants (500 clones) were saved at -70°C till use.

Preparation of master plates

Master plates (LB agar containing 100 ug/ml ampicillin) were prepared from the original library plates (stored at -70°C) using microtitre well

plate combs (96 teeth). These replica plates after inoculation of colonies were incubated at 37°C, for 16 hours. Master plates were stored at 4°C.

Preparation of colony filters

Next step was to make colony filters for screening the library by colony hybridization using total DNA as probe. For this purpose, Genescreen filters were smoothly placed on the LB agar plates (containing 100 ug/ml ampicillin) avoiding any air bubble between the membrane and the medium. Cultures were inoculated on the membrane in the same way as done for the master plates. These plates with the membrane were incubated at 37°C for 16 hours for growth of individual colonies. Next day, individual colonies appeared on the filter. Filters were further treated and used for screening the library by colony hybridization.

2.7 Screening the library for low copy/repetitive clones by colony hybridization

In 1975, Grunstein and Hogness (17) described a method for in situ lysis of bacterial colonies on nitrocellulose filters and non-covalent attachment of released DNA to the filters. The DNA could then be hybridized to appropriately labelled nucleic

acid probes. This method remains the most commonly used technique to date to identify individual bacterial colonies carrying plasmids or cosmids that contain sequences of interest and screen the library for different types of sequences. Various solid supports like nitrocellulose filters, nylon and Whatman filters can be used for in situ hybridization of lysed bacterial colonies. To liberate the plasmid DNA, recombinant colonies on the filter were lysed essentially based on original procedure of Grunstein and Hogness (17). 4 pieces of Whatman paper (3 MM) were cut and placed on the bottom of 4 pyrex baking dishes (20 cm x 20 cm). The first step involved the lysis of bacterial colonies, wherein, one piece of paper was saturated with 10% SDS and excess liquid was poured off. Genescreen filter having overnight grown recombinants was peeled off from the plate using blunt-ended forceps (eg Millipore forceps) and was placed on the SDS impregnated Whatman paper with colony side up for 3 minutes. This treatment limits the diffusion of the plasmid DNA during denaturation and neutralization resulting in a sharper hybridization signal. Filter was then transferred to the second Whatman sheet, saturated with the denaturing solution (0.5 N NaOH, 1.5 M NaCl), for 5 minutes. This brings about in situ denaturation of released plasmid DNA. It was later transferred to the third sheet of Whatman paper,

saturated with the neutralizing solution (1.5 M NaCl, 0.5 M Tris. Cl, pH 8.0), for 5 minutes. Filter was then rested on a dry sheet of Whatman paper with colony side up. It was allowed to dry at room temperature for 30-60 minutes. The filters were sandwiched between two sheets of dry Whatman 3 MM paper. Liberated, denatured DNA on the filter was fixed by baking it in vacuo at 80°C for 1-2 hours. These filters were used for hybridization with P-32-labelled probes for screening the library.

2.8 Southern blotting

The transfer of DNA from agarose gels to Hybond N/Hybond C (Amersham) membranes or Nytran filters is a routine procedure and may be carried out by different methods like capillary transfer, electrophoresis or vacuum transfer (2). In 1975, Southern (18) introduced a blotting technique for the transfer of electrophoretically separated DNA to a membrane support. This method was subsequently extended to RNA analysis (19). Specific nucleic acid sequence immobilized on the membrane can be detected by hybridization of filter bound DNA or RNA with a radioactively labelled probe. These blot hybridization procedures have found very wide applications in gene mapping, transcription analysis, and pre-natal investigations of

human genetic disorders.

The transfer of DNA was carried out essentially according to the capillary transfer protocol of Southern (18) and Sambrook et al. (2). To begin with, the gels were soaked in 0.25 N HCl at room temperature for 10 minutes which results in depurination of DNA. The DNA was then denatured by soaking the gel in 1.5 M NaCl and 0.5 N NaOH for 30 minutes, twice, at room temperature with constant gentle shaking. Neutralization of the gel was done by soaking the gel in 1M Tris, pH 7.4 and 1.5 M NaCl, 30 minutes, twice, at room temperature with constant shaking. The transfer of DNA onto the membrane was carried out for 12-15 hours in 20 X SSC, after which the membranes were rinsed in 2 X SSC, air dried and baked for 2 hours in vacuo at 80°C (for nitrocellulose filters) and without vacuo for nylon membranes. The filters were stored at room temperature between sheets of Whatman 3 MM paper till further use.

2.9 Labelling the probe DNA and it's purification

Nucleic acid hybridization is a pivot step in molecular cloning and virtually every aspect of cloning, characterization and analysis of genes involves hybridization of one DNA strand to other. The labelled probe when hybridizes to unlabeled DNA strand

on the filter gives a signal on the autoradiogram, depending on sequence homology between two fragments. The efficiency of such hybridization experiments depends on the specific activity of the radio labelled probe. I have essentially used two procedures namely, Nick translation and Prime-a-Gene system (or Random primer method) for obtaining a radio-labelled probe (20,21,22).

The nick translation method described by Rigby et al. (20) is based on the ability of DNaseI to introduce randomly distributed nicks into DNA and the simultaneous ability of E. coli DNA polymerase I to synthesize DNA complementary to the intact strand in a 5'--->3' direction, using the 3'---OH termini of the nick as the primer. The 5'--->3' exonucleolytic activity of DNA polymerase I removes nucleotides in the direction of synthesis. The nick translation reaction mixture consisted of 1 ug DNA, 5mM MgCl₂ and 20 mM each of the unlabelled dATP, dGTP and dTTP. Commercial alpha P-32 dCTP (Amersham) (specific activity 3000 Ci/m Mole) was added to a final concentration of 90-100 picomoles. E. coli DNA polymerase I was added to a final concentration of 5U/ug. The reaction was carried out at 15°C for 3 hours and was terminated by incubating reaction mixture at 65°C, for 5 minutes.

Alternatively, the probes were labelled with alpha P-32 dCTP (specific activity 3000 Ci/m Mole (Amersham) by random priming method of Fienberg and Vogelstein (21,22) which is based on the hybridization of a mixture of all possible hexanucleotides to the DNA to be labelled. With this straight forward method, it is possible to generate probes of extremely high specific activity ($> 10^9$ cpm/ug DNA), even using DNA fragments directly isolated from agarose gels. The length of DNA fragments to be labelled does not influence the reaction. Since input DNA serves as a template and remains intact during the reaction, it is possible to label minimal amounts of DNA (10 ng) to a high specific activity.

In a typical reaction, 25-50 ng of the probe in a suitable volume of TE (Tris 10 mM, EDTA 1 mM, pH 8.0) was denatured by boiling for 10 minutes and was immediately chilled on ice. To the denatured DNA fragment, following reagents from Amersham multiprime kit were added in series : 5 ul primer, 5ul labelling buffer, 4 ul each of 10 X dATP, 10 X dGTP and 10 X dTTP, 5 ul (50 uCi) of alpha P-32 dCTP (specific activity 3000 Ci/m Mole) and 2 ul klenow fragment (4U), final volume was made to 50 ul with sterile water. Reaction was carried out at room temperature for 2-3 hours and terminated by adding 10 X stop dye

to a final concentration of 1 X [10% SDS, 125 mM EDTA, 10 mg/ml stock BPB (1ml/ml)].

To prevent background or non-specific hybridization in any experiment involving radioactivity, it is necessary to separate the probe from unincorporated labelled nucleotides. This is achieved by purification of the probe by spun column chromatography (2). To begin with, a 1.5 ml microfuge tube was packed with Sephadex G-50 to the level of 1 ml by centrifugation in a Remi table-top microfuge at 1000 rpm. The entire mixture containing labelled probe and labelled free nucleotides was pipetted on this exclusion column. 50 ul of 25mM EDTA/1% SDS was then added directly on the top of the column, and after waiting for 30 seconds, the column was spun at 1000 rpm for 20-25 seconds. After spinning, the probe eluted in a volume of about 350-400 ul was collected and used for the hybridization.

2.10 Prehybridization, hybridization, post-hybridization washing and autoradiography

To prevent the non-specific binding of the probe onto membranes, filters were prehybridized for 3-4 hours. Prehybridization was carried out at 62°C in heat sealed plastic bags containing the following mixture : 5 X SSC, 0.1% SDS, 5 X Denhardt's solution

(1% Ficoll, 1% Polyvinyl pyrrolidone, 1% BSA), 100 ug/ml sonicated salmon sperm DNA and 5 mM sodium pyrophosphate. After prehybridization, the solution was removed and replaced by hybridization solution. The latter had the same constituents of the prehybridization mixture except that it contained the probe. The purified probe (specific activity $1-2 \times 10^8$ cpm/ug) was denatured by boiling it for 10 minutes and immediately chilled on ice. Hybridization of the probe to the DNAs on the filter was continued overnight at 62°C with gentle shaking.

After overnight hybridization, probe was removed and filters were washed off from the unbound/unhybridized probe first with solution I (5 X SSC, 0.1% SDS) for 10 minutes at room temperature with gentle shaking. This was then replaced by fresh solution (5 X SSC, 0.1% SDS), prewarmed at 62°C. This washing was done twice, 15 minutes each, at 62°C, with gentle shaking. Washing was continued with two stringent washes, first with 2 X SSC, 0.1% SDS followed by 0.5 X SSC, 0.1% SDS. Moist filters were saran wrapped and exposed to XAR-5 X-ray film (Fuji) for 4 hours to 4 days, at -70°C using Cronex Lightening-Plus intensifier screens. After exposure, the films were developed with Agfa X-ray developer IPC-163, washed, dried and photographed with a 35mm SLR camera (Minolta X 700 with

macrophotography and zoom lens system) on Agfa copex film (25 ASA).

2.11 Copy number estimation

In order to quantify the copy number of a repeat sequence in the rice genome, different amounts of total rice DNA (10 ng, 100 ng and 1000 ng) and recombinant plasmid DNA were dot blotted onto nitrocellulose (Hybond C, Amersham) membrane. The DNA samples were denatured by boiling for 10 minutes prior to application on the membrane. Filters were prehybridized for 3-4 hours, and hybridized in the same solution as described earlier using P-32- labeled pOSM1C-2 insert as probe (0.1 ug, 10^8 cpm/ug). After overnight hybridization, filter was sequentially washed with decreasing concentrations of SSC (5X, 2X, 1X, and 0.1X) containing 0.1% SDS at 62°C, for 20 minutes each time, and exposed to X-ray film. Copy number was estimated by a laser densitometric scan of the autoradiogram (after drawing a line for self hybridization). Copy number of repeat sequence was calculated according to Rivin et al. (23).

2.12 RFLP analysis

In the present study, RFLP analysis was used to estimate the level of genomic variability among 19

cultivars of rice belonging to indica subspecies. RFLP probing was done using a dispersed repeat (pOSM1C-2) (1.55 kbp PstI fragment), identified and isolated from the rice partial genomic library. The logic behind using this particular clone as probe was to assess whether pOSM1C-2 exhibited a potential to decipher genetic/genomic variability among the selected germ-plasm. For this purpose, radio-labelled pOSM1C-2 insert was challenged against the digests of 19 cultivars with EcoRI, EcoRV, XbaI, DraI, ScaI, and HindIII and the autoradiograms generated for each of the enzyme probe combinations were analysed. Every DNA fragment detected by Southern analysis, when probed with pOSM1C-2 insert, was treated as an independent unit. All hybridizing fragments detected in a single variety were used to make comparisons among the cultivars. The Nei's F values for different combinations of accessions were calculated based on the number of shared fragments. To estimate F (proportion of fragments shared by 2 populations), the following formula proposed by Nei was used (24,25).

$$F \text{ (similarity index)} = 2n \text{ X Y} / (nX + nY)$$

where nX & nY are the numbers of fragments in populations X & Y respectively, where as nXY is the number of fragments shared by the two populations.

Using the above formula, the nature of polymorphism detected for 18 cultivars was presented in the form of a matrix. Although a large amount of information is contained in such a matrix, it is in a form which is difficult to visualize. For this reason a number of methods have been used to cluster taxa and produce dendrograms based on similarity and distance values (26). The three main methods used for analysing electrophoretic data are:

1. The unweighted pair group method with arithmetic mean (UPGMA) (27)
2. The phylogenetic tree method of Fitch and Margoliash (28).
3. The modified Wagner tree procedure of Farris (29).

I have used the first method, namely, UPGMA for the construction of dendrogram. Here, the clustering takes place beginning with the most similar pair of species and this pair is hence forth treated as a single entity. The data matrix is recalculated, and a species which exhibits the highest level of similarity to a cluster is admitted to the previously formed clusters. Recalculation is continued in a cyclic fashion until all the species are clustered. In UPGMA, the average similarity of a potential member to a cluster

is determined by weighing equally each species in the cluster. The resulting dendrogram is a useful visual summary of the relationships among species as group of species (26).

REFERENCES

1. Shure, M., Wessler, S., Federoff, N. (1983)
Cell, 35: 225-233.
2. Sambrook, J., Fritsch, E.F. and Maniatis, T.
(1989)
In: Molecular Cloning : A Laboratory Manual, 2nd
edition, Cold Spring Harbor Laboratory Press, New
York.
3. Birnboim, H.C. and Doly, J. (1979)
Nucleic Acids Res., 7: 1513-23.
4. Ish-Horowicz, D. and Burk J.F. (1981)
Nucleic Acids Res., 9: 2989-2998.
5. Holmes, D.S. & Quiley, M. (1981)
Anal. Biochem., 114: 193.
6. Haltiner, M., Kemple, T. and Tijan, R. (1985)
Nucleic Acids Res., 13: 1015-25.
7. Serghini, M.A, Ritzenthaler, C. and Pinck, L.
(1989)
Nucleic Acids Res., 17(9), 3604.
8. Yanisch-Perron, C., Vieira, J. and Messing, J.
(1985)
Gene, 33: 103.
9. Messing, J. (1983)
Meth. Enzymol., 101: 20.

10. Norrander, J., Kempe, T., and Messing, J. (1983)
Gene, 26: 101.
11. Messelsen, M. and Yuan R. (1968)
Nature, 217: 110.
12. Hanahan, D. (1983)
J. Mol. Biol., 166: 557-580.
13. Young, N.D., Miller, J.C., Tanksley, S.D. (1987)
Nucleic Acids Res., 15: 9339-9348.
14. Dretzen, G., Bellard, M., Sassone-Corsi, P., Chambon, P. (1981)
Anal. Biochem., 112: 295-298.
15. Mandel, M. & Higa, A. (1970)
J. Mol. Biol., 53: 159.
16. Cohen, S.N., Chang, A. C. Y., Hsu, L. (1972)
Proc. Natl. Acad. Sci/USA., 69: 2110.
17. Grunstein, M., Hogness, D. (1975)
Proc. Natl. Acad. Sci/USA., 72: 3961-65.
18. Southern, E.M. (1975)
J. Mol. Biol., 98: 503-517.
19. Alwine, J.C., Kemp, D.J. & Stark, C.R. (1977)
Proc. Natl. Acad. Sci/USA., 74: 5350-5354.
20. Rigby, P., Dieckmann, M., Rhodes, C., Berg, P. (1977)
J. Mol. Biol., 113: 237-251.

21. Feinberg, A.P., Vogelstein, B. (1983)
Anal. Biochem., 132: 6-13.
22. Feinberg, A.P., Vogelstein, B. (1984)
Anal. Biochem., 137: 266-267.
23. Rivin, C.J., Cullis, C.A., Walbot, V. (1986)
Genetics, 113: 1009-1019.
24. Nei, M. and Li, W. H. (1979)
Proc. Natl. Acad. Sci./USA, 76: 5269-5273.
25. Nei, M (1987) : Molecular evolutionary genetics.
Columbia University Press, New York, pp. 106-107 .
26. Fergusson, A. (1980)
In. Biochemical Systematics and Evolution.
Blackie and Sons Limited, Bishopbriggs, Glasgow,
U.K.
27. Sneath, P.H.A. and Sokal, R.R. (1973)
Numerical Taxonomy, W.H. Freeman & Co.,
San Francisco.
28. Fitch, W.M. and Margoliash, E. (1967)
Science, 155: 279-284.
29. Farris, J.S. (1972)
Amer. Nat., 106: 645-668.

CHAPTER THREE

RESULTS

SECTION I

3.1 Construction of a partial genomic library of rice

My interest was mainly in studying the repeated DNA sequences in rice, which comprise about 52% of total genome (1). Rice DNA in general, when digested with different restriction enzymes, shows presence of repeat families with very few enzymes, such as AvaI, EcoRV, and DraI. Moreover, one cannot isolate and characterize a particular repeat sequence/or family unless it is cloned. This implies that construction of a library and further screening it with suitable probe/(s) becomes a first step in identifying & isolating such sequences.

3.1.1 Cloning of rice DNA for generation of a partial genomic library and screening it for repetitive clones

The library used in this study was constructed in the plasmid vector pUC18 using restriction enzyme PstI and rice O. sativa cv. Malkolam. The pUC18 plasmid was selected because of easy screening procedures and direct use of the cloned fragment for expression studies and sequencing. PstI, a methylation sensitive enzyme, recognizes the sequence 5' -- CTGCAG -- 3' and is sensitive to cytosine methylation in the 5' location (2). Clones represented in the library used

in this study are, therefore, likely to be from the regions of the genome which are not methylated at this site. Malkolam which is an indica rice was selected because it is locally cultivated and is a landrace. This implies that it has a conserved gene pool which is quite different from the modern cultivars and is also likely to contain genes of commercial value. It has a few elite characters like early maturity, medium slender (MS) type of grain and good cooking qualities.

For the library construction, rice DNA was digested to completion with PstI and the region of the fragments corresponding to 0.2-2kbp was eluted and ligated to the PstI restricted vector DNA (Fig 3.1A and B). Chimeric plasmids were transformed into E. coli DH5 α , and the recombinants were selected on X gal-Amp-LB plates. Colony lifts from the partial library were probed with P - 32-labelled total DNA to screen the clones for repetitive sequences. The clones were identified as repeated sequences based on their strong hybridization to labelled rice genomic DNA, while remaining clones revealing faint or no hybridization mainly contained low/single copy sequences (Fig. 3.2). The library represented approximately 45% repeat clones and 55% low/single copy clones.

3.1.2 Determination of insert sizes in a few random genomic clones

In order to confirm the presence of inserts and determine the insert sizes, DNA from a few random genomic clones were isolated and digested either with PstI (Fig.3.3 A, lanes b-l) or EcoRI plus HindIII, (Fig.3.3 B, lanes c-n), followed by electrophoresis on 1% agarose gels.

Fig.3.3 A lanes b-l reveals that the recombinants (pOSM2B-2-pOSM2B-12) are heterogeneous with respect to the length of rice DNA fragment inserted in them, and the size of the insert ranges from 0.2-2.0 kbp. The DNA band at a fragment length of 2.7 kbp represents the linearised pUC18 vector DNA.

The presence and size of the inserts was alternatively determined by digestion of a few random clones (pOSM1C-1-pOSM1C-12) with EcoRI plus HindIII (Fig.3.3 B, lanes c-n). While all the clones reveal presence of inserts, again in the same range, two clones namely, pOSM1C-5 (lane g) and pOSM1C-10 (lane l), show two insert bands. This may be due to the presence of either EcoRI or HindIII sites in the cloned fragment.

3.1.3 Identification of a repeat sequence pOSM1C-2

In my initial DNA analysis for distinguishing three rice cultivars (namely IR36, Malkolam & IR50) a few probes (repetitive and low copy) from the partial PstI genomic library were challenged against the DNA digests of these cultivars with a few restriction enzymes. During this study, my attention was attracted by a repeat sequence which showed a rather unique pattern of hybridization. For eg. when the digests of these cultivars with EcoRI, BamHI, HindIII and PstI were probed with clones like pOSM2D-5, pOSM1E-8, pOSM1A-4, pOSM1A-6 and pOSM1C-2, pOSM1C-2 showed a distinct pattern of hybridization with prominent bands on a smeary background, in the EcoRI digests. This repeat, in general, appeared to be dispersed through the rice genome. In rice, as on today, very few dispersed repeats have been identified and characterized (3,4). I, therefore, decided to concentrate on this particular repeat sequence and study its nature, organization, structural features, and abundance in the rice genome.

Fig.3.1 A

Agarose gel electrophoresis of rice DNA (cv. Malkolam)
digested with PstI.

Lane a: Undigested rice DNA as control

Lane b: Lambda DNA digested with EcoRI+HindIII as
molecular weight marker

Lane c: rice DNA digested with PstI

Fig.3.1 B

Agarose gel electrophoresis of vector (pUC18) DNA

Lane a: Undigested vector pUC 18

Lane b: pUC 18 DNA digested with PstI

Lane c: Baptized linearised pUC18 DNA

Lane d: Ligation of vector and insert DNA

Electrophoresis was carried out on 1% neutral agarose
gel in 1 X TAE (pH 8.1) at a constant current of 30
mA.

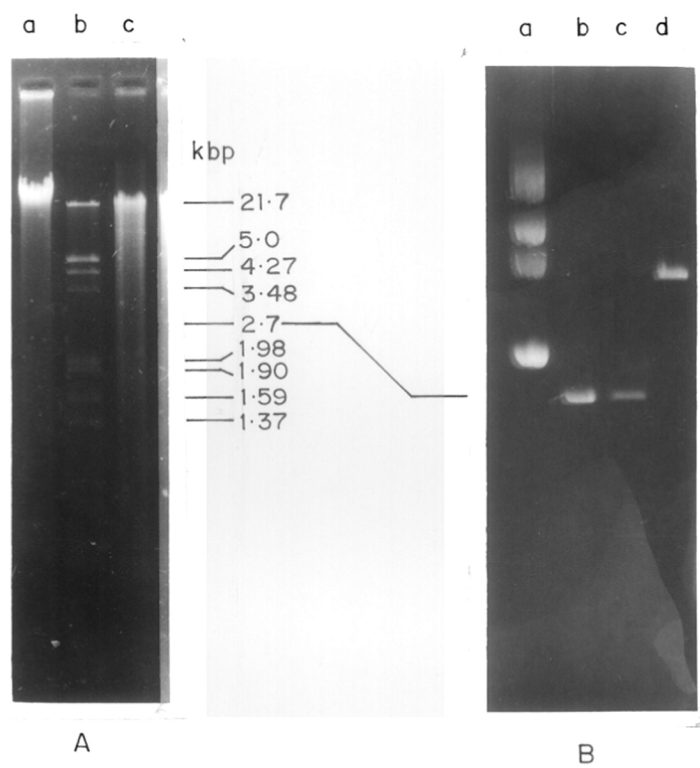


FIG. 3.1

Fig 3.2

Colony hybridization of recombinants (on Genescreen Plus Filter) with P-32-labelled rice (var.Malkolam) genomic DNA for screening the library.

The intense signals represent clones containing highly repeated DNA sequences.

POSMIC-2 —

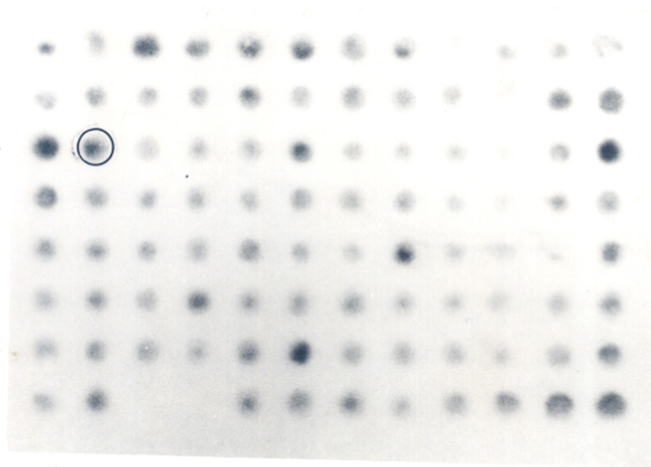


FIG. 3·2

Fig 3.3 A

Agarose gel electrophoresis pattern of random genomic clones from the rice genomic library digested with PstI.

Lane a: pCA8 DNA digested with HindII as molecular weight marker.

Lane b: pOSM2B-2, Lane c: pOSM2B-3, Lane d: pOSM2B-4

Lane e: pOSM2B-5, Lane f: pOSM2B-6, Lane g: pOSM2B-7

Lane h: pOSM2B-8, Lane i: pOSM2B-9, Lane j: pOSM2B-10

Lane k: pOSM2B-11 Lane l: pOSM2B-12.

Electrophoresis was carried out on 1% neutral agarose slab gels in TAE buffer (pH 8.1) at a constant current of 30 mA.

In lanes b-l, the band at 2.7 kbp is the linear band of pUC18 while bands in the range of 0.5-2.0 kbp are the inserts.

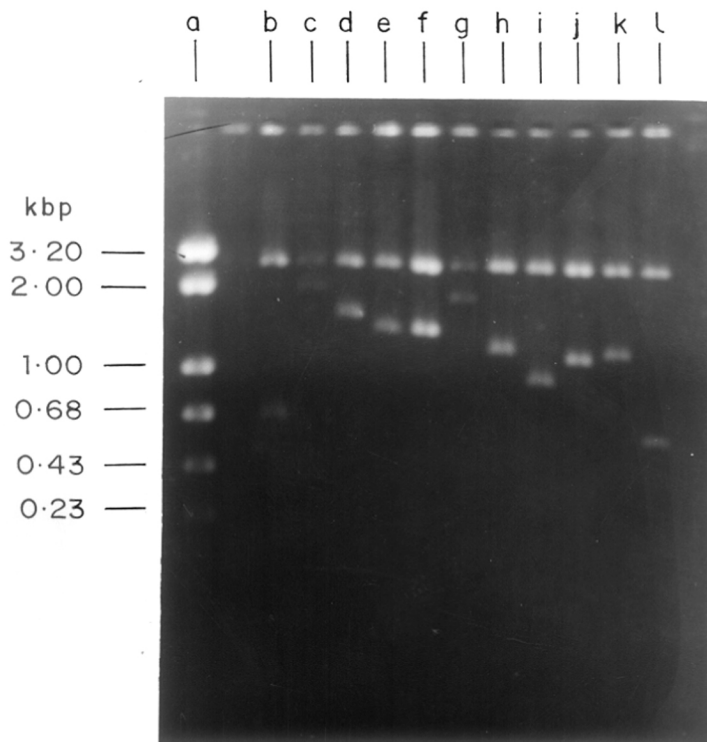


FIG. 3·3 A

Fig.3.3 B

Agarose gel electrophoresis pattern of random genomic clones from the rice genomic library digested with EcoRI+HindIII.

Lane a: Lambda DNA digested with EcoRI+HindIII as molecular weight marker.

Lane b: pOSM1C-1 (undigested DNA).

Lane c: pOSM1C-1, Lane d: pOSM1C-2, lane e: pOSM1C-3,

Lane f: pOSM1C-4, Lane g: pOSM1C-5, lane h: pOSM1C-6,

Lane i: pOSM1C-7, Lane j: pOSM1C-8, lane k: pOSM1C-9,

Lane l: pOSM1C-10, Lane m: pOSM1C-11, lane n: pOSM1C-12.

Electrophoresis was carried out on 1% neutral agarose slab gels in TAE buffer (pH 8.1) at a constant current of 30 mA.

Fig.3.3 C

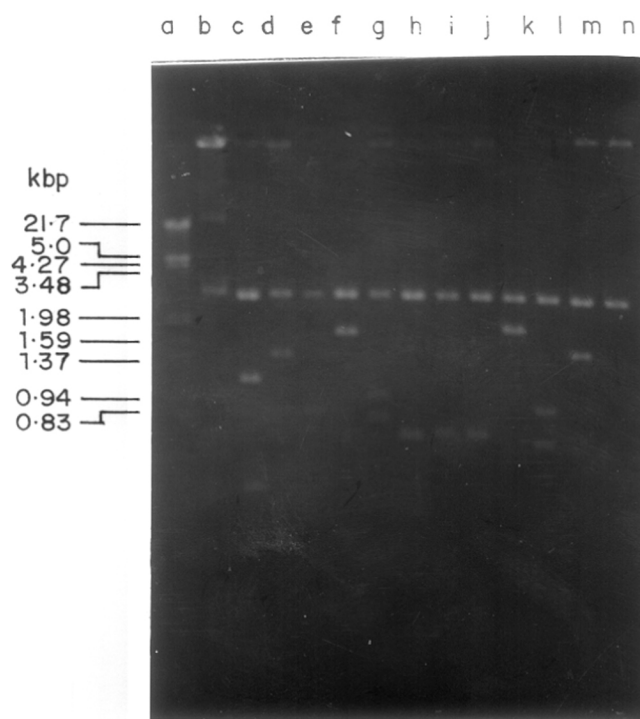
Agarose gel electrophoresis pattern of pOSM1C-2 DNA digested with PstI.

Lane a: Undigested pOSM1C-2 DNA.

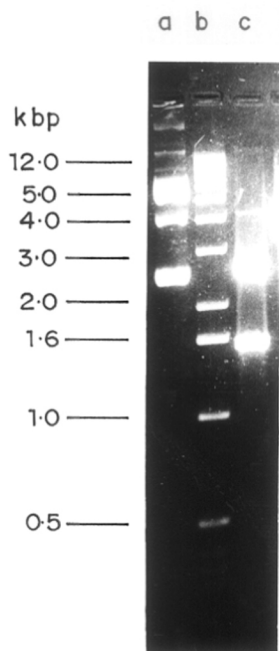
Lane b: Lambda DNA 1kb ladder as molecular weight marker.

Lane c: pOSM1C-2 digested with PstI.

The band at 2.7 kbp is pUC 18 DNA while the band at 1.55 kbp is the pOSM1C-2 insert.



B



C

FIG. 3.3

SECTION II

3.2 Characterization of the repeat sequence pOSM1C-2

As explained in the previous section, pOSM1C-2 is a repeat sequence identified and isolated from the partial PstI genomic library of rice O. sativa cv. Malkolam in pUC18. It gave a moderately intense signal, when probed with P-32-labelled total rice genomic DNA, on the colony autoradiogram (Fig.3.2). Digestion of pOSM1C-2 with PstI revealed an insert of 1.55 kbp (Fig.3.3 C, lane c). Its unique and interesting pattern of hybridization with three rice DNA digests namely, IR36, Malkolam and IR50, particularly with EcoRI, prompted me to characterize this sequence in detail. This section describes the characterization of pOSM1C-2 with respect to its nature and organization in rice genome, structural features, abundance and methylation status. I anticipate that this information could be useful in arriving at a general picture of organization of dispersed repeated sequences in rice.

3.2.1 Nature of pOSM1C-2 repeat sequence in the rice genome

Repetitive sequences are usually organized either as tandem repeats or dispersed repeats located in various loci on the genome (5). In order to address

the question of how pOSM1C-2 sequence is organized in the rice genome, a series of partial and complete digestions of rice DNA with different amounts of PstI were set up and digestion products were separated by electrophoresis on 1% agarose gel (Fig.3.4 A). The digests were then challenged with P-32-labelled pOSM1C-2 insert (Fig.3.4 B). It is clearly seen from the figure that no distinct bands other than two hybridizing fragments namely 1.55 kbp and 1.76 kbp, superimposed on a faint background smear are detected in all the digests. Tandem organization of pOSM1C-2 sequence would result in a very characteristic ladder of oligomers as the products of incomplete hydrolysis that hybridize to the radioactive probe. The results of this experiment strongly indicate non-tandem or dispersed character of pOSM1C-2 sequence in the genome. In other words, the repeat sequence appears to be scattered throughout the genome and interspersed with unique or other non-homologous repeated sequences. The repeat typically shows "clustered and scrambled" organization, wherein, the pOSM1C-2 sequence would be frequently found in clusters of repeated sequences that are scrambled in different genomic locations.

Fig.3.4 A

Electrophoresis on 1.0% agarose of Malkolam DNA digested with different amounts of PstI.

Lanes a & b: High molecular weight Malkolam DNA as control.

Lane c : Lambda DNA 1kb ladder as molecular weight marker.

Lane d: 1U/ug, Lane e: 6U/ug, Lane f: 10U/ug

Lane g: 20 U/ug, Lane h: 40 U/ug.

Fig.3.4 B

Southern hybridization of P-32-labelled pOSM1C-2 insert (1.55 kbp PstI fragment) with the PstI digests of Malkolam DNA. Lanewise order is the same as described in Fig. 3.4 A.

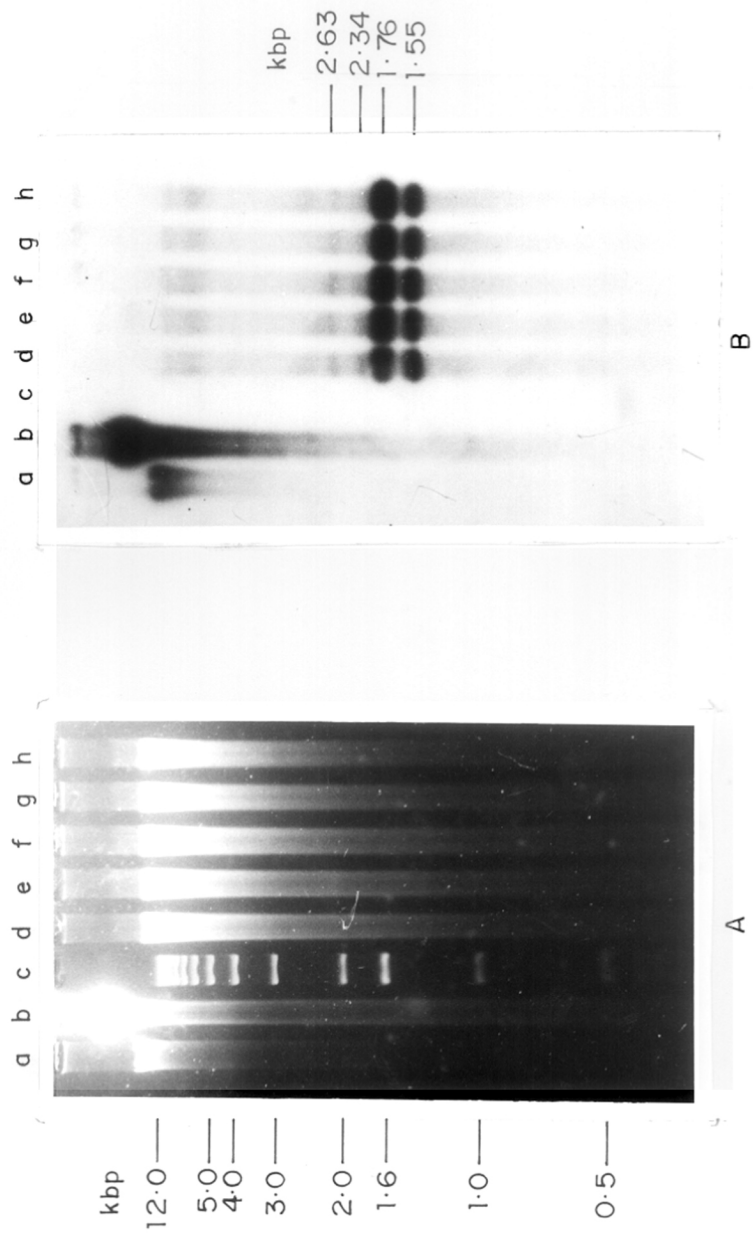


FIG. 3.4

3.2.2 Organization of the pOSM1C-2 repeat sequence in the rice genome

After confirming that the pOSM1C-2 sequence is a dispersed repeat, its organization was studied in the rice genome. For this purpose, rice DNA was digested with fourteen different restriction enzymes, namely EcoRI, BamHI, HindIII, BglII, PstI, HaeIII, BglI, XbaI, PvuII, SacI, ScaI, SalI, EcoRV, and BstEII, the digests size fractionated on 1% agarose gel (Fig.3.5 A) and probed with P-32-labelled pOSM1C-2 insert (Fig.3.5 B). An overall pattern of 2-3 prominent bands superimposed on a faint background smear was observed in most of the digests. The organization can be grouped into two categories. In the case of BamHI, SacI, SalI and BstE-II digests, the hybridization of the probe is restricted to high molecular weight DNA (Fig.3.5 B, lanes c,k,m and o). This may be due to less frequent sites of these enzymes in the repetitive stretches homologous to the probe. Similar observations were reported for a BamHI family in Nicotiana tabacum, where most of the DNA, homologous to HSR 60.1, was localized in cases of XbaI, BglII and AluI digests in the undigested "relic" DNAs (6). However, the products of digestion by EcoRI, HindIII, BglII, PstI, HaeIII, BglI, XbaI, PvuII, ScaI and EcoRV, when separated on agarose gel resulted in

strong discrete bands in the blot hybridization experiments (Fig.3.5 B, lanes b,d,e,f,g,h,i,j,l & n). The appearance of these distinct bands on the autoradiogram may be due to the occurrence of frequent sites for these enzymes in the genome, either in the vicinity or in the region that is homologous to the probe. When the pOSM1C-2 sequence is found in different genomic locations, it may have different repetitive flanking sequences and hence different external restriction sites. The occurrence of strong discrete bands of hybridization as seen in Fig.3.5 B can be considered as an evidence for the clustered/ scrambled arrangement of pOSM1C-2.

3.2.3 The abundance of pOSM1C-2 repeat sequence in the rice genome

The copy number of pOSM1C-2 repeat sequence in the rice genome was determined by quantitative dot blot hybridization. A dilution series of the cloned repeat was used as copy number standard (Fig.3.6 Column A). Three different amounts of genomic DNA were loaded on the same filter (Fig.3.6 Column B). The extent of hybridization was quantitated by tracing the autoradiogram with a densitometer as described under materials and methods. For the copy number calculation, the size

Fig.3.5 A

Electrophoresis pattern of Malkolam DNA digested to completion with different enzymes.

Lane a: Lambda EcoRI+HindIII digest as molecular weight marker.

Lane b : EcoRI, lane c : BamHI, Lane d : HindIII,

Lane e : BglII, Lane f : PstI, Lane g : HaeIII,

Lane h : BglI, Lane i : XbaI, Lane j : PvuII,

Lane k : SacI, Lane l : ScaI, Lane m : SalI,

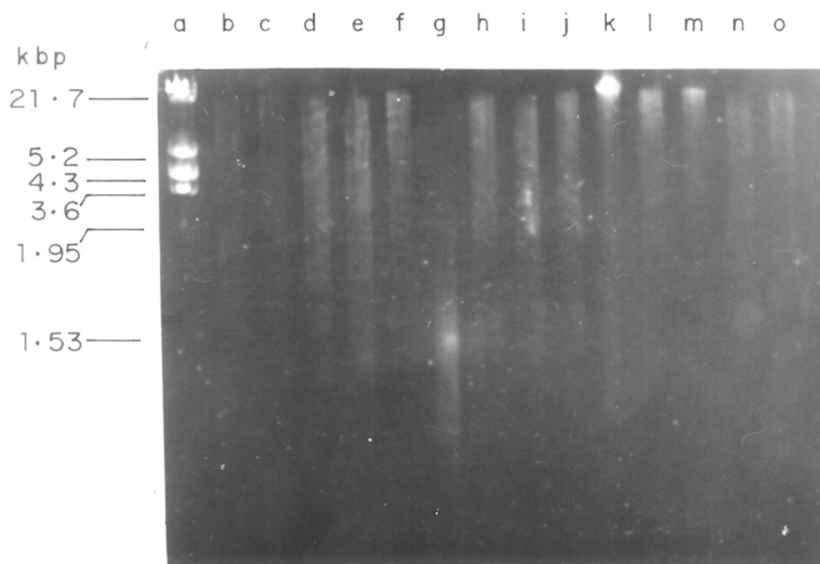
Lane n : EcoRV, Lane o : BstEII.

Electrophoresis was carried out on 1% neutral agarose slab gels in TAE (1X) buffer pH 8.0, at a constant current of 30 mA.

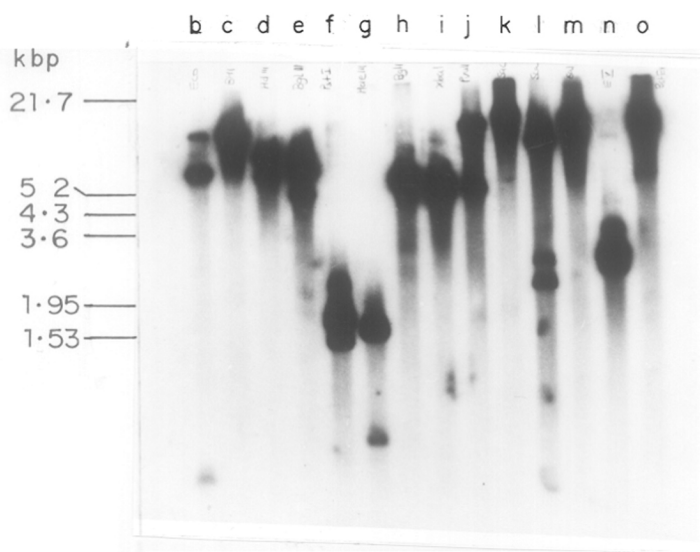
Fig.3.5 B

Southern hybridization of P-32-labelled pOSM1C-2 insert (1.55 kbp PstI fragment) with Malkolam DNA digested with different enzymes.

Lanewise order of the enzymes is same as described in Fig 3.5 A.



A



B

FIG. 3.5

Fig.3.6

Copy number determination of pOSM1C-2 repeat sequence in rice genome.

Different amounts of DNA were dot-blotted on to Hybond-C membrane (Amersham). Copy number standards (same as the probe) were used in column A. The numbers to the left of the column refer to copy numbers.

Different amounts of total rice DNA were spotted in column B. top lane 10ng, middle lane: 100 ng; bottom lane 1000ng.

The filter was probed with P-32-labelled pOSM1C-2 insert (1.55 kbp PstI fragment).

Quantitation of hybridization signal was determined by tracing X-ray film with a densitometer.

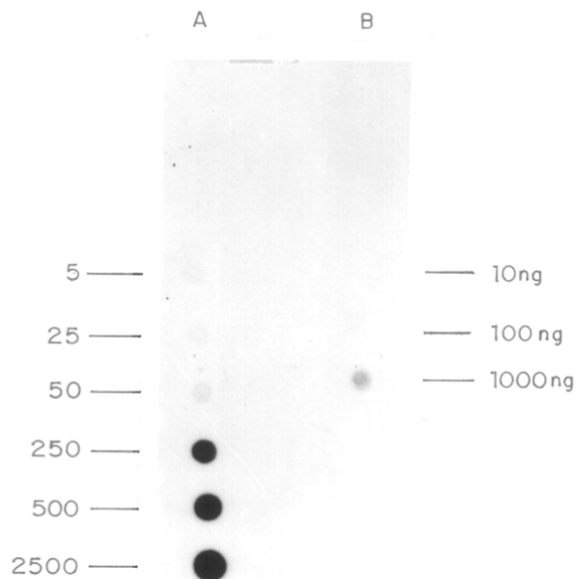


FIG. 3·6

of rice haploid genome was taken as 1.0×10^9 bp (7). The copy number was estimated to be 150 copies per haploid genome. Depending on the copy number, length of the cloned repeat sequence and rice haploid genome size, this repeat sequence comprises 0.037% of rice genome.

3.2.4 Restriction mapping of pOSM1C-2

Different hexacutter restriction endonucleases were used for mapping the pOSM1C-2 sequence. Initial mapping was done using the enzymes which have unique sites and are from the multiple cloning site region of the vector pUC18. These included EcoRI, HindIII, AccI, XbaI, BamHI, SmaI & KpnI. Fig.3.7 A shows the electrophoresis pattern of pOSM1C-2 sequence digested with each of these enzymes. In the case of HindIII (lane c), XbaI (lane e), BamHI (lane f) & KpnI (lane h) digests, a single fragment of 4.1-4.2 kbp is seen, ie there is just linearization of the cloned repeat sequence, thus indicating the absence of these restriction endonuclease sites in the pOSM1C-2 sequence. In the case of EcoRI (lane b), AccI (lane d) and SmaI (lane g) digests, on the other hand, two fragments are observed, with molecular weights of 4.00 kbp and 0.230 kbp in EcoRI digest, 3.80 kbp & 0.450 kbp

in AccI digest and 3.3 kbp & 0.9 kbp in SmaI digest. This suggests that the cloned repeat possesses single/unique site for these enzymes. The restriction map of the repeat with respect to these enzymes sites is depicted in Fig.3.7 D. Among other restriction enzymes which were tried for mapping were enzymes like EcoRV, ClaI, BglI, HincII, PvuII and ScaI. EcoRV and ClaI have no site in pUC18, BglI and PvuII have two sites and HincII and ScaI have single site in the plasmid pUC18 DNA. As seen from the Fig.3.7 B, double digestion of pOSM1C-2 with PstI and EcoRV (lane b) and PstI + ClaI (lane c) gives two bands of 2.7 kbp and 1.55 kbp; the former corresponds to the linearized vector, while the latter is pOSM1C-2 insert. The pattern of digestion is the same as digestion of pOSM1C-2 with PstI alone (Fig.3.7 B, lane a). This indicates that the cloned repeat does not possess sites for both EcoRV & ClaI in it.

In double digest of pOSM1C-2 with PstI and BglI (Fig.3.7 B, lane d), four fragments of 1.55 kbp, 1.40 kbp, 1.07 kbp and 0.168 kbp are observed. BglI has two sites in pUC18 DNA at position 245 and 1813 respectively. 1.55 kbp band corresponds to pOSM1C-2 insert DNA obtained by PstI digestion of pOSM1C-2. The three latter fragments in the digest (namely 1.40 kbp, 1.07

kbp and 0.168 kbp) correspond to BglI-BglI, BglI-PstI and PstI-BglI fragments of pUC18 DNA respectively. The above pattern of restriction of pOSM1C-2 by BglI suggests the absence of this endonuclease site in pOSM1C-2

The digestion pattern of pOSM1C-2 with PstI and HindII is shown in Fig.3.7.B, lane g. Four fragments of 2.7 kbp, 0.90 kbp, 0.450 kbp and 0.20 kbp are observed. 2.7 kbp fragment corresponds to the linearized pUC18 DNA. There is a single HindII site in pUC18 at 419 bp. The 0.90 kbp, 0.20 kbp and 0.45 kbp fragments correspond to HincII (site in pUC18)-HincII (site 1 in insert), HincII (site 1 in insert)-HincII (site II in insert) and HincII-PstI fragments respectively, thus suggesting the presence of two HincII sites in the cloned repeat. The restriction map of pOSM1C-2 with respect to HincII sites is shown in Fig.3.7 D.

Digestion pattern of pOSM1C-2 with PvuII is shown in Fig.3.7 C, lane a. PvuII has two sites in pUC18 DNA at 306 and 628 bp respectively. Two fragments of molecular weights 2.36 kbp and 1.87 kbp are observed in the PvuII digest of pOSM1C-2. The 2.36 kbp band corresponds to the pUC18 PvuII fragment, while 1.87 kbp band appears due to the insert DNA fragment, flanked by 2 PvuII sites at 306 and 628 in pUC18. 1.87 kbp frag-

Fig.3.7

Digestion of pOSM1C-2 with different restriction endonucleases.

Fig. 3.7 A

Lane a : pCA8 HindII DNA digest as molecular weight marker.

Lane b : EcoRI, Lane c: HindIII, Lane d: AccI,

Lane e : XbaI, Lane f: BamHI, Lane g: SmaI,

Lane h : KpnI.

Fig.3.7 B

Lane a : PstI, Lane b : PstI+EcoRV,

Lane c : PstI+ClaI, Lane d : PstI+BglI,

Lane e & f : Lambda 1kb DNA ladder as molecular weight marker.

Lane g : PstI+HindII.

Fig.3.7 C

Lane a : PvuII.

Lane b : ScaI.

Electrophoresis was carried out on 1% neutral agarose gels in TAE (1X) pH 8.0 at a constant current of 30 mA.

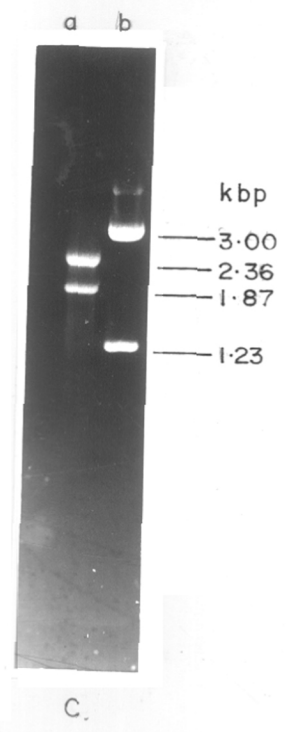
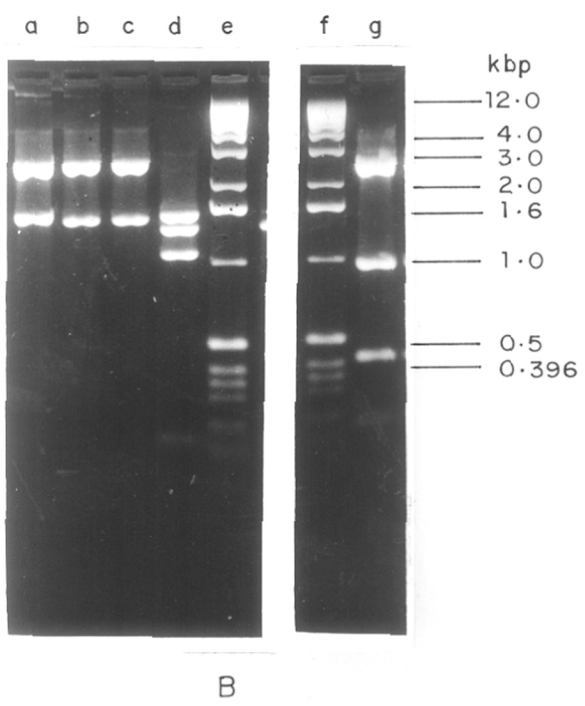
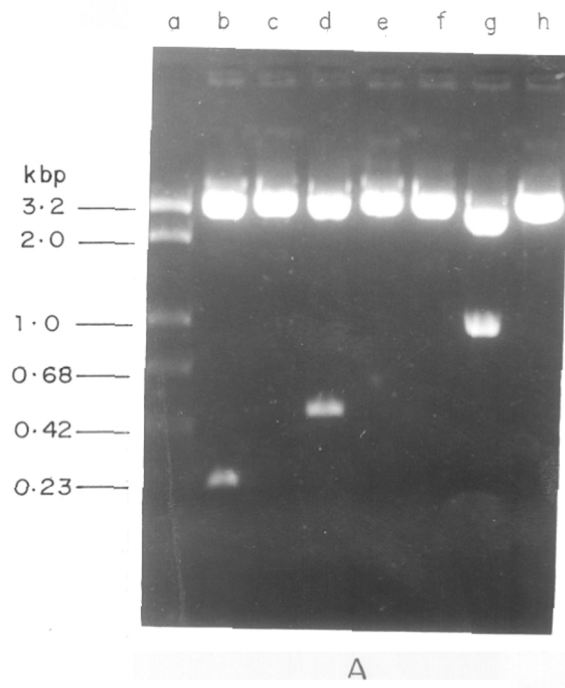
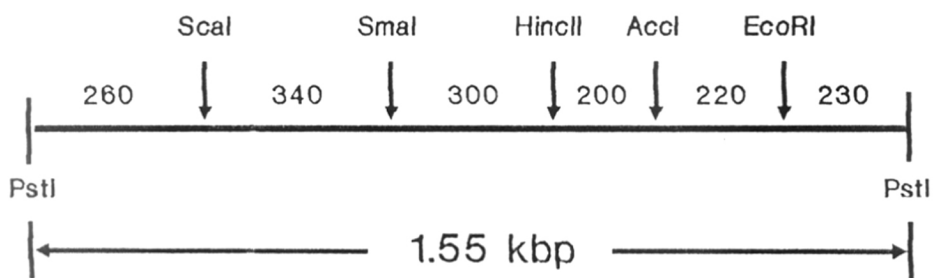


FIG. 3.7

Fig.3.7 D

Line diagram showing the restriction enzyme sites in the cloned sequence pOSM1C-2.



Restrictase Map of pOSM1C-2

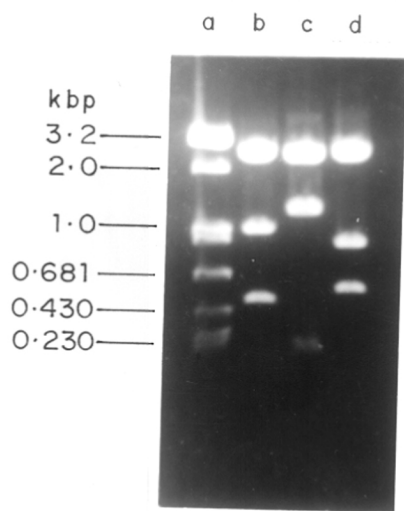
Figure 3.7D

ment includes 107 bp PvuII-PstI and 1550 bp Pst-PstI and 215 bp PstI-PvuII (at 628) fragments. This again indicates the absence of PvuII site in the cloned repeat.

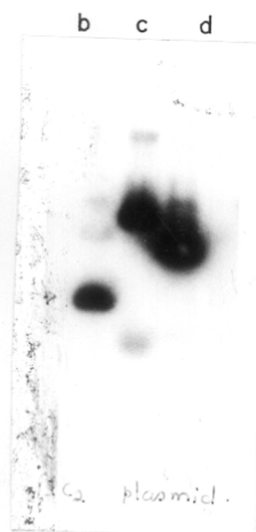
The products of digestion of pOSM1C-2 with ScaI are shown in Fig.3.7 C, lane b. pUC18 DNA has a single site for ScaI at 2177. Absence of ScaI site in the cloned repeat would then give a single linearized band of pOSM1C-2. The digests however, show two bands of 3.0 kbp and 1.2 kbp respectively, suggesting the presence of a single site in the cloned repeat fragment. The ScaI site has been shown in Fig. 3.7 D.

3.2.5 Homology studies within the repeat sequence pOSM1C-2

From the restriction mapping of the pOSM1C-2 fragment, I noticed that the five restriction enzymes, having unique sites in this repeat, produced many smaller fragments (fragment length 200-340 bp) all along the repeat. (Fig.3.7 D). Hence, I was interested in assessing the interrelationships within the repeat units in terms of sequence homologies. For this purpose, the AccI-PstI fragment (0.450 kbp) (Fig.3.8 A, lane b) was randomly selected and used as a probe against the double digests of cloned repeat pOSM1C-2



A



B

FIG. 3-8

Fig. 3.8 A

Agarose gel electrophoresis of double digests of pOSM1C-2 DNA

Lane b: Pst-AccI

Lane c: PstI-EcoRI

Lane d: PstI-SmaI

Lane a: HindII digested pCA8 DNA as molecular weight marker

Electrophoresis was carried out on 1% neutral agarose slab gel run in TAE (1X) buffer, pH 8.0 at a constant current of 30 mA.

Fig 3.8 B

Southern blot hybridization of P-32-labelled 0.45 kbp AccI-PstI fragment to the double digests of pOSM1C-2 DNA described in Fig.3.8 A.

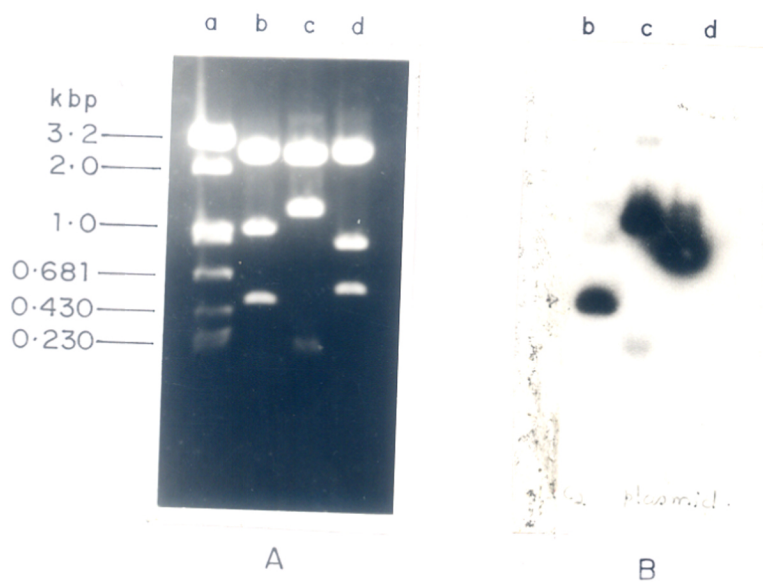


FIG. 3·8

with PstI-AccI (Fig.3.8 A, lane b), PstI-EcoRI (Fig.3.8 A, lane c) and PstI-SmaI (Fig.3.8 A, lane d). Fig.3.8 B, lane b shows a very faint hybridization signal with the PstI-AccI larger fragment (1.1 kbp) while very intense hybridization signal with AccI-PstI smaller fragment (0.45 kbp) which is the probe itself. In the PstI-EcoRI digest (Fig.3.8 B, lane c), the probe shows hybridization with both the PstI-EcoRI fragments, while in case of PstI-SmaI digest (Fig.3.8 B, lane d) the probe strongly hybridizes to 0.9 kbp PstI-SmaI fragment, with practically no hybridization signal in the other part of the repeat sequence (0.6 kbp SmaI-PstI fragment). These observations led me to suspect the probable presence of at least two independent repeats namely, PstI-AccI (1.1 kbp) repeat and AccI-PstI (0.450 kbp) repeat within the repeat unit pOSM1C-2, and these two will be termed as LSR (Large subrepeat) and SSR (small subrepeat) respectively.

3.2.6 Organization of LSR and SSR in the genome

During the study of organization of the two subrepeats namely LSR and SSR in the rice genome, genomic DNA was digested to completion with EcoRI and PstI and these digests were probed with P-32-labelled LSR and SSR (Fig.3.9 A & B). The two fragments differ

in hybridization patterns with both the enzymes. The absence of a 'ladder' on the genomic blots again depict their non-tandem organization. In the EcoRI digest (Fig.3.9 A, lane a), 'LSR' shows hybridization signals in high molecular weight region (18.47 kbp and 9.9 kbp), whereas 'SSR' though shows hybridization in the same region, gives additional bands in the high molecular weight (15.29) as well as low molecular weight regions 0.4 kbp and 0.238 kbp respectively (Fig.3.9 A, lane b). In the PstI digests, though two hybridizing bands are shared by LSR and SSR (1.55 kbp and 1.76 kbp) (Fig.3.9 B, lanes a and b), SSR again gives two additional faint hybridization bands at 2.32 kbp and 2.64 kbp (Fig.3.9 B, lane b). This suggests that SSR, though occurs as a part of PstI repeat (1.55 kbp fragment), is an independent unit, occurring in a different environment in the genome. LSR, however, occurs as a part of PstI repeat only and has no independent occurrence in the genome.

3.2.7 Methylation status of pOSM1C-2 repeat sequence at different stages of development in rice (namely embryo, shoot and leaf)

Dhar et al. (8) have shown predominance and tissue specificity of adenine methylation in

Fig. 3.9 A

Southern blot hybridization of P-32-labelled LSR & SSR
to the digest of Malkolam DNA with EcoRI.

Lane a : LSR as probe (PstI-AccI fragment 1.1 kbp)

Lane b : SSR as probe (AccI-PstI fragment 0.45 kbp)

Fig 3.9 B

Southern blot hybridization of P-32-labelled LSR & SSR
to the digest of Malkolam DNA with PstI.

Lane a : LSR as probe (PstI-AccI fragment 1.1 kbp)

Lane b : SSR as probe (AccI-PstI fragment 0.45 kbp)

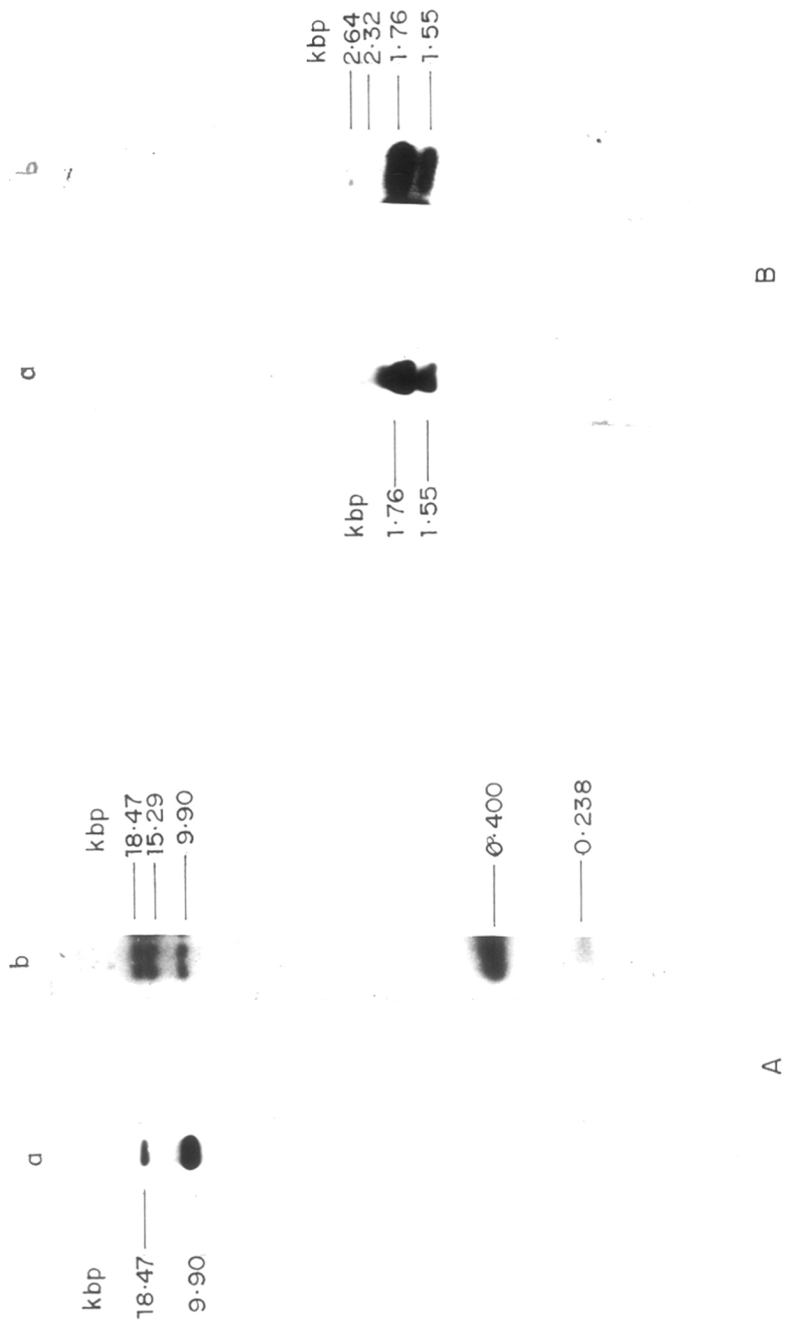


FIG. 3-9

total, repetitive and specific cloned repeat sequence in rice O. sativa cv. Basmati-370 . In the present work, I describe the restriction enzyme analysis of total rice DNA cv. Malkolam using methylation-specific enzymes namely MboI , Sau3AI , DpnI , MspI and HpaII. To investigate whether the methylation patterns are tissue/stage specific, studies were carried out using 3 tissue systems namely embryo, shoot and leaf. The methylation status of a cloned repeat sequence pOSM1C-2 was also studied during the transition from embryo to shoot and leaf stages.

MboI, Sau3AI and DpnI recognize the sequence 5' -GATC--3' but show differential sensitivity towards adenine and cytosine methylation. DpnI cleaves the sequence only if adenine is methylated, while MboI does not cleave the sequence if adenine is methylated, and is insensitive to cytosine methylation. Sau3AI will not cleave the sequence if cytosine is methylated and is insensitive to adenine methylation.

Figure 3.10 A shows the digestion pattern of embryo DNA with MboI (lane b), Sau3AI (lane c) and DpnI (lane d). It is clearly seen that there is practically no digestion of embryo DNA with DpnI (lane d), suggesting the probable absence of adenine methylation in the sequence 5'-GATC-3'. MboI (lane b) on the other

hand, digests the DNA extensively, suggesting the dearth of methylated adenine residues in the sequence 5'-GATC-3' at the embryonic stage in rice. Compared to MboI digestion (lane b), the digestion by Sau3AI (lane c) is slightly less indicating that part of 5'-GATC-3' sequences are methylated at 'C'. It is clear from Fig.3.10, B & C that DpnI digests the leaf DNA (Fig.3.10 C, lane c) more than the shoot DNA (Fig.3.10 B, lane e), which in turn is more than the embryo DNA (Fig.3.10 A, lane d). Based on the DpnI digestion patterns of embryo, shoot and leaf DNA, it appears that there is an increase in the number of methylated adenine residues during rice development. The MboI/Sau3AI digestion patterns however, indicate that the level of "C" methylation remains more or less the same.

The digestion patterns of rice embryo (lanes e and f), shoot (lanes f and g) and leaf (lanes d and e) DNAs with MspI and HpaII are shown in Fig. 3.10 A,B and C respectively. Here, it is observed that the extent of digestion with MspI is more than that in HpaII in all the three DNAs, indicating that the level of external C methylation is lower as compared to internal C methylation in the sequence 5'-CCGG-3'. This also indicates higher frequency of mCpG dinucleotide as compared to mCpC dinucleotide in these genomes.

Fig 3.10 A

Agarose gel electrophoresis of rice cv. Malkolam embryo DNA digested with methylation sensitive enzymes.

Lane a : Lambda DNA 1kb ladder as molecular weight marker.

Lane b : MboI, Lane c : Sau3AI, Lane d: DpnI,

Lane e : MspI, Lane f : HpaII.

Lane g : Lambda DNA 1kb ladder as molecular weight marker.

Fig 3.10 B

Agarose gel electrophoresis of rice cv. Malkolam shoot DNA (8-10 days old seedling) with methylation sensitive enzymes.

Lane a : Undigested rice DNA as control.

Lane b : Lambda DNA 1kb ladder as molecular weight marker.

Lane c : MboI, Lane d : Sau3AI, Lane e : DpnI,

Lane f : MspI, Lane g : HpaII.

Fig.3.10 C

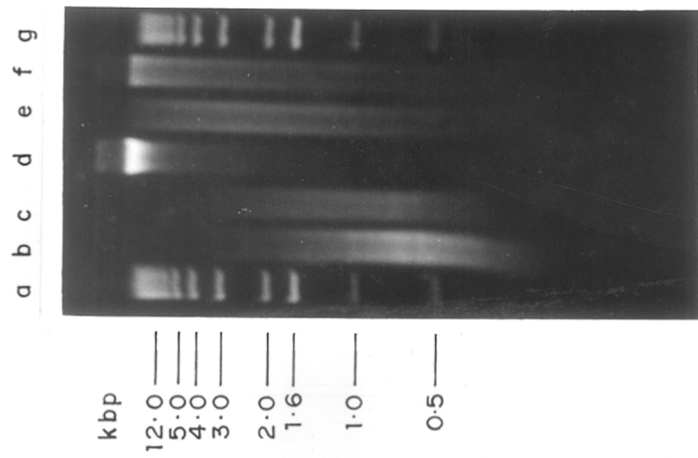
Agarose gel electrophoresis of rice leaf DNA (3-4 weeks old) (cv. Malkolam) digested with methylation sensitive enzymes.

Lane a : MboI, Lane b : Sau3AI, Lane c : DpnI,

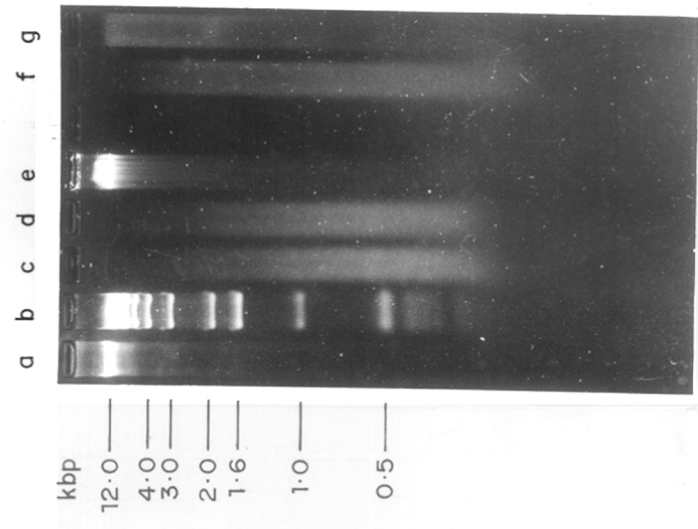
Lane d : MspI, Lane e : HpaII.

Lane f : Lambda DNA 1kb ladder as molecular weight marker.

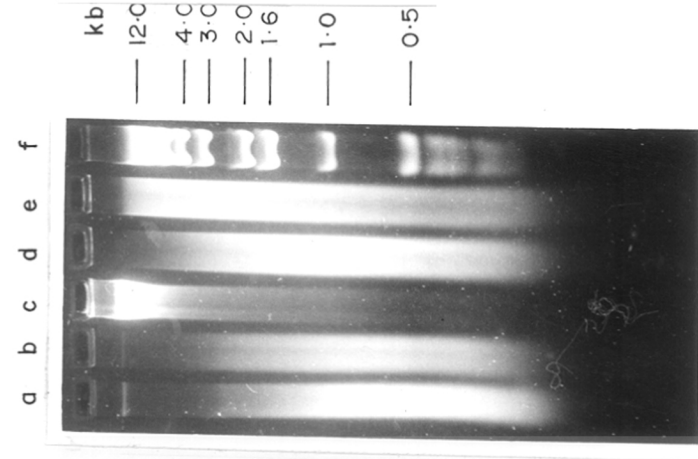
Electrophoresis was carried out on 1% neutral agarose slab gels in TAE 1 X, pH 8.0, at a constant current of 30 mA.



A



B



C

FIG 3.10

In order to study the methylation status of the repeat sequence pOSM1C-2, rice tissue at different stages (namely, embryo, shoot and leaf) was digested with MboI, Sau3AI, DpnI, MspI and HpaII (Fig.3.10) and the digests were probed with P-32-labelled pOSM1C-2 insert (Fig. 3.11). The patterns with MboI and Sau3AI in Fig.3.11, A (lanes b and c), B (lanes c and d) and C (lanes a and b) show that the repeat hybridizes to embryo DNA giving a series of bands in the range of 0.2-3.0 kbp as against only a few bands at shoot and leaf stages. This difference in the hybridization pattern of the repeat may be due to genomic rearrangement /reorganisation partly arising due to differential methylation in these genomes.

The hybridization pattern of P-32-labelled pOSM1C-2 insert with MspI and HpaII digests of embryo DNA is shown in Fig.3.11 A, lanes e & f. In the MspI digest (lane e), the repeat hybridizes giving two distinct bands at 1.27 kbp and 0.87 kbp. The length of two bands is less than the length of the repeat sequence pOSM1C-2 (1.55 kbp). This suggests the presence of an internal site for this enzyme in the repeat itself, where the sequence 5'-CCGG-3' may or may not be methylated at the internal 'C' residue. HpaII digest of

Fig 3.11 A

Southern blot hybridization of P-32-labelled pOSM1C-2 insert with the digests of rice cv. Malkolam embryo DNA. Lane wise order is same as described in Fig 3.10 A.

Fig 3.11 B

Southern blot hybridization of P-32-labelled pOSM1C-2 insert with the digests of rice cv Malkolam shoot DNA (8-10 days old seedlings). Lane wise order is the same as described in Fig 3.10 B.

Fig 3.11 C

Southern blot hybridization of P-32-labelled pOSM1C-2 insert with the digests of rice cv. Malkolam leaf DNA (3-4 weeks old). Lane wise order is the same as described in Fig 3.10 C.

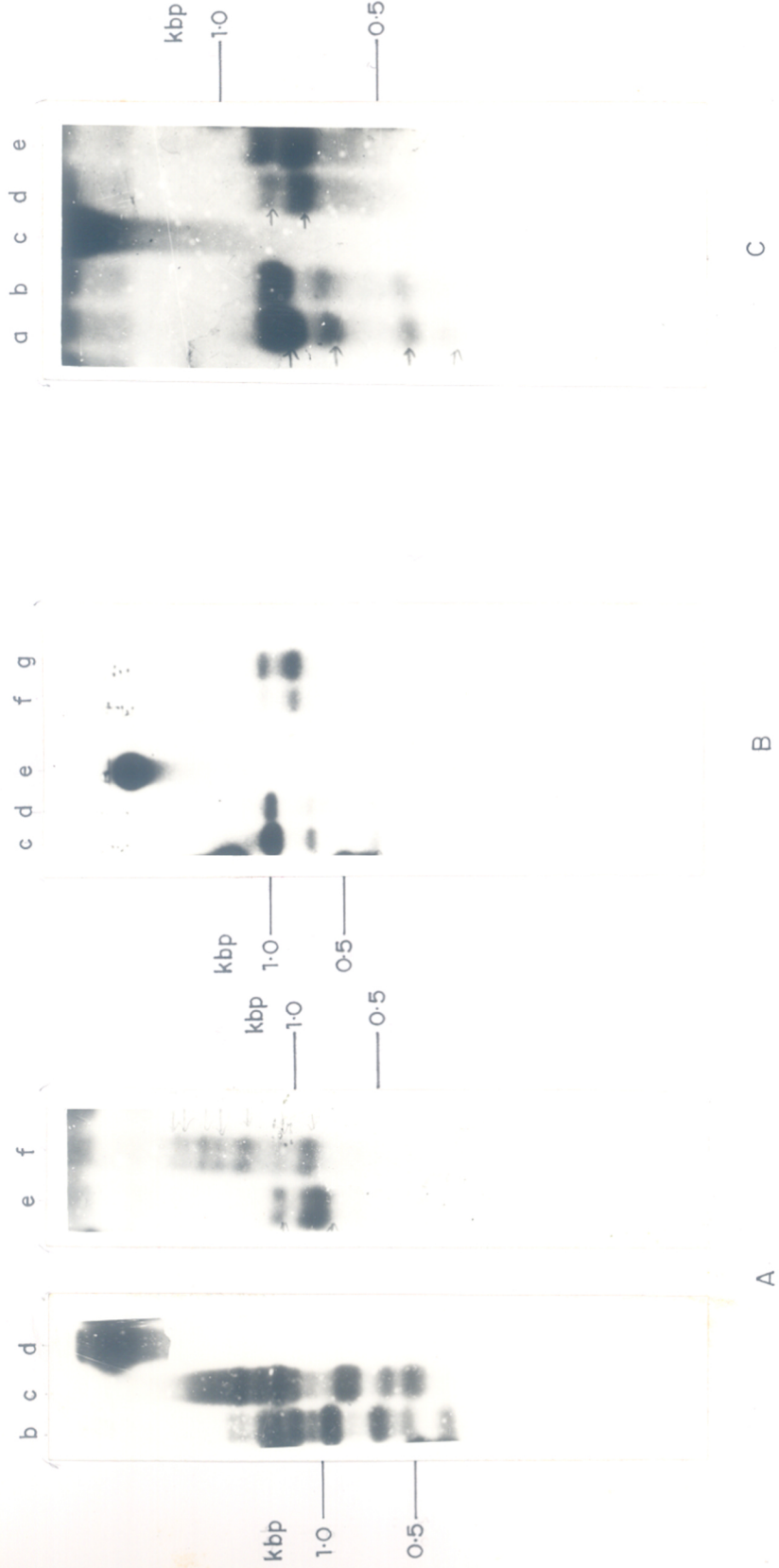


FIG. 3.11

embryo DNA challenged with P-32-labelled pOSM1C-2 insert, however, reveals a number of hybridizing bands (lane f) out of which seven hybridizing bands are in the range of 0.9 kbp-4 kbp. This hybridization pattern suggests the presence of 5'-C^mCGG-3' type of sequences rather than 5'-^mCCGG-3' in the embryo genome. Thus this implies a high frequency of ^mCpG dinucleotide as compared to ^mCpC dinucleotide in the 5'-CCGG-3' sequence in rice embryonic DNA.

The pattern of hybridization of P-32-labelled pOSM1C-2 insert with the MspI and HpaII digests of rice shoot and leaf DNAs are shown in [Fig.3.11 B, lane f, Fig.3.11 C, lane d] and [Fig.3.11 B, lane g, Fig.3.11 C, lane e] respectively. The repeat hybridizes giving two bands at 1.27 kbp and 0.874 kbp, with both MspI (lanes f & d) and HpaII (lanes g & e) at shoot and leaf stages. Compared to the hybridization pattern of pOSM1C-2 with HpaII digest of embryo DNA, (Fig.3.11 A, lane f) the number of hybridizing bands are reduced to only 2 in case of shoot and leaf DNAs digested with HpaII. This indicates less methylation of internal 'C' residues in the sequence 5'-CCGG-3' in the rice genome homologous to the probe. Thus the frequency of internal 'C' methylation is reduced from embryonic stage to the shoot and leaf stage of rice.

Fig. 3.12 A

Electrophoresis pattern of different genome types of rice digested with ScaI.

Lane a : O. eichingeri, Lane b : O. longistaminata,

Lane c : O. australiensis, Lane d : O. latifolia,

Lane e : O. alta,

Lane f : O. officinalis (SD1047),

Lane g : O. officinalis (SD1049)

Lane h : Basmati-370, Lane i : IR-64.

Electrophoresis was carried out on 1% neutral agarose gel in 1X TAE (pH 8.1), at a constant current of 30 mA.

Fig 3.12 B

Southern blot hybridization of P-32-labelled pOSM1C-2 insert (1.55 kbp PstI fragment) with the digests described in Fig.3.12 A.

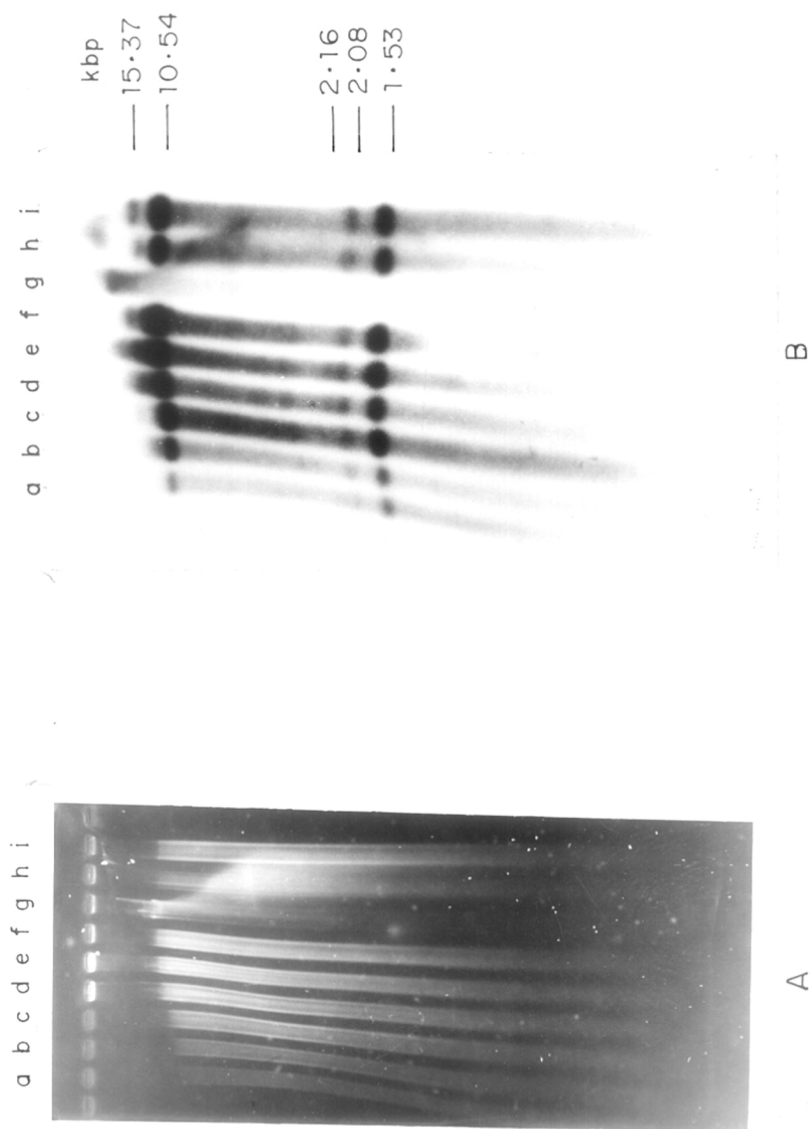


FIG. 3.12

3.2.8 Presence and distribution of pOSM1C-2 repeat sequence in four genome-types of rice

I was interested in knowing if this particular dispersed repeat exhibited genome or species specificity. A number of species were chosen to represent various genomic karyotypes namely AA, CC, CCDD & EE. Rice entries belonging to these genomes were digested with ScaI, the digests were electrophoresed on 1% agarose gel (Fig.3.12 A) and challenged with P- 32-labelled pOSM1C-2 insert (Fig.3.12 B). From Fig.3.12B it is seen that the sequences homologous to pOSM1C-2 are present in all genomes, since the probe shows hybridization signal with all the rice entries analysed. All the rice genomes under investigation display more or less identical pattern with the probe with three prominent hybridizing bands at 1.53 kbp, 2.08 kbp and 10.54 kbp. A general pattern of prominent bands superimposed on a background smear depicts the dispersed as well as conserved nature of the probe in different rice genomes.

SECTION III

3.3 Assessment of genomic/genetic variability among rice germplasm using the dispersed repeat sequence pOSM1C-2

In the last two sections namely 3.1 and 3.2, I have described how I identified and isolated a dispersed repeat sequence pOSM1C-2 from the partial rice library and its characterization with respect to its nature, organization, restriction mapping, methylation status and abundance in the rice genome. My objective next was to study the genomic/genetic interrelationships amongst the elite Maharashtra rice germplasm, using the pOSM1C-2 sequence per se. i.e. I wanted to assess if this repeat sequence proved to be an informative molecular marker in deciphering genomic variability amongst the selected rice germplasm. For this purpose, I used the most recent and fairly simple technology namely 'RFLP' approach. The basic RFLP technique is a powerful tool for plant breeders to characterise and protect lines & varieties. The relatedness between individual progeny and parental material can be evaluated at the seedling or callus stage thus allowing one to speed up the breeding program (9). The variation observed in the size of restriction fragments homologous to the specific probe has been called as 'Restriction Fragment Length Polymorphism' (RFLP)

(10). These differences are a consequence of heritable changes in DNA and include point mutations which create or abolish restriction endonuclease sites in DNA, and DNA rearrangements involving insertions or deletions. It is well known that RFLP technology can be applied to complex genetic problems like development of saturated linkage maps, analysis of quantitative trait loci and assessment of genetic variability among closely related populations. In fact, RFLP markers serve as a powerful tool in detecting, measuring and monitoring variation among closely related population at the molecular level (11). An important aspect of variability is that it can be used by plant breeders and geneticists as a genetic marker (9).

In the present work, 19 cultivars of O. sativa ssp indica were selected for analysis of their genetic variability using RFLP approach. In India, predominantly indica cultivars are grown and used in the breeding programs. (12). There is, therefore, a need towards the understanding of the genetic/genomic interrelationships amongst the elite rice germplasm. All the cultivars are of local origin and are grown in various agroclimatic zones of Maharashtra. All varieties indicated are considered potentially useful parents for future rice breeding programs. The varieties were selected on the basis of some elite characters like scent, fineness,

yield and duration. The list of cultivars used with their elite characters is given in Table 2.1.

The major pre-requisite for RFLP analysis is the digestion of cultivars with restriction endonucleases and subsequent RFLP probing of the digests with genomic or cDNA clones. The digestion patterns of 19 cultivars with six hexacutter endonucleases and subsequent probing of these digests with P-32-labelled pOSM1C-2 insert are described in the following sections .

3.3.1 Restriction enzyme analysis of rice cultivars

Leaf tissue (3-4 weeks old) was harvested and used as source of DNA for all rice varieties (Fig.3.13 A and B). DNA preparations with fragment length higher than 23 kbp were used for all the studies (Fig.3.14). A preliminary analysis of all the DNAs began with their digestion with a few random hexacutters like EcoRI, EcoRV, XbaI, DraI, ScaI and HindIII. These enzymes were selected because (i) they produced an even gradation of fragment sizes from about 500 bp-20 kbp which is a favourable range for RFLP analysis (ii) most of the hexacutters used are insensitive to methylation and (iii) these enzymes are reliable on even partially purified DNA and are inexpensive (except DraI).

Fig 3.13 A & B

Glasshouse photograph of rice germplasm grown in pots
at Agriculture College, Pune.

•



FIG. 3·13 A



FIG. 3·13 B

Fig 3.14

Electrophoresis pattern of high molecular weight DNAs from leaf tissue of different rice cultivars.

Lane a : RP4-14, Lane b : Rasi, Lane c : VDN 7850,

Lane d : KJT-184, Lane e : Jaya,

Lane f: 1kb DNA ladder as molecular weight marker

Lane g : RDN 185-2, Lane h : BG 380-2, Lane i : Jalgaon-5,

Lane j : VDN 3-51-18, Lane k : RTN-73,

Lane l : Lambda 1kb DNA ladder as molecular weight marker

Electrophoresis was carried out on 1% neutral agarose gel in TAE (1X), pH 8.1 at a constant current of 30 mA.

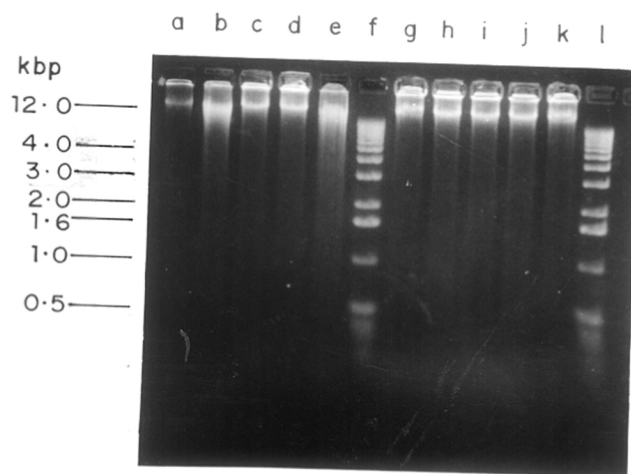


FIG. 3.14

The agarose gel electrophoresis pattern of the DNAs of 19 cultivars with EcoRV is shown in Fig.3.15 A. All the cultivars in general show a continuous smear of fragments ranging from 20 kbp to about 1 kbp. Figure 3.15 B depicts the digestion of 19 DNAs with EcoRI. All the DNAs show digestion to different extent with this enzyme, barring one DNA, namely, VDN-3-51-18 (lane j), which shows practically no digestion.

The restriction enzyme XbaI recognizes the sequence 5'-TCTAGA-3' where the action of the enzyme is inhibited if 'C' or 'A' are methylated. It is clear from Fig.3.15 C, that there is a fairly good digestion of cultivars by XbaI, except cultivars like RP4-14 (lane a), Rasi (lane b), BG 380-2 (lane h), VDN 3-51-18 (lane j) and Pusa-33 (lane s) where digestion leaves a smear of high mol.wt. DNA. This may be either due to the presence of a few recognition sites of this enzyme in the genomes of these accessions or predominance of 'C' or 'A' methylation in the enzyme recognition sequence.

Electrophoresis pattern of DraI digests of 19 cultivars is shown in Fig.3.15 D. All the cultivars except VDN 3-51-18 (lane j) and PLG 39-4-1 are fairly well digested with this enzyme. Since DraI recognizes the sequence 5'-TTTAAA-3' and the repeat sequences in

Fig 3.15 A

Agarose gel electrophoresis pattern of DNAs (5ug/lane) from 19 rice cultivars digested to completion with EcoRV.

Lane a : RP4-14, Lane b : Rasi, Lane c : VDN-7850,

Lane d : KJT-184, Lane e : Jaya,

Lane f : Lambda 1 kb DNA ladder as molecular weight marker.

Lane g : RDN 185-2, Lane h : BG 380-2, Lane i : Jalgaon-5,

Lane j : VDN 3-51-18, Lane k : RTN 73, Lane l : Ratna,

Lane m : VDN-4288, Lane n : RTN-24, Lane o : RTN-711,

Lane p : RTN-1,

Lane q : HindIII digested Lambda DNA digest as molecular weight marker.

Lane r : PLG 39-4-1, Lane s : Pusa-33, Lane t : ACK 5,,

Lane u : KJT-35-3.

Electrophoresis was carried out on 1.0% neutral agarose slab gels in TAE (1X) buffer (pH 8.0) at a constant current of 20 mA for 12-15 hrs.

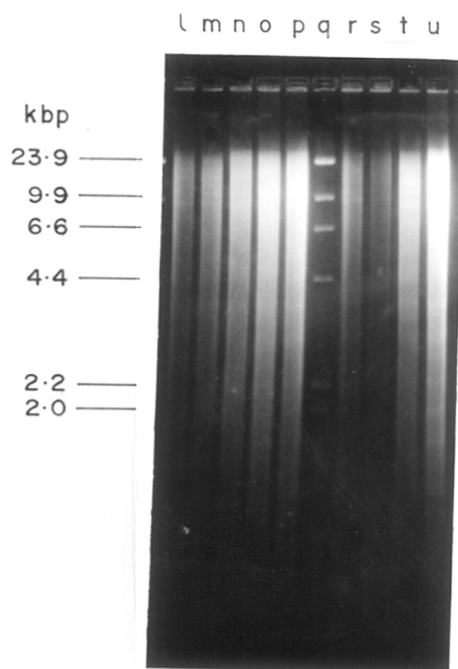
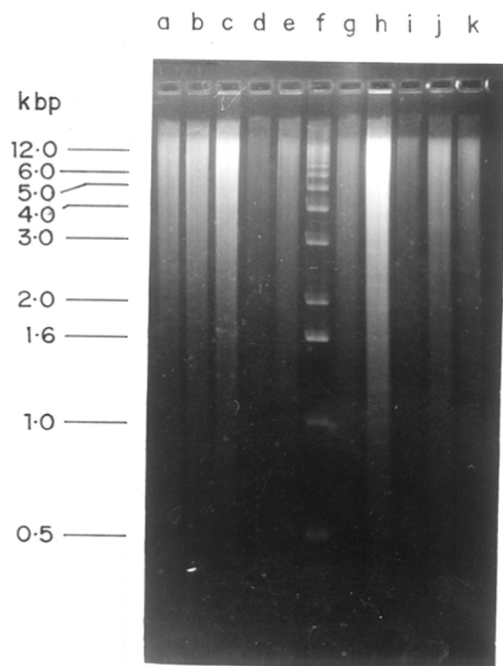


FIG. 3.15A

Fig 3.15 B

Agarose gel electrophoresis pattern of DNAs from 19 rice cultivars digested with EcoRI.

Lane a : RP4-14, Lane b : Rasi, Lane c : VDN-7850,

Lane d : KJT-184, Lane e : Jaya,

Lane f : Lambda DNA 1kb ladder as molecular weight marker,

Lane g : RDN 185-2, Lane h : BG 380-2, Lane i : Jalgaon-5,

Lane j : VDN 3-51-18, Lane K : RTN 73, Lane l : Ratna,

Lane m : VDN-4288, Lane n : RTN-24, Lane o : RTN-711,

Lane p : RTN-1,

Lane q : Lambda DNA 1kb ladder as molecular weight marker.

Lane r : PLG 39-4-1, Lane s : Pusa-33, Lane t : ACK 5,,

Lane u : KJT-35-3.

Electrophoresis was carried out on 1.0% neutral agarose slab gels in TAE (1X) buffer (pH 8.0) at a constant current of 20 mA for 12-15 hrs.

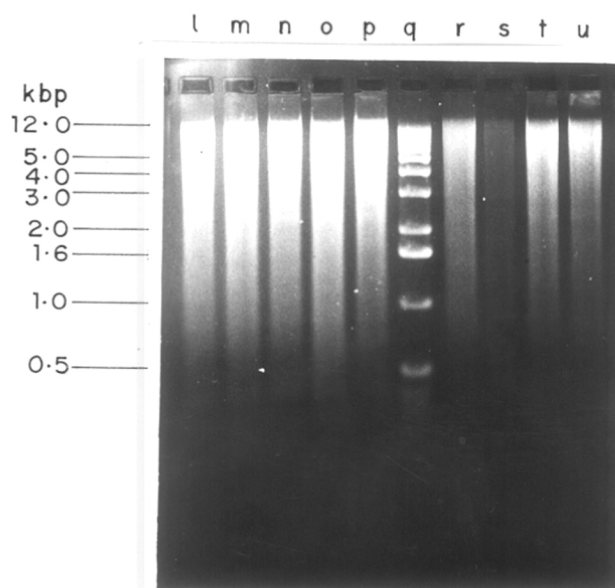
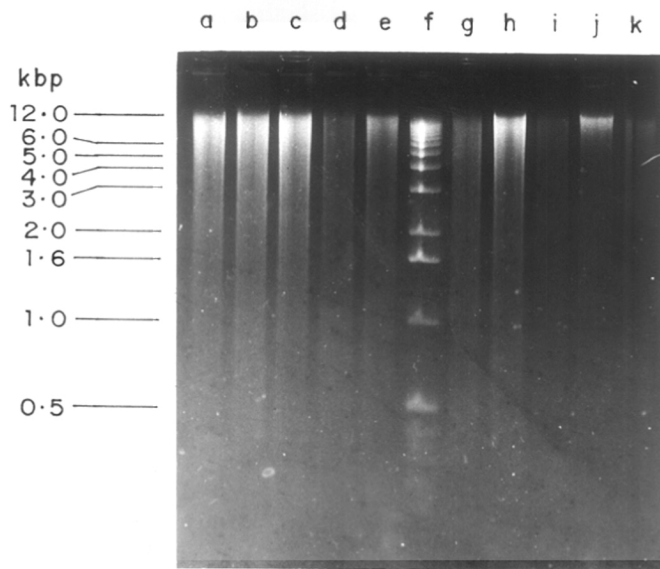


FIG. 3.15 B

Fig 3.15 C and D

Agarose gel electrophoresis pattern of DNAs from 19 rice cultivars digested with XbaI (C) and DraI (D) .

Lane a : RP4-14, Lane b : Rasi, Lane c : VDN-7850,

Lane d : KJT-184, Lane e : Jaya,

Lane f : Lambda DNA 1kb ladder as molecular weight marker,

Lane g : RDN 185-2, Lane h : BG 380-2, Lane i : Jalgaon-5,

Lane j : VDN 3-51-18, Lane k : RTN 73, Lane l : Ratna,

Lane m : VDN-4288, Lane n : RTN-24, Lane o : RTN-711,

Lane p : RTN-1,

Lane q : Lambda DNA 1kb ladder as molecular weight marker.

Lane r : PLG 39-4-1, Lane s : Pusa-33, Lane t : ACK 5,,

Lane u : KJT-35-3.

Electrophoresis was carried out on 1% neutral agarose slab gels, in TAE (1X) buffer pH 8.1, at a constant current of 20 mA for 12-15 hrs.

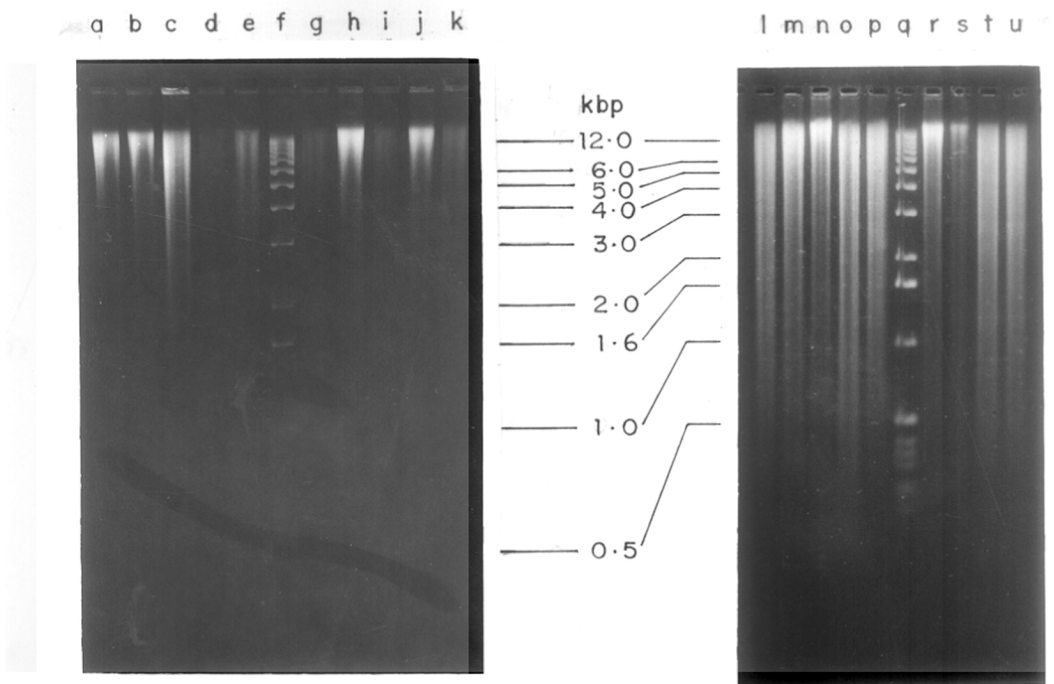


FIG. 3-15C

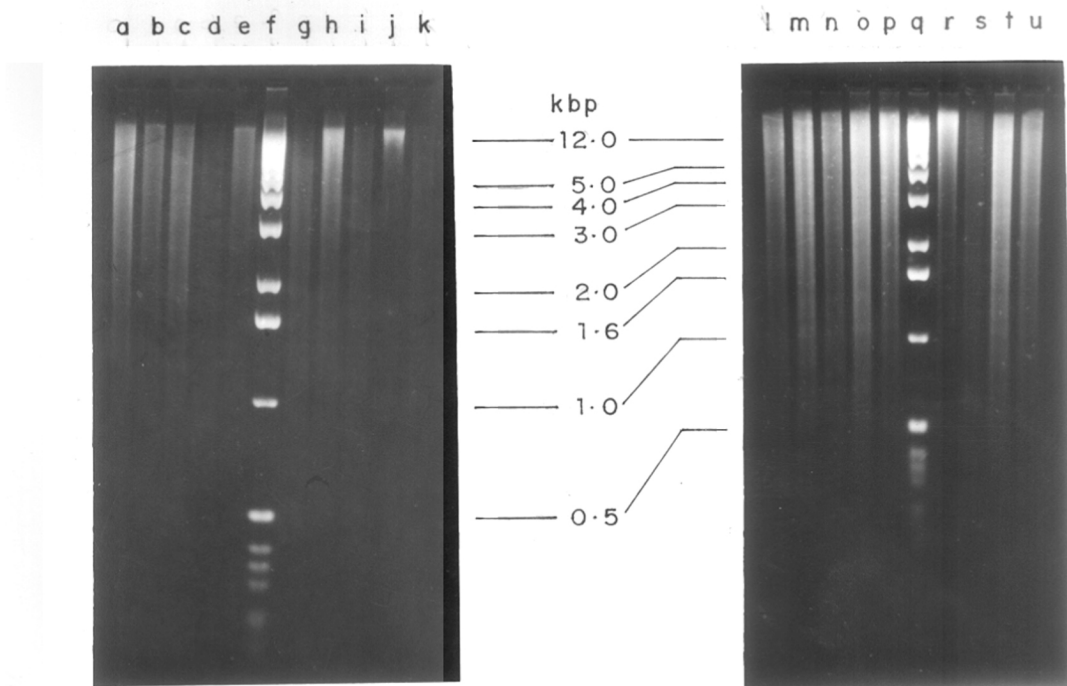


FIG. 3-15 D

rice are G+C rich the latter are expected to remain in the high molecular weight region in the digests.

Agarose gel electrophoresis pattern of ScaI digest (Fig.3.15 E) reveals digestion with fragment sizes ranging from 100 bp to 20 kbp with all the DNAs except cultivars like Rasi (lane b), Jalgaon-5 (lane i), and RTN-73 (lane k) which show comparatively less digestion leaving a smear of fragments mostly in the high molecular weight region. Ethidium bromide staining also reveals a good smear of fragments in the region below 2 kbp in the DNAs of several cultivars. The restriction enzyme Hind III digests most of the DNAs with the exception of cultivars like Jalgaon 5 (lane i), VDN-3-51-18 (lane j), RTN-73 (lane k), wherein the digestion with this enzyme leaves a high molecular weight smear (Fig. 3.15 F).

3.3.2 Analysis of hybridization patterns of P-32-labelled pOSM1C-2 insert with the digests of DNAs of 19 cultivars

In order to find out the potential of pOSM1C-2 in assessing the genetic/genomic variability amongst indica cultivars, RFLP probing was done using pOSM1C-2 repeat sequence. This probe was of choice for analysis because (i) it has an inserted rice DNA frag

Fig 3.15 E and F

Agarose gel electrophoresis pattern of DNAs from 19 rice cultivars digested to completion with ScaI (E) and Hind III (F).

Lane a : RP4-14, Lane b : Rasi, Lane c : VDN-7850,
Lane d : KJT-184, Lane e : Jaya,
Lane f : Lambda HindIII digest as molecular weight marker,
Lane g : RDN 185-2, Lane h : BG 380-2, Lane i : Jalgaon-5,
Lane j : VDN 3-51-18, Lane k : RTN 73, Lane l : Ratna,
Lane m : VDN-4288, Lane n : RTN-24, Lane o : RTN-711,
Lane p : RTN-1,
Lane q : Lambda HindIII digest as molecular weight marker,
Lane r : PLG 39-4-1, Lane s : Pusa-33, Lane t : ACK 5,,
Lane u : KJT-35-3.

Electrophoresis was carried out on 1% neutral agarose slab gels, in TAE (1X) buffer pH 8.1, at a constant current of 20 mA for 12-15 hrs.

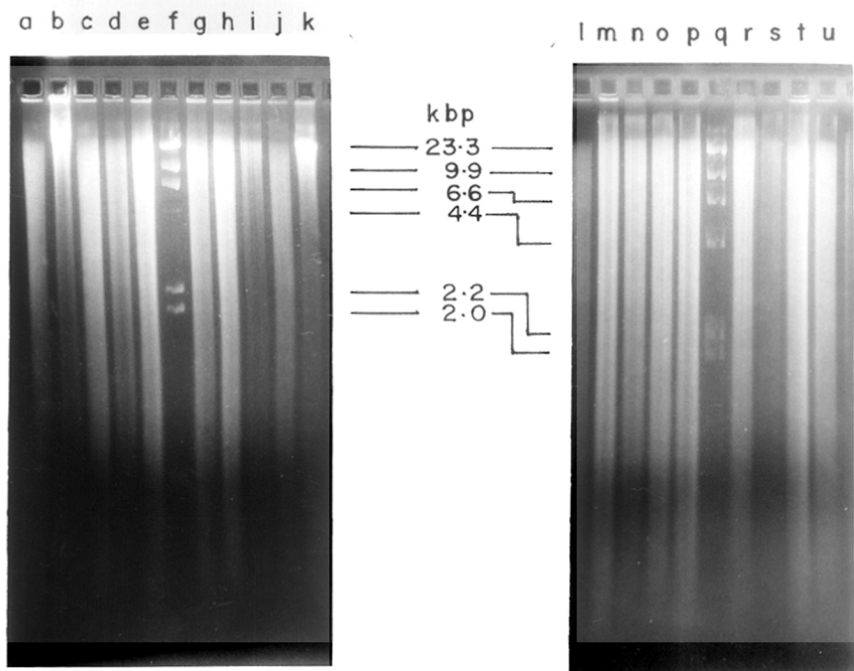


FIG. 3-15 E

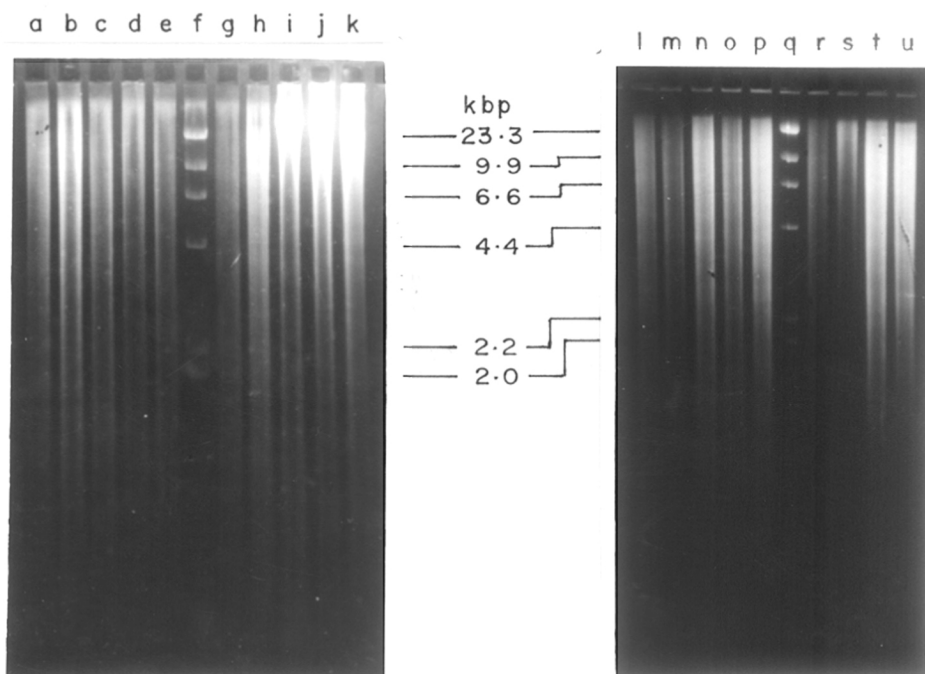


FIG. 3-15 F

ment of 1.55 kbp, a size useful for RFLP analysis. and (ii) it is a dispersed repeat. In rice, there are no reports so far where dispersed repeat has been used for analysis of genetic/genomic variability .

I have in the following paragraphs discussed the pattern of hybridization of P-32-labelled pOSM1C-2 insert with the digests of the DNAs of 19 cultivars with 6 restriction endonucleases. Every DNA fragment detected by Southern analysis when probed with pOSM1C-2 insert was treated as an independent unit. On the basis of RFLP data, the Nei's F values for different combinations of accessions were calculated (Nei, 1979). Details regarding Nei's formula are given in Materials and Methods (Section 2.12). A computer generated dendrogram was then constructed by the UPGMA method using the program written by us.

The pattern of hybridization of P-32-labelled pOSM1C-2 insert with the EcoRV digests of DNAs of 19 cultivars is shown in Fig.3.16 A. It is clearly seen that the repeat hybridizes giving one very prominent band (4.32 kbp) in all the cultivars. The intensity of this band, however, shows a distinct variation. The band, for example, is intense in cultivars like Rasi (lane b), VDN 7850 (lane c), BG 380-2 (lane h), RTN-1 (lane p), ACK-5 (lane t) and KJT-35-3 (lane u) ,

Fig.3.16 A

Southern blot hybridization of P-32-labelled pOSM1C-2 insert with DNA digests of rice cultivars with EcoRV. Lanewise order of the cultivars is the same as described in Fig. 3.15 A.

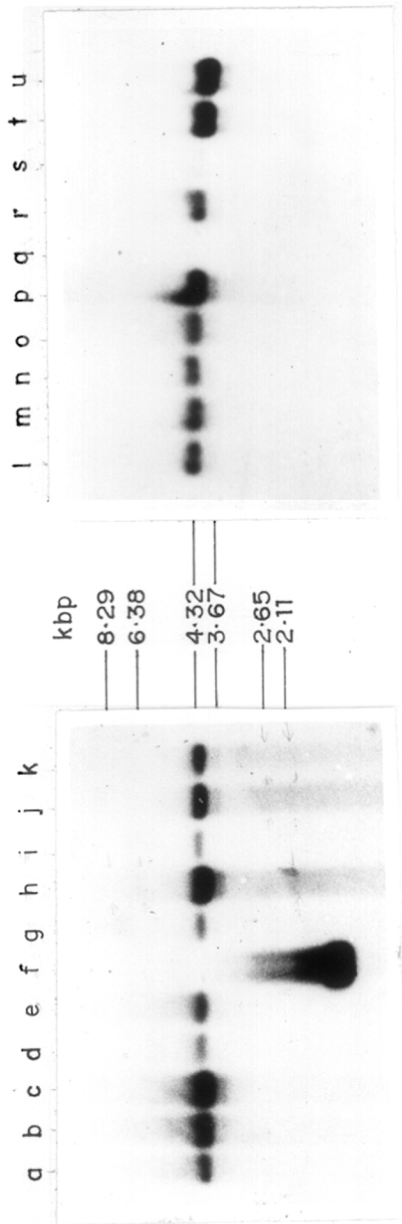


FIG. 3.16A

distinct in RP 4-14 (lane a), Jaya (lane e), VDN 3-51-18 (lane j), RTN 73 (lane k), Ratna (lane l), VDN 4288 (lane m), RTN 24 (lane n), RTN 711 (lane o) and PLG 39-4-1 (lane r) and faint in cultivars like KJT 184 (lane d), RDN 185-2 (lane g) and Jalgaon-5 (lane i). In addition to the band at 4.32 kbp, another very faint band is present in VDN-7850 (lane c), BG 380-2 (lane h), VDN 3-51-18 (lane j), RTN-73 (lane k) and RTN-1 (lane p) at a fragment length of 3.67 kbp.

In the EcoRI autoradiogram (Fig.3.16 B) the DNAs of all the cultivars show one high molecular weight hybridizing band at 9.9 kbp and two bands at 400 bp and 238 bp (not clearly seen in the figure). While characterizing the repeat sequence, it has been found that the repeat has two subrepeats namely LSR & SSR (Section 3.2.5). The high molecular weight band occurs due to both the subrepeats (namely LSR and SSR) while the two low molecular weight hybridizing bands occur solely due to the SSR. From this it is clear that LSR (PstI-AccI fragment 1.1 kbp) and SSR (AccI-PstI fragment 0.450 kbp) are two independent repeat units present in same as well as in different environment. The three, one major band (9.90 kbp) and two minor bands (400 bp and 238 bp) may be due to differential organization of these subrepeats.

Fig 3.16 B

Southern blot hybridization of P-32-labelled pOSM1C-2 insert with DNA digests of rice cultivars with EcoRI. Lanewise order of the cultivars is the same as described in Fig. 3.15 B.

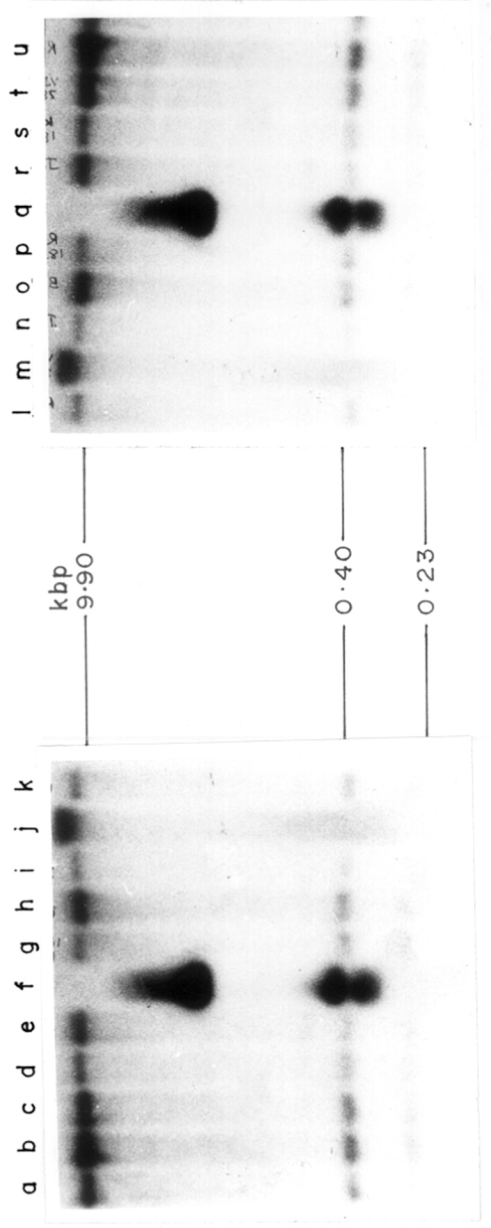


FIG. 3.16 B

The pattern of hybridization of P-32-labelled posM1C-2 insert with the XbaI DNA digests is shown in Fig.3.16 C. In case where there is a fairly good digestion of the varieties with XbaI, a single prominent hybridizing band at 7.30 kbp is observed. Examples are accessions like VDN-7850 (lane c), KJT-184 (lane d), Jaya (lane e), RDN-185-2 (lane g), Jalgaon-5 (lane i), RTN-73 (lane k), Ratna (lane l), VDN 4288 (lane m), RTN 24 (lane n), RTN-711 (lane o) and ACK 5 (lane t). Such hybridization of the probe at high molecular weight region (7.30 kbp) suggests that XbaI sites are far away in the region of genomic DNA homologous to the probe. For those DNAs which are digested to a lesser extent, probe hybridizes giving 2-3 bands in high molecular weight region (7.98 kbp, 9.10 kbp and 10.35 kbp). Examples are RP 4-14 (lane a), Rasi (lane b), BG 380-2 (lane h), VDN 3-51-18 (lane j), and KJT 35-3 (lane u) .

In the DraI autoradiogram (Fig. 3.16 D) most of cultivars show presence of 5 hybridizing bands of molecular weights 9.06 kbp, 8.30 kbp, 7.83 kbp, 6.02 kbp and 1.76 kbp. These include cultivars like RP 4-14 (lane a), Rasi (lane b), Jaya (lane e), BG380-2 (lane h), RTN- 73 (lane k), Ratna (lane l), VDN 4288 (lane m), Ratna-24 (lane n), RTN-711 (lane o), RTN-1 (lane p), ACK-5 (lane t) and KJT 35-3 (lane u). A few culti

Fig 3.16 C

Southern blot hybridization of P-32- labelled pOSM1C-2 insert with DNA digests of rice cultivars with XbaI.

Lanewise order of the cultivars is the same as described in Fig. 3.15 C.

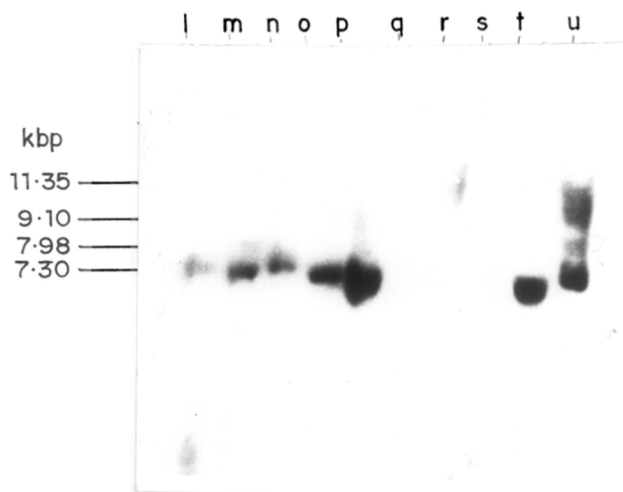
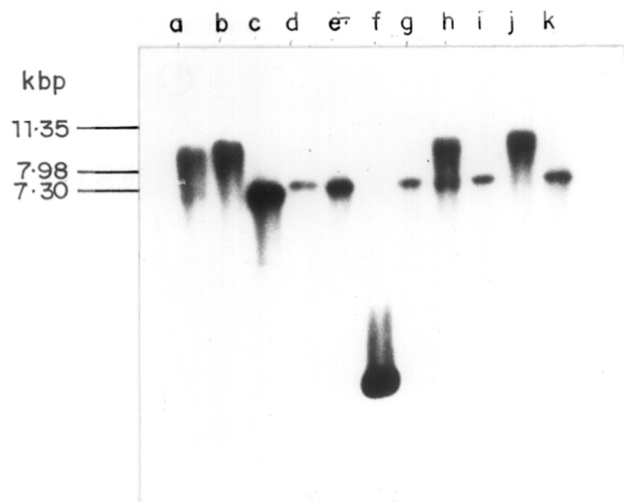


FIG. 3.16 C

Fig.3.16 D

Southern blot hybridization of P-32-labelled pOSM1C-2 insert with DNA digests of rice cultivars with DraI. Lanewise order of the cultivars is the same as described in Fig. 3.15 D.

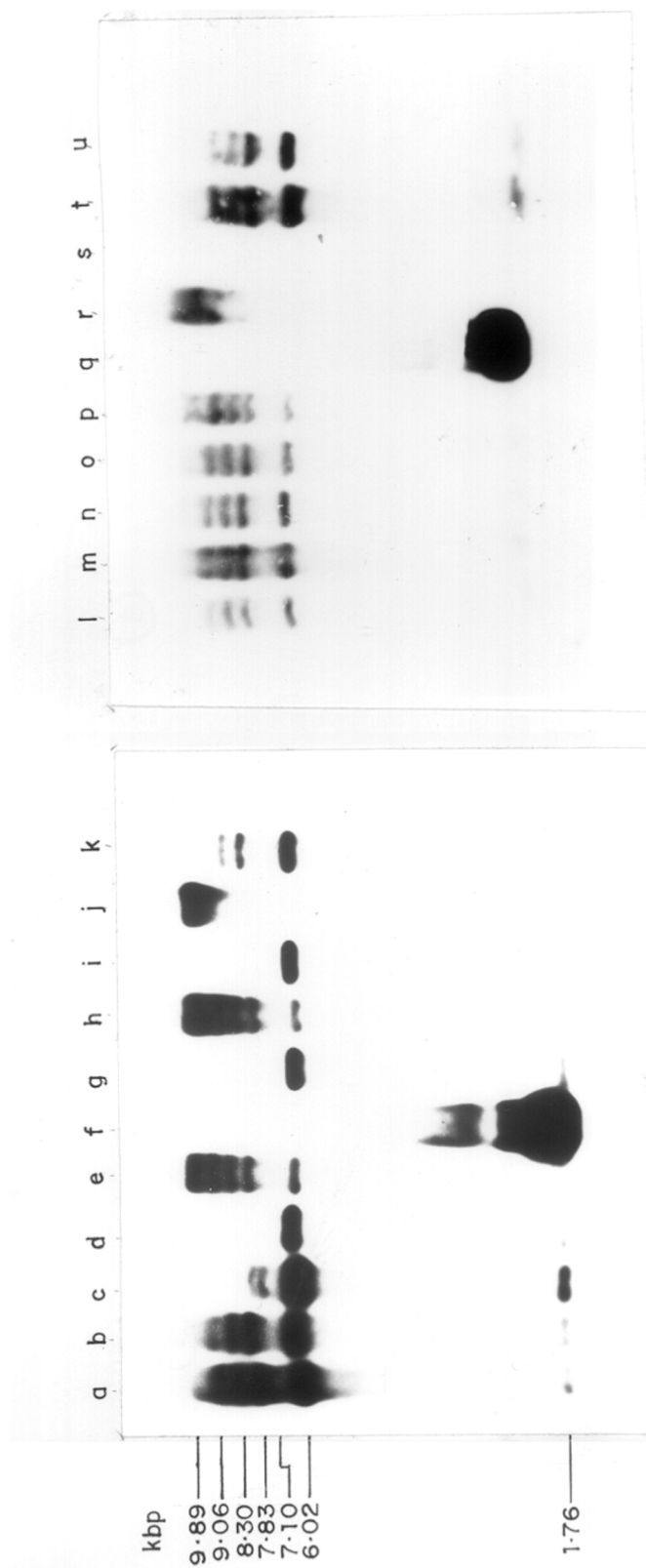


FIG. 3.16 D

vars like Jaya (lane e) and BG 380-2 (lane h) reveal an additional high molecular weight hybridizing band at 9.89 kbp. A band at 7.1 kbp appears only in VDN-7850 (lane c). Cultivars like VDN-7850 (lane c), KJT 184 (lane d), RDN 185-2 (lane g) and Jalgaon-5 (lane i) show absence of 9.06 kbp band. The repeat thus appears to be polymorphic with DraI.

The pattern of hybridization of P-32-labelled pOSM1C-2 insert with Sca I DNA digests is shown in Fig. 3.16 E. Repeat shows polymorphism with presence to absence pattern of bands in the 18 cultivars. A prominent hybridizing band at 10.54 kbp occurs in all the 18 cultivars. A band at 11.60 kbp however shows presence to absence pattern. The low molecular weight band at 1.54 kbp is present only in KJT-184 (lane d), Jaya (lane e), VDN 3-51-18 (lane j), Ratna (lane l), VDN 4288 (lane m), RTN 24 (lane n), RTN 711 (lane o), RTN-1 (lane p), PLG 39-4-1 (lane r), ACK 5 (lane t) and KJT 35-3 (lane u).

The autoradiography pattern of pOSM1C-2 with Hind III was not taken for the analysis of genomic variability, since the pattern in this case was inconclusive.

Fig.3.16 E

Southern blot hybridization of P-32- labelled pOSM1C-2 insert with DNA digests of rice cultivars with ScaI. Lanewise order of the cultivars is the same as described in Fig. 3.15 E.

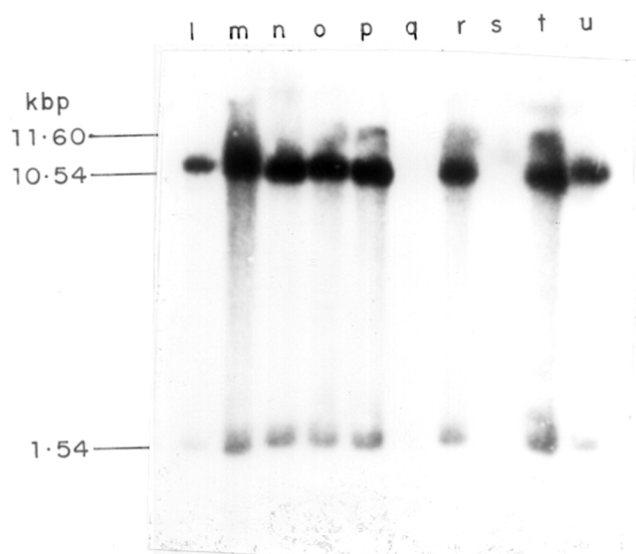
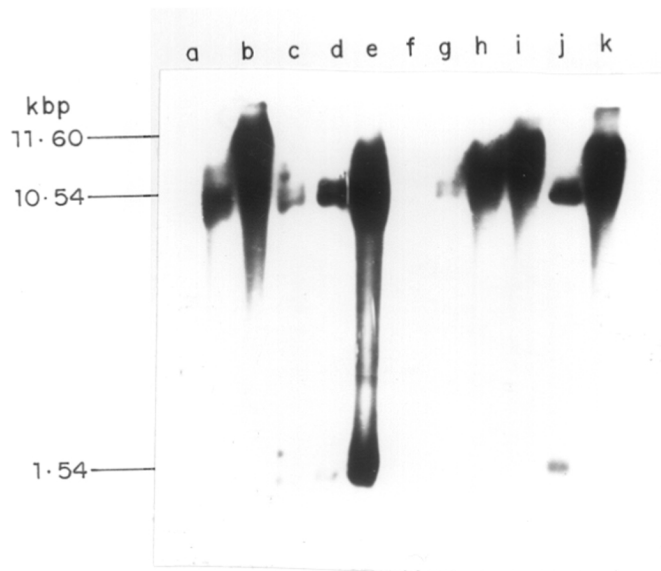


FIG. 3.16E

3.3.3 Analysis of the dendrogram generated based on Nei's F values for polymorphic enzyme-probe combinations

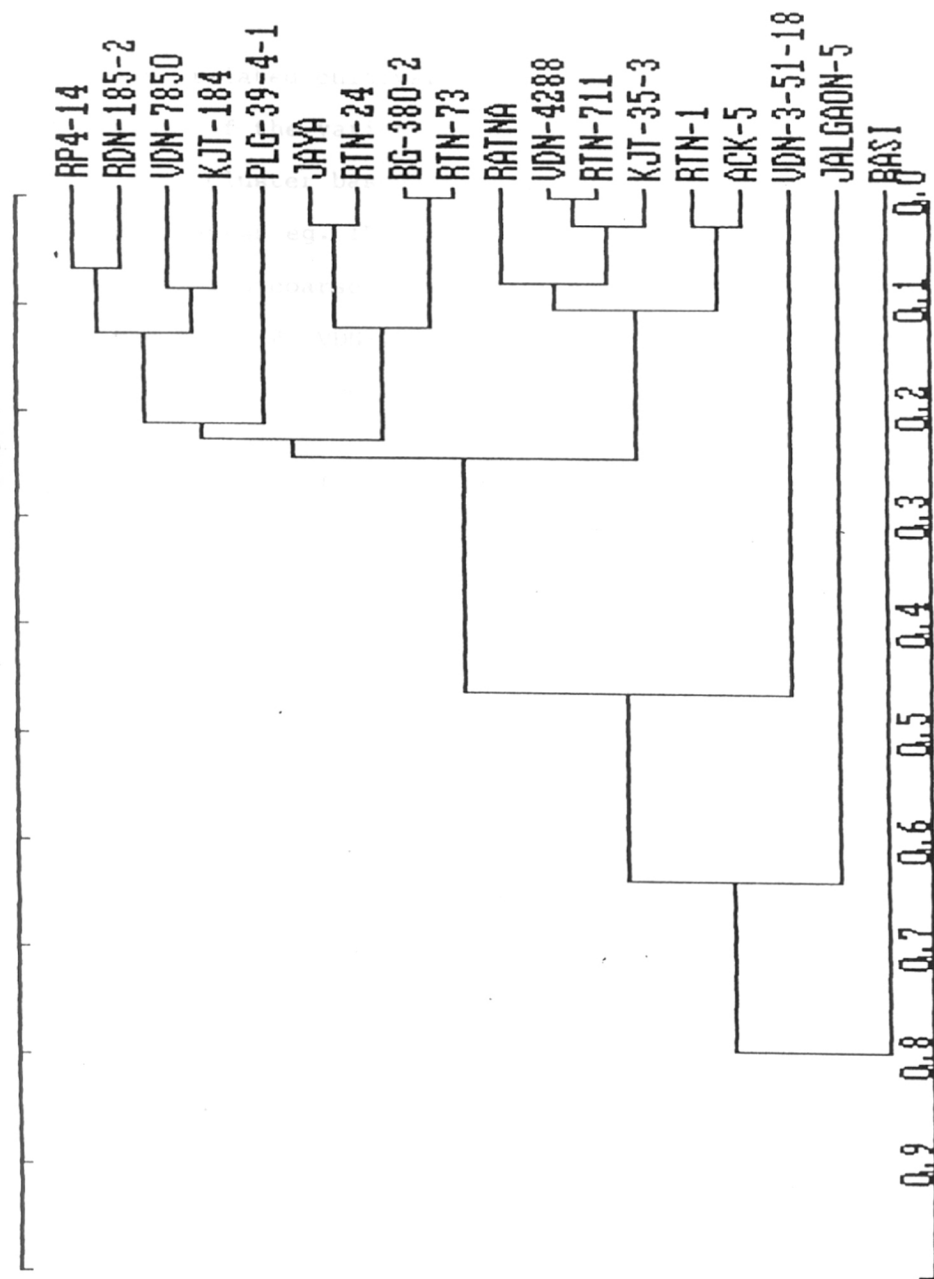
A computer generated dendrogram constructed using 3 enzymes (namely EcoRV, ScaI and DraI) and pOSM1C-2 as probe for 18 cultivars is shown in Fig. 3.17. One cultivar, namely, Pusa-33 was excluded all the time for the analysis of variability because autoradiography pattern with this cultivar was inconclusive. The dendrogram based on Nei's F values revealed two major clusters. Cultivars like Ack-5, RTN-1, KJT 35-3, RTN-711, VDN-4288 form one cluster, while cultivars like RP 4-14, RDN 185-2, VDN 7850, and KJT 184 form the other cluster. Interestingly, the analysis of dendrogram reveals that the grouping of cultivars is based on their pedigree. In other words, most of the cultivars from the former cluster have IR 8 as the common parent in their pedigree, while the latter cluster has cultivars having TN (1) as the common parent in their pedigree. This observation is in agreement with the fact that the breeding programs in India have been mostly based on a narrow genetic base comprising IR 8 and TN (1) (12). The IR 8 cluster and TN (1) cluster are further grouped into two sub-clusters. The resolution as seen from this analysis is high

Fig 3.17

A computer generated dendrogram of 18 varieties of Oryza sativa by RFLP analysis.

1-F values are used for the construction of the dendrogram. Hence herein the value 0 corresponds to more homology instead of 1 due to 1-F values instead of F values. The scale inversion facilitates better analysis of the interrelationships.

FIG. 3-17



enough to bring out subtle differences even amongst closely related cultivars. For example, when the attributes of the varieties are analysed, it is found that they cluster based on grain quality i.e. fineness and coarseness eg. IR 8 cluster has Jaya, BG 380-2 and RTN-73 having coarse grain quality, while TN (1) cluster has RP 4-14, VDN-7850 and KJT-184 having fine grain quality.

References

1. Deshpande, V.G. and Ranjekar, P.K. (1980)
Hoppe-Seyler's Z. Physiol. Chem., 361: 1223-33.
2. Nelson, M., McClelland, M. (1987)
Nucleic Acids Res., 15 Suppl: r219-230.
3. Pental, D., and Barnes, S.R. (1985)
Theor. Appl. Genet., 70: 185-191.
4. Kikuchi, S., Takaiwa, F. and Oono K. (1987)
Mol. Gen. Genet., 210: 373-380.
5. Jelinek, W.R. and Schmid, C.W. (1982)
Ann. Rev. Biochem., 51: 813-844.
6. Koukalova, B., Reich, J., Matyasek, R., Kuhrova, V.
and Bezdek, M. (1989)
Theor. Appl. Genet., 78: 77-80.
7. Bennett, M.D., Smith, J.B. and Heslop-Harrison J.C.
(1982)
Proc. R. Soc. Lond. B, 216: 179-99.
8. Dhar, M.S., Pethe, V.V., Gupta, and Ranjekar, P.K.
(1990)
Theor. Appl. Genet., 80: 402-408.
9. Vedel, F. and Delseny, M. (1987)
Plant Physiol. Biochem., 25(2): 191-210.
10. Kochert, G. (1989)
In: Introduction to RFLP mapping and Plant Breeding
applications.

The Rockefeller Foundation, International Program on Rice Biotechnology; pp 1-4.

11. Wang, Z.Y. and Tanksley, S.D. (1989)
Genome, 32:1113-1118.
12. Krishnamurthy, A: High yielding rice varieties of India. Central Rice Research Institute Trenchinca
Bulletin No. 1, India, (1987)

CHAPTER FOUR

DISCUSSION

4.1 Present status of repeats in plants

In the recent years, repeated DNA sequences are proving themselves to be very promising tools and lucrative targets in the areas of molecular biology and plant breeding . They have attracted a lot of attention in view of their role particularly in conferring species-specific and genome-specific characters (1-11), in possessing transposable element like properties (12-17) and as potential RFLP markers (18-24). For example, repeated DNA sequences isolated from P. hybrida have been used as probes to distinguish species with 14 chromosomes from those with 18 chromosomes (7). In addition, three repeat sequences are able to distinguish white-flowered species from pink flowered species . In oats (*Avena*), a cloned repetitive DNA sequence has been used to differentiate the 'C' genome bearing species from the other genomes namely, A and D (8). Here, the sequence RS-1 is abundant in 'C' genome *Avena* species, but is present at much reduced levels in A and D genomes. Species-specific repeated sequences have been used to detect and screen B. vulgaris having addition chromosome of B. procumbens (25). Recently, the potential of dispersed repeated sequences in genetic analysis has been investigated because of their ability to define many genetic loci (26). However, this potential has

largely remained unexploited, except for the example of hypervariable human minisatellite sequences which are used to finger-print rice cultivars (27,28,29). The human minisatellite family consists of many dispersed arrays of short (10-50 bp) tandem direct repeats that contain variants of a common 'core' sequences. The arrays exhibit a high degree of length variability probably due to the changes in the number of tandem repeats. These polycore probes namely 33.6 and 33.15 have been chosen due to their ability to detect polymorphic regions in vertebrate genomes. When these probes are challenged against the digests of O. sativa and O. glaberrima cultivars, the pattern of hybridization is found to be cultivar specific. Moreover, this pattern remains largely unchanged after regeneration of plants from tissue culture (29).

Recently, Lee et al. (26) have described the isolation of a pDR1 repeat sequence in pea and have shown its use to follow segregation of many loci simultaneously in conventional RFLP analysis. pDR1 is present in a region between two legumin genes, leg J and leg K in pea, and contains a dispersed repeated sequence of 5 kbp. This sequence has a structural homology to copia and Ty elements and is, therefore, related to the retrovirus like transposable element.

Levy et al. (30) have demonstrated the use of a repeated DNA sequence both as a population and phylogenetic marker in M. grisea. Rice blast fungus (Pyricularia grisea Sacc.) is economically the most important pathogen and is composed of clonal isolates that infect rice to cause the devastating rice blast disease. Hamer et al. (31) have identified a family of dispersed repetitive DNA sequences, called MGR, that is diagnostically conserved in M. grisea rice pathogen genomes. Several features of MGR sequences suggested their use in population level genetic analysis. For example, they occur in approximately 50 copies per haploid genome and were dispersed among all chromosomes. Blot hybridization experiment using pCB 586 as probe yields mitotically stable and isolate-specific EcoRI RFLP profiles, which could consequently serve as genotype specific finger-prints.

4.2 Highlights of work on repeated DNA sequences in rice in my laboratory and my own research

My laboratory has been engaged in the analysis of molecular aspects of rice genome since last 15 years. Following are some of the highlights of rice genome organization and repetitive DNA sequences.

[1] Repetitive DNA content in rice is 52% of the total genome (32).

[2] Rice shows a novel DNA sequence organization, where the length of uninterspersed repeated DNA sequence is 6.5 kbp (33).

[3] Further analysis reveals absence of interspersion in rice at a DNA fragment length as high as 20 kbp (34).

[4] The repeat sequences have high thermal stability (34).

[5] There is predominance of adenine methylation which is tissue specific in three indica rices, namely, Basmati-370, IR-24 and Sona (35).

[6] There is a modulation of a repeat sequence (pRL7) in terms of adenine methylation during transition from embryo to shoot stages in rice (35,36).

In view of the above research work on rice repeated DNA sequences in my laboratory, I was interested in characterizing rice genome in greater details. During my work, I hit upon a random genomic clone, namely pOSM1C-2 from the rice partial genomic library. Interestingly, the repeat turned out to be dispersed throughout the rice genome. This prompted me to characterize this sequence in some details. Further, I studied the methylation status of this sequence at three different stages of development in rice, namely, embryo, seedling (8-10 days old) and leaf (3-4 weeks

old). Lastly, it's dispersed nature led me to assess it's role as a molecular marker in distinguishing different rice varieties belonging to the indica sub-species.

4.3 Why construct a partial PstI rice genomic library?

Low level of "C" methylation in rice repeats

For the assessment of genomic/genetic variability between different parental lines, genomic or cDNA libraries have been constructed in a number of plant species as source of candidate probes. Detection of the most diverse parents is a pre-requisite for construction of genetic linkage maps based on RFLPs. After detecting an informative probe-enzyme combination for a pair of parental lines of interest, and using a segregating population therein, genetic linkage maps have been or are being, constructed for monocots like rice (37,38), maize (39,40), wheat & barley (41) and dicots like tomato (42), lentil (43), Brassica (44), lettuce (45), soybean (46) and potato (47). The libraries at large have been constructed using two enzymes, namely, PstI & EcoRI. PstI, a methylation sensitive enzyme recognizes the sequence 5'-CTGCAC-3' and is sensitive to 'C' methylation in the 5' location (48).

Thus clones represented in the library are likely to be from the regions of the genome which are not methylated at this site. Such undermethylated sequences are enriched for single copy sequences, possibly because they represent coding regions (40). Recently, McCouch *et al.* (37) have constructed a PstI genomic library of rice O. sativa cv. IR36, in pUC8 . A total of 300 genomic clones have been surveyed for copy number and polymorphism by probing them on to the filter containing total DNA from IR 34583 and Bulu Dalam, that have been digested with 11 different restriction enzymes. The clones are categorized as follows:

- [1] Single copy sequences (SC) , where more than 90% of the signal in the lane is accounted by one or two discrete bands.
- [2] Multiple copy sequences (MC), where 3 or more bands are present and no single band accounts for less than 20% or more than 80% of the signal and
- [3] Repeated sequences which appear as smear, but no discrete bands.

The estimated single copy fraction from their library is 58% (37).

It is interesting to compare the percentage of single copy versus multicopy probes generated by

PstI and EcoRI. McCouch *et al.* (37) have already made such a comparison between their rice library and two other libraries of tomato and maize. Rice and tomato have a small genome size (0.6 pg - 1.0 pg) (49) and a haploid complement of 12 chromosomes. In tomato, 40% of the clones derived from the EcoRI library are estimated to contain unique sequences (50,51). However, in the PstI library of tomato, single copy fraction of the clones rises to 92% (51). Such a high frequency (>75%) of low-copy-number probes are also observed in the PstI partial genomic libraries of Brassica species and maize (52,39). In rice, PstI libraries of two cultivars namely Malkolam and IR-36, reveal a similar trend wherein the fraction of single copy clones significantly decreases suggesting that PstI library in rice is less biased towards single copy sequences than in tomato, maize and Brassica. A comparison of cutting efficiencies of 'C' methylation sensitive and insensitive restriction endonucleases on rice and tomato DNA indicates that rice DNA is less methylated than tomato DNA, though both plant species have same 'C' value (37). In tomato, most of the repeated sequences are hypermethylated whereas single copy sequences are undermethylated. In rice, repeated sequences appear to have low level of methylation. Simi-

lar results have been reported in lentil (Lens culinaris Medik), a self-pollinating diploid ($2n=2x=14$) grain legume, wherein PstI and EcoRI partial genomic libraries give approximately equal frequencies of low-copy-number probes (41% and 38% respectively) (53). These results indicate no obvious advantage in generating probe libraries from hypomethylated regions in rice and lentil. Hence, a PstI library of rice could be equally efficient in generating repetitive probes as any other library like EcoRI.

In summary, repeated sequences in rice have low level of "C" methylation. In spite of using a methylation sensitive enzyme like PstI for the construction of the library, nearly equal number of repetitive and low copy number sequences are generated. As stated above, PstI thus is less biased towards single copy sequences in rice.

4.4 Random organization of pOSM1C-2 sequence in rice genome

In the present work, I have isolated and characterized one repeat sequence namely pOSM1C-2 which was randomly selected from rice, O. sativa ssp. indica cv. Malkolam. The first most interesting observation of pOSM1C-2 sequence is that it typically

shows a "clustered and scrambled " organization in the genome. In such an arrangement, the PstI repeat can be found in clusters of repeated sequences that are scrambled in different genomic locations. Thus at certain locations in the genome, it would have one set of neighbouring sequences, while at the other locations, the flanking sequences would themselves be repetitive in nature. The dispersed repeat pOSM1C-2 may also be present elsewhere in the genome as part of longer/larger repeating units (tandem or dispersed). Evidence for this comes directly from the genomic blot hybridization pattern of pOSM1C-2 with rice DNA digested with different restriction enzymes. The finding of a 'ladder' in any of the enzyme digest indicates that the repeat is also present as a tandem array. Our blot hybridization experiments reveal absence of such pattern indicating that the repeat sequence is present as part of larger dispersed repeat units. Thus, a small dispersed repeat unit (1.55 kbp) may also occur as part of longer (full length) repeat units elsewhere in the genome. It is thus reasonable to assume that the PstI fragment, cloned in pUC18, is just internal parts of much longer units in the genome. Evans and Barnes (54) have reported a B1000 sequence (BamHI repeat, 1 kb in

length, from Allium sativum), that is repeated upto 10^6 times in the genome and probably forms a part of longer repeating units (54). Similar pattern of organization of dispersed repeat elements (about 300-1300 bp in length) is reported in maize wherein some small repeat elements exist as part of larger repeating units elsewhere in the genome (55).

4.5 Differential organization of two subrepeats of pOSM1C-2: Probable role as transposable elements

Another interesting structural feature of pOSM1C-2 is presence of two subrepeats, LSR (large subrepeat) and SSR (short subrepeat) occurring as a part of one single repeat of 1.55 kbp at most of the times, while SSR showing an additional independent organization in the genome. Rice is known to have long (4-20 kbp) and short (300-400 bp) repeats interspersed amongst themselves as well as with each other (34). This implies that the LSR & SSR adjacent to each other (1.55 kbp single repeat) may be a part of long repeats alone or may be parts of long and short repeats respectively which are interspersed with each other. The SSR, in addition to this, may be existing in different environment in rice genome. Hence, the sequence pOSM1C-2

could be referred to as a "junction or a border" sequence of two internal repeat units having differential organization and frequencies of reiteration in the rice genome. Such repeats have been reported earlier in plants like Zea mays and Lilium henryi (55,56). In maize, for instance, some small repeats, 300 bp in length exist as a part of a larger repeating units elsewhere in the genome. In lily, although the two repeat sequences (2 kbp and 5 kbp) are usually adjacent, with occasional absence of internal BamHI site, they result in the 7 kbp fragment. A smear of homologous sequences overlaying the 2-, 5- and 7 kbp bands occurs due to subregions of BamHI family being interspersed elsewhere in the genome (56).

It is further suggested that the BamHI family elements of lily have been amplified and dispersed due to transposition (56). The repeated DNA sequences/families dispersed in the cereal genomes presumably result from permanent genome rearrangements, mainly due to amplification and translocation of individual DNA segments (57). It is suggested that mobile genetic elements take part in most of such genome rearrangement events. In other words, transposition events are common for the occurrence of dispersion of the sequences in the genome. Dialect-1, a species-

specific repetitive DNA sequence of barley Hordeum vulgare, is found to be dispersed through all barley chromosomes with a copy number of 5000 per genome (58). The involvement of transposable elements in the D-1 H.v. repeat is indicated by its species-specificity and by existence of varying forms of related sequences in different barley species. It is also postulated that the duplication or amplification processes occur quite frequently after dispersion process. Based on the dispersed nature and differential organization of the internal subrepeats of pOSM1C-2, it is likely that similar transposition events might have also occurred in case of rice. This is probably the first report to our knowledge where a repeat sequence of this kind has been identified and partially characterized from a small genome plant species like rice.

4.6 Predominance and stage specific adenine methylation in rice genome

It is now generally accepted that DNA methylation plays a major role in the regulation of gene expression especially during development. Transcriptional regulation is emerging as a complicated multi-level process probably involving the interactions of a large number of trans-acting factors with specific

cis-acting elements. One controlling element in this scheme is the local configuration of a gene which may determine its accessibility to the various cellular factors required for transcription. DNA methylation provides one such mechanism for stably altering the local structure of the gene and thus may play an important role in regulation of gene activity (59).

In my work, I have studied the methylation status of Malkolam rice genome with respect to pOSM1C-2 sequence at three different stages of development, namely embryo, shoot and leaf. The DNAs of these tissues were digested with 5 methylation sensitive enzymes namely MboI, Sau3AI, DpnI, MspI and HpaII and probed with P-32-labelled pOSM1C-2 insert.

The restriction endonuclease data using MboI, Sau3AI and DpnI indicates that the pattern of adenine methylation is tissue specific, i.e. there is predominance of 'A^m' in leaf than in shoot and in shoot as compared to embryo. A similar stage specific methylation pattern is also reported in three other varieties of rice, namely, Basmati-370, Sona and R-24 (35). This information about predominance and tissue specificity of adenine methylation may be of structural or functional significance in rice development.

In case of animals and microorganisms, methylated adenine has defined roles (60,61,62). One possible role for DNA methylation in vertebrate cell reported by Hare et al. (61) is strand discrimination in the repair of mismatches. In E. coli, methylated cytosines in the lac repressor spontaneously mutate to thymine at a rate many times higher than that of unmodified cytosine. When lac repressor DNA resides in the host that does not methylate cytosines, mutation rate at these sites returns to normal (63). Recently the role of 6-mA in DNA repair in plant cells has been envisaged by Pintor-Toro (64). Though zein genes are methylated at 'A', the experimental evidence suggests that it is unrelated to their regulation of expression. From all this information, it appears that functional significance of DNA methylation in higher plants still remains speculative.

4.7 Role of pOSM1C-2 as a RFLP marker

After a fairly detailed characterization of pOSM1C-2, my objective next on these lines was to assess if this sequence had a potential to decipher genetic/genomic variability among the selected rice germplasm. In other words, the aim here was to study whether pOSM1C-2 proved itself to be an informative

marker or a polymorphic probe. The knowledge of the amount and distribution of genetic variability among the cultivars (within and between species) is important to the optimal use of these genetic resources in plant breeding programs. In this study, RFLP analysis has been used to estimate the genomic interrelationships between the varieties of rice belonging to indica subspecies. Based on RFLP variation, I have constructed a dendrogram depicting genetic similarities among rice varieties.

The development of RFLP technology has opened a door to detecting, monitoring and manipulating genetic variation in plants in a manner not previously possible (65). RFLPs are the most sensitive tools for the detection of differences within or between species. One of the most ready applications of RFLP analysis is in assessing genetic variation in natural populations and exploring the evolutionary relationships among plant taxa (65). In fact, DNA markers have been popularly used as tools to measure genetic affinities/distances between species and varieties in a number of plants (66). The potential usefulness of RFLPs is, however, dependent on the degree of polymorphism existing among the organisms.

The studies on analysis of genetic variation using RFLP approach have been reported in a number of plant species. Some of the examples of monocot species are rice (65,67), wheat and barley (68) and dicots like tomato (69,70), soybean (71), lettuce (72), lentil (53) and Vicia (73). Recently, Wang and Tanksley (65) have used RFLP markers to study genetic variation among 70 accessions of O. sativa from three subspecies. Here, 10 single copy probes from rice IR36 library in pUC 8, and one phytochrome cDNA are hybridized with rice DNA digested with 5 different restriction enzymes for a total of 50 probe-enzyme combinations. On the basis of polymorphism, a dendrogram is constructed showing genetic interrelationships among the varieties. These studies have demonstrated the amount of variation in O. sativa and provided a clearer picture of evolutionary relationships between cultivated rice and its wild ancestors. From their dendrogram, conventional indica type varieties (tropical rice) cannot be grouped as a single unit, while all japonicas (temperate rice) do cluster into a single group. Several indicas were also in the japonica group.

Kangle et al. (67) have on the similar lines

studied RFLPs between cultivated and wild rice O. glumacpalata . On the basis of dendrogram, genotypes of japonica and indica were grouped respectively and it was found that indica accessions were closer relatives of wild rices than japonica. The japonica variety Ketan Nangta which showed wide compatibility was found to be closer to indica.

The RFLP analysis of genetic relatedness reveals that 18 rice varieties cluster based on their pedigree into two major clusters, having IR 8 and TN (1) as common parent in their pedigree. This observation is in agreement with the fact that breeding programs in India have been mostly based on a narrow genetic base comprising IR 8 and TN (1) (74). However, the resolution as seen from the subclustering by our RFLP analysis is high enough to bring out subtle differences even amongst closely related cultivars. When the attributes of the varieties were analysed, it is found that they cluster based on the grain quality ie. fineness and coarseness e.g. TN (1) cluster has RP 4-14, VDN 7850 and KJT 184 having fine grain quality, whereas Jaya, BG 380-2 and RTN 73 having coarse grain quality.

These findings suggest the potential use of the present data in designing breeding programs. The

breeding can be done to obtain two types of results, viz.(i) a wide spectrum of variation by resorting to inter-cluster crosses and (ii) a narrow spectrum of variation by resorting to intra-cluster crosses.

References

1. Saul, M.W. and Potrykus, I. (1984)
Pl. Cell Reports, 3:65-67.
2. Rayburn, A.L. and Gill, B.S. (1987)
Amer. J. Bot., 74:574-580.
3. Appels, R. and Moran, L.B. (1984)
Molecular Analysis of Alien Chromatin Introduced
into Wheat - In Gustafson, J.P., (Ed.):16th
Stadler Genetics Symposium, pp 529-557. New York:
Plenum. N.Y., Furlek, 1984.
4. McIntyre, C.L., Clarke, B.C. and Appels, R. (1988)
Pl. Syst. Evol., 160:39-59. 81:130.
5. Schweizer, G., Ganai, M., Ninnemann, H. and
Hemleben, V. (1988)
Theor. Appl. Genet., 75:679-684.
6. Metzlauff, M., Troebner, W., Baldauf, F., Schlegel,
R. and Cullum, J. (1986)
Theor. Appl. Genet., 72:207-210.
7. Shepherd, A.L., Anderson, S. and Smith, S.M.
(1990)
Plant Sci., 67:57-62.
8. Crowhurst, R.N. and Gardner, R.C. (1991)
Theor. Appl. Genet., 81:71-78.
9. Fabijanski, S., Fedak, G., Armstrong, K. and
Altosaar, I. (1990)
Theor. Appl. Genet., 79:1-7.
10. Rayburn, A.L. and Gill, B.S. (1986)

- Pl. Mol. Biol. Rept., 4:102-109.
11. Zhao, X., Wu, T., Xie, Y. and Wu, R. (1989)
Theor. Appl. Genet., 78:201-209.
 12. Vershinin, A.V., Salina, E.A., Svitashhev, S.K. and
Shumny, V.K. (1987)
Theor. Appl. Genet., 73:428-432.
 13. Baker, B., Schell, J., Lorz, H. and Fedoroff, N.
(1986)
Proc. Natl. Acad. Sci. (USA), 83:4844-4848.
 14. Fedoroff, N.V., Furtek, D.B. and Nelson, O.E.
(1984)
Proc. Natl. Acad. Sci. (USA), 81:3825-3829.
 15. Paz-Ares, J., Weinand, V., Peterson, P.A. and
Saedler, H. (1986)
EMBO J., 5:829-834.
 16. Martin, C., Carpenter, R.M., Sommer, M., Saedler,
H. and Coen, E.S. (1985)
EMBO J., 4:1625-1630.
 17. Vedel, F. and Delseny, M. (1987)
Pl. Physiol. Biochem., 25:191-210.
 18. Botstein, D., White, R., Skolnick, M. and Davis,
R.W. (1980)
Amer. J. Hum. Genet., 32:314-331.
 19. Tanksley, S.D. and Bernatsky, R. (1985)
Molecular markers in plant breeding - Crop Sci.
Amer. Argon Meeting Abstr., 25:136.

20. Beckmann, J.S. and Soller, M. (1983)
Theor. Appl. Genet., 67:35-43.
21. Landry, B.S., Kesseli, R.V., Farrara, B. and
Michelmore, R.W. (1987a)
Genetics, 116:331-337.
22. Landry, B.S., Leung, H. and Michelmore, R.W.
(1987b)
Theor. Appl. Genet., 74:646-653.
23. Helentjaris, T., King, G., Slocum, M., Sieden
strang, C. and Wegman, S. (1985)
Pl. Mol. Biol., 5: 109-118.
24. Evola, S.V., Burr, F.A. and Burr, B. (1986)
Theor. Appl. Genet., 71:765-771.
25. Schmidt, T., Junghans, H. and Metzlauff, M. (1990)
Theor. Appl. Genet., 79:177-181.
26. Lee, D., Ellis, T.H.N., Turner, L., Hellens, R.P.
and Cleary, W.G. (1990)
Pl. Mol. Biol., 15:707-722.
27. Jeffreys, A.J., Wilson, V. and Thein, S.L. (1985)
Nature, 314:67-73.
28. Jeffreys, A.J., Wilson, V. and Thein, S.L. (1985)
Nature, 316:76-79.
29. Dallas, J.F. (1988)
Proc. Natl. Acad. Sci. (USA), 85:6831-6835.
30. Levy, M., Ramao, J., Marchetti, M.A. and Hamer,
J.E. (1991)
The Plant Cell, 3:95-102.

31. Hamer, J.E., Farrall, L., Orbach, M.J., Valent, B. and Chumley, F. (1989)
Proc. Natl. Acad. Sci. (USA), 86:9981-9985.
32. Deshpande, V.G. and Ranjekar, P.K. (1980)
Hoppe-Seylers Z. Physiol. Chem., 361:1223-1233.
33. Gupta, V.S., Gadre, S.R. and Ranjekar, P.K. (1981)
Biochim. Biophys. Acta., 656:147-154.
34. Dhar, M.S., Dabak, M.M., Gupta, V.S. and Ranjekar, P.K. (1988)
Plant Sci., 55:43-52.
35. Dhar, M.S., Pethe, V.V., Gupta, V.S. and Ranjekar, P.K. (1990)
Theor. Appl. Genet., 80:402-408.
36. Gupta, V.S., Dhar, M.S., Patil, B.G., Narvekar, G. S., Rawat, S.R. and Ranjekar, P.K. (1990)
J. Biosci., 15(4):261-269.
37. McCouch, S.R., Kochert, G., Yu, Z.H., Wang, Z.Y., Khush, G.S., Coffman, W.R. and Tanksley, S.D. (1988).
Theor. Appl. Genet., 76: 815-829.
38. Tanksley, S., Ahn, N., Causse, M., Coffman, R., Fulton, T., McCouch, S., Second, G., Tai, T., Wang, Z., Wu. K. and Yu, Z. (1990)
RFLP mapping of rice genome. Fourth Annual International Program on Rice Biotechnology, held at IRRI, Philippines during May 9 - 12.

39. Helentjaris, T. (1987)
Trends in Genet., 3:217-221.
40. Burr, B., Burr, R.A., Thompson, K.H., Albertsen, M.C. and Stuber, C.W. (1988)
Genetics, 118:519-526.
41. Chao, S., Sharp, P.J., Kilian, A. and Gale, M.D. (1988)
RFLP based genetic maps of homeologous 7 chromosomes of wheat and barley. Plant Genetics and Breeding Research, Annual Report, pp 23-24.
42. Helentjaris, T., Slocum, M., Wright, S., Schaefer, A. and Nienhuis, J. (1986)
Theor. Appl. Genet., 72:761-769.
43. Havey, M.J. and Muehlbauer, F.J. (1989)
Theor. Appl. Genet., 77:395-401.
44. Slocum, M.K., Figdore, S.S., Kennard, W.C., Suzuki, J.Y. and Osborn, T.C. (1990)
Theor. Appl. Genet., 80:57-64.
45. Landry, B.S., Kesseli, R.V., Farrara, B. and Michelmore, R.W. (1987)
Genetics, 116:331-337.
46. Keim, P., Diers, B.W., Olson, T.C. and Shoemaker, R.C. (1989)
Genetics, 126:735-742.
47. Bonierbale, M.W., Plaisted, R.L. and Tanksley, S.D. (1988)
Genetics, 120:1095-1103

48. Nelson, M. and McClelland, M. (1987)
Nucleic Acids Res., 15:Suppl:r219-230.
49. Bennett, M. D. and Smith, J. B. (1976)
Philos. Trans. R. SOC> London, Ser. B 274: 227-274.
50. Tanksley, S.D., Miller, J., Paterson, A. and Bernatzky, R. (1987)
Molecular mapping of plant chromosomes. In: Gustafson J.P., Appels, R.A. (Eds) Chromosome structure and function. Plenum Press, New York, pp. 157-173.
51. Zamir, D. and Tanksley, S.D. (1988)
Mol. Gen. Genet., 213:254-261.
52. Figdore, S.S., Kennard, W.C. Song, K.M., Slocum, M.K. and Osborn, T.C. (1988)
Theor. Appl. Genet., 75:833-840
53. Havey, M.J. and Muehlbauer, F.J. (1989)
Theor. Appl. Genet., 77:839-843.
54. Evans, I.J., James, A.M. and Barnes, S.R. (1983)
J. Mol. Biol., 170:803-826.
55. Gupta, M., Shepherd, N.S., Bertram, I. and Saedler, H. (1984)
The EMBO J., 3 (1): 133-139.
56. SENTRY, J.W. and Smyth, D.R. (1985)
Chromosoma, 92:149-155.
57. Flavell, R.B. (1980)

Improvement

- Ann. Rev. Plant Physiol., 31:569-596.
58. Sonina, N.V., Lushnikova, A.A., Tihanov, A.P. and Ananiev, E.V. (1989)
Theor. Appl. Genet., 78:589-593.
59. Cedar, H. (1988)
Cell, 53:3-4.
60. Cheng, S., Gail, H. and Modrich, P. (1985)
J. Biol. Chem., 260:191-194.
61. Hare, J.T. and Taylor, J.H. (1985)
Proc. Natl. Acad. Sci. (USA), 82:7350-7354.
62. Pukkila, P.J., Peterson, J., Herman, G., Modrich, P. and Messelson, M. (1983)
Genetics, 104:571-582.
63. Miller, J. H. (1982)
Cell, 31:5-7.
64. Pintor-Toro, J.A. (1987)
Biochem. Biophys. Res. Comm., 147:1082-1087.
65. Wang, Z.Y. and Tanksley, S.D. (1989)
Genome, 32:1113-1118.
66. Tanksley, S.D., Young, N.D., Paterson, A.H. and Bonierbale, M.W. (1989)
Bio/Technology, 7:257-264.
67. Kangle, Z., Shen, B., Yu, F., Zhao, C., Qi, X., Xu, X. and Xiong, Z. (1989)
Restriction Fragment Length Polymorphism in Rice.
Second Annual Meeting on "Biotechnology Development for Rice Improvement in China", held at Beijing,

- during November, 24-25.
68. Liu, Y.G., Mori, N. and Tsunewaki, K. (1990)
Jap. J. Genet., 65 (5):367-380.
 69. Miller, J.C. and Tanksley, S.D. (1990)
Theor. Appl. Genet., 80:437-448.
 70. Miller, J.C. and Tanksley, S.D. (1990)
Theor. Appl. Genet., 80:385-389.
 71. Keim, P., Shoemaker, R.C. and Palmer, R.G. (1989)
Theor. Appl. Genet., 77:786-792.
 72. Landry, B.S., Kesseli, R., Leung, H. and
Michelmore, R.W. (1987)
Theor. Appl. Genet., 74:646-653
 73. M. van de Ven, Powell, W., Ramsay, G.
and Waugh, R. (1990)
Heredity, 65:329-342.
 74. Krishnamurthy, A: High yielding rice varieties
of India. Central Rice Research Institute
Technical Bulletin No 1, India (1987).

CURRICULUM VITAE

CURRICULUM VITAE

Name : Miss Vaijayanti V. Pethe
Current address : Division of Biochemical Sciences,
National Chemical Laboratory,
Pune 411 008. India.
NCL Tele fax : (0212) 330 233
Telephone no. : (212) 33 6451-52-53.
Permanent address : C/O Dr. V.P. Pethe
Dashabhuj Ganesh Co-operative
Society, Near Paud Fata, Karve Road,
Pune 411 008. India.
Telephone no. 330253.
Nationality : Indian.
Date of birth : 15th October 1963
Marital Status : Unmarried
Academic qualifications :

Degree	University	Year	Class	Subject
B.Sc	Poona, India	1983	I (79%)	Chemistry (Principal) Microbiology & Zoology (Subsidiary)
M.Sc.	Poona, India	1985	I (69%)	Biochemistry
M.Phil.	Poona, India	1987	'O' grade (75%)	Biochemistry
Ph.D.	Poona, India	Thesis submitted		Chemistry

Title of the Ph.D. Thesis : A molecular approach for the assesment of variability in indica rices.

Prizes/Scholarships/Fellowships

- Recipient of three prizes for standing first in B.Sc. Chemistry, while in M.E.S. College of Poona University, India (1983).
- Recipient of National Merit Scholarship during M.Sc. course (1983-1985).
- Recipient of award of University Research Stipend during M.Phil course (1985-1986).
- Recipient of Junior Research Fellowship of the Council of Scientific and Industrial Research (CSIR), New Delhi, India (1986-1988).
- Recipient of Senior Research Fellowship of the Council of Scientific and Industrial Research (CSIR), New Delhi, India (1988-1990).
- Invited as a Pre-Doc fellow at Perpignan, France (in the laboratory of Dr. Michel Delseny) to learn advanced techniques in Plant Molecular Biology and Genetic Engineering under Rockefeller Foundation fellowship for a period of three months (1st May 1989- 31st July 1989).

Membership of Scientific Bodies

- Member, Society of Biological Chemists, India.
- Member, Indian Society of Cell Biology.

Participation at Workshops/Symposia

- Participated in the fourth annual seminar for Research students of the University of Poona, India on 7th and 8th February 1986, during M. Phil course.
- Attended workshop on 'Genetic Engineering in India' held at Indian Institute of Science (IISc), Bangalore from 26/7/88 to 28/7/88.
- Was one of the instructing faculty at the workshop on 'Plant Genetic Engineering: Cloning and analysis of a specific plant DNA fragment" held in the Division of Biochemical Sciences, National Chemical Laboratory, Poona, India, from 10/09/1987 to 30/09/1987.
- Was one of the instructing faculty at the workshop on "Cloning and analysis of a plant DNA fragment and Agrobacterium mediated transformation in plants" held at National Chemical Laboratory, Poona, India, from 28/03/89 to 17/04/89.

List of Publications

1. Vaijayanti Pethe, Meena Lagu, P.K.Chitnis, Vidya Gupta, and P.K. Ranjekar.
Restriction Fragment Length Polymorphism. A recent approach in plant breeding.
Indian Journal of Biochemistry and Biophysics. 26(5): 285-288 October 1989.
2. M.S. Dhar, V.V. Pethe, V.S. Gupta, P.K. Ranjekar.
Predominance and tissue specificity of Adenine methylation in Rice.
Theoretical and Applied Genetics: 80: 1-7, 1990.
3. Pethe, V.V. Chitnis, P.K., Lagu, M.D., Gupta, V.S. and Ranjekar, P.K.
Characterization of a rice RFLP marker and demonstration of its potential in rice breeding.
Communicated for Publication.

Presentations/Abstracts:

1. Pethe V.V., Panchal R, Bhide S.V., Kale N.R.
Temperature dependence of the specific interaction of Concanavalin A (Con A) with chromogenic ligand Blue dextran 2000.
Presented at the 7th Annual Seminars for Research students, University of Poona, February, 1986.

2. Pethe, V.V., Bhide S.V., Kale N.R.
Influence of Phosphorylated derivatives of D-glucose and D-fructose and Con A- Glycogen interaction.
Abstract, " 55th annual meeting of Society of Biological Chemists" (India), at Trivandrum 1986.
3. P.K. Chitnis, V.V. Pethe, M.D. Lagu, V.S. Gupta and P.K. Ranjekar
Genetic variation and mapping in rice using RFLP approach. Abstract, In IVth Annual Meeting of the Rockefeller Foundations International Program on Rice Biotechnology, May 9-12, 1990
4. V.V. Pethe, P.K. Chitnis, M.D. Lagu, V.S. Gupta and P.K. Ranjekar
Genetic variability amongst the elite Maharashtra rice germplasm using RFLP approach.
Abstract, In International Symposium on Rice Research: New Frontiers, November 15-18, 1990, Hyderabad, India.
5. Vidya S. Gupta, Madhu S. Dhar, Vaijayanti V. Pethe and P.K. Ranjekar
Cloning and restriction mapping of rice long repetitive DNA
Abstract, In A Symposium on Genetic Engineering Research in India, July 26-28, 1988, Bangalore, India.

6. V.V. Pethe, P.K. Chitnis, M.D. Lagu, V.S. Gupta and P.K. Ranjekar

Use of RFLP techniques in selective rice breeding.
Abstract, International conference on Genetic Engineering and Biotechnology. Nepal Biotechnology Association (NBA), April 15-19, 1991

Research Experience

1. Two years research experience in Protein Chemistry.
2. Three years research experience in Plant Molecular Biology and Genetic Engineering involving gene cloning and library construction, Southern hybridization and DNA sequencing technique.

References

1. Dr. P. K. Ranjekar
Division of Biochemical Sciences
National Chemical Laboratory
Pune 411008, India.
2. Dr. Michel Delseny
Laboratoire De Physiologie Végétale
Université De Perpignan
66025 Perpignan Cedex, France.
3. Dr.(Mrs) Vidya Gupta
Division of Biochemical Sciences
National Chemical Laboratory
Pune 411008, India.
4. Dr. A.S. Reddy
C/o Dr. Terry L. Thomas
Associate Research Scientist
Texas A & M University
Dept of Biology
College Station, Texas.
77843-3258, USA

Predominance and tissue specificity of adenine methylation in rice

M.S. Dhar, V.V. Pethe, V.S. Gupta and P.K. Ranjekar*

Division of Biochemical Sciences, National Chemical Laboratory, Pune–411008, India

Received November 21, 1989; Accepted May 15, 1990

Communicated by G.S. Khush

Summary. Using 'A' and 'C' methylation-specific restriction enzymes, namely, MboI, Sau3AI, DpnI, MspI, and HpaII, total rice cv Basmati 370 DNA, repetitive DNAs, and a specific repeat sequence indicated an abundance of adenine methylation. Although cytosine methylation in 5'-CCGG-3' sequences suggested more CpC methylation than CpG, the 'C' methylation in sequence 5'-GATC-3' was comparatively less than 'A' methylation. Furthermore, the presence of adenine methylation was tissue specific; it was predominant in rice shoot DNA as compared to embryo DNA. This pattern was also observed in two other cultivars of rice, i.e., R-24 and Sona, and was again confirmed using a cloned probe of a specific repeat sequence. Besides the changes in adenine methylation, there was also a qualitative change in 5mC from CpG to CpC dinucleotides in these two tissue systems.

Key words: Rice – Adenine methylation – Tissue specificity – Genomic rearrangement

Introduction

The role of 5-methyl cytosine, 5mC, has been well examined in animal genomes as compared to plants. It has been shown that the presence of 5mC is negatively correlated with gene activity. Organ- or tissue-specific variations in methylation patterns have also been observed in many animal genes, such as rabbit B-globin and chicken ovalbumin (Mandel and Chambon 1979; Compere and Palmiter 1981). In plants, methylation patterns of rDNA have been thoroughly studied. In general, the ribosomal genes seem to be heavily methylated in a few plant species

investigated so far (Goldsbrough and Ellis 1981; Uchimiya et al. 1982; Ellis et al. 1983; Segel and Kolacz 1983; Olmedilla et al. 1984), except the rice rDNA unit. In rice, there is no change in rDNA methylation from seedlings grown in aerobic and anaerobic conditions (Aspart et al. 1983), although there is a change in the synthesis of rDNA genes. Tissue- or stage-specific variations are observed in radish rDNA (Delseny et al. 1984) and zein-light chain genes (Spena et al. 1983). Recently it has been shown in onion, broad bean, and rice that repeat DNA sequences undergo a change during differentiation and dedifferentiation events in cell culture conditions, and such variations are referred to as 'modulation' (Durante et al. 1977; Cremonini et al. 1981; Bassi et al. 1984; Zheng et al. 1985; Kikuchi et al. 1987).

In the present work, we describe the restriction enzyme analysis of total as well as repetitive DNA fractions of rice cv Basmati 370 using methylation-specific enzymes, i.e., MboI, Sau3AI, DpnI, MspI, and HpaII. To assess whether the methylation patterns are tissue/stage-specific, studies were also carried out in two tissue systems, i.e., embryo and shoot. Finally, to examine if changes in the methylation status of rice DNA during development are a general feature of the rice genome or if they are variety specific, similar experiments were carried out using the DNAs of two more rice varieties, namely, R-24 and Sona.

Materials and methods

Isolation of total DNA and repetitive DNA

Native, high-molecular-weight DNAs were prepared from embryos and shoots of three different varieties of rice cultivars, Basmati 370, R-24, and Sona, according to the procedures of Marmur (1961) and Ranjekar et al. (1976).

* To whom correspondence should be addressed

Long (9–20 kbp) and short (0.2–0.3 kbp) repeat DNA sequences reassociating by Cot 0.1 and Cot 50 were isolated according to Dhar et al. (1988).

Restriction enzyme analysis, gel electrophoresis, and Southern hybridization

Restriction enzyme digestions of all the DNAs with a few methylation-specific enzymes, such as MboI, Sau3AI, DpnI, MspI, and HpaII, were carried out according to Maniatis et al. (1982). The DNA digests were analyzed on 1% and 1.4% agarose gels or 10% polyacrylamide gels, using both the high- and low-molecular-weight DNA markers. The total pRL7 DNA (3.6 kbp long rice repetitive DNA cloned in pBR 325) and the Sona 2.6 kbp repeat family were labelled by nick-translation as described by Rigby et al. (1977). Southern hybridization experiments were essentially according to Maniatis et al. (1982).

Results

Methylation status of rice shoot DNA

Figure 1A shows the restriction enzyme digestion patterns of total rice cv Basmati 370 shoot DNA with MboI, Sau3AI, and DpnI. From this figure, it is clear that MboI (lane b) digestion of rice DNA is less than that of Sau3AI (lane c). Since it is known that the restriction enzyme Sau3AI does not cut the sequence 5'-GATC-3' if C is methylated, whereas MboI does not cut the sequence 5'-GATC-3' if A is methylated, the difference in the extent of digestion of rice shoot DNA with these two enzymes indicates the abundance of methylated adenine as compared to methylated cytosine in the sequence 5'-GATC-3'. This is further confirmed by the occurrence of extensive digestion of shoot DNA with DpnI (lane d), which digests the DNA only if A is methylated in the sequence 5'-GATC-3'.

Restriction enzyme digestion of total rice shoot DNA with MspI and HpaII is shown in Fig. 1B. These two restriction enzymes are known to be sensitive towards C methylation in the sequence 5'-CCGG-3'. MspI does not cut the 5'-^mCCGG-3' sequence, whereas it is insensitive to internal C methylation. On the other hand, HpaII is sensitive to internal C methylation and does not cut the sequence, 5'-C^mCCGG-3'. It can be seen from Fig. 1B that HpaII (lane c) shows comparatively more digestion of DNA than that with MspI (lane b), suggesting the abundance of 5'-^mCCGG-3' type of sequences rather than 5'-C^mCCGG or 5'-CCGG-3' type in rice genome. Thus, this suggested a high frequency of ^mCpC dinucleotide as compared to ^mCpG dinucleotide in the 5'-CCGG-3' sequence in rice DNA.

Presence of methylated adenine in rice-repetitive DNA

Since the total rice genomic DNA showed peculiar patterns of 'A' and 'C' methylation, we were interested in studying methylation patterns of its repetitive DNA se-

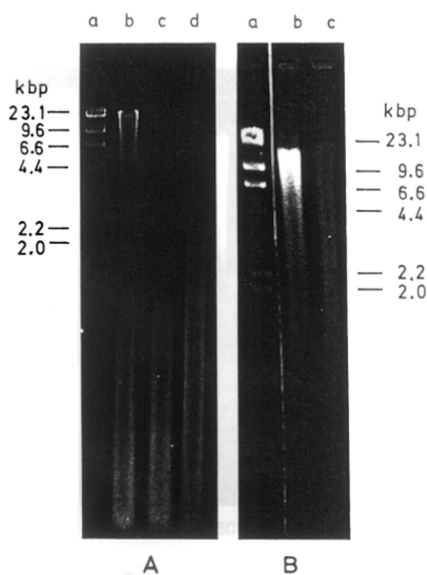


Fig. 1A and B. Digestion of total rice cv Basmati 370 shoot DNA with MboI, Sau3AI, DpnI (A) and MspI, HpaII (B). A Lambda HindIII digest lane a; MboI digest lane b; Sau3AI digest lane c; DpnI digest lane d. B Lambda HindIII digest lane a; MspI digest lane b; HpaII digest lane c. Electrophoresis was carried out on 1.4% (A) and 1% (B) neutral agarose slab gels in TAE buffer, pH 8.1, at a constant current of 30 mA

quences. For this purpose, long (9–20 kbp) and short (0.2–0.3 kbp) repeats of Cot 0.1 and Cot 50 fractions were isolated and digested with the above two sets of restriction enzymes and the results are shown in Fig. 2A and B. From Fig. 2A, it is evident that the digestion of the long repeats of Cot 0.1 DNA is less with MboI (lane b) than that with Sau3AI (lane c), thus revealing the presence of methylated adenine in 5'-GATC-3' sequences in rice highly repetitive DNA fractions. Though not as distinct as Cot 0.1 DNA, a similar methylation pattern is also observed in the case of Cot 50 DNA (Fig. 2A). This is further confirmed by the reasonable digestion of the DNAs with DpnI (lane d) compared to MboI (lane b) in both cases. Comparatively more digestion of these repeat sequences by Sau3AI (lane c) also suggests the absence of 'C' methylation in the sequence 5'-GATC-3'. Figure 2B shows the digestion patterns of long repeats of Cot 0.1 and 50 DNA with MspI (lane b) and HpaII (lane c). Here, HpaII digests the DNAs more than that with MspI, suggesting again the abundance of 5'-^mCCGG-3'/5'-CCGG-3' type of sequences in these repeat fractions. When the digests of short repeat sequences of both Cot 0.1 and Cot 50 with the methylation-specific enzymes were examined by polyacrylamide gel electrophoresis, it was observed that the short repeats were

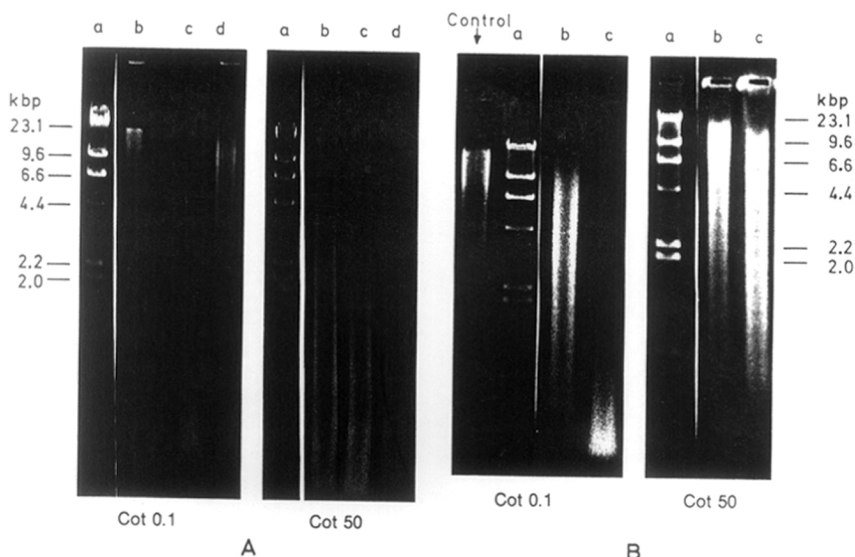


Fig. 2 **A and B.** Digestion of long repeats of Cot 0.1 and Cot 50 of rice cv Basmati 370 with MboI, Sau3AI, DpnI (A) and MspI, HpaII (B). **A** Lambda HindIII digest lane a; MboI digest lane b; Sau3AI digest lane c; DpnI digest lane d. **B** Lambda HindIII digest lane a; MspI digest lane b; HpaII digest lane c. Electrophoresis was carried out on 1.4% (A) and 1% (B) neutral, agarose slab gels in TAE buffer, pH 8.1, at a constant current of 30 mA

too small to reveal any distinct differences (data not shown).

Methylation status of rice embryo DNA

To assess if the methylation patterns of rice shoot DNA are tissue specific, rice embryo DNA from cv Basmati 370 was digested with MboI, Sau3AI, DpnI, MspI, and HpaII. It is clearly seen from Fig. 3A and B that the trend of 'A' and 'C' methylation is reversed in embryo DNA. For example, the restriction enzyme DpnI, which is specific for A methylation in the sequence 5'-GATC-3', cuts the rice shoot DNA extensively (Fig. 1A, lane d). However, in rice embryo DNA, the digestion by DpnI (Fig. 3A, lane d) is very small, suggesting low frequency of A methylation in the sequence 5'-GATC-3'. Furthermore, Fig. 3B shows MspI (lane b) digestion to be greater than HpaII (lane c), suggesting a predominance of 5'-C^mCGG-3' type of sequences. Thus 5^mC occurs as a part of CpG dinucleotides rather than of CpC dinucleotides in the embryo DNA. Apart from differences in DNA methylation, tissue-specific repeat families are also seen. The Basmati 370 embryo DNA shows a number of bands with both MspI and HpaII in the range of 0.6 to 4.0 kbp, suggesting the presence of these tissue-specific repeat families that are absent in the corresponding shoot DNA digests.

Distribution of pRL7 sequence in rice embryo and shoot DNA

The presence or distribution of a specific repeat sequence was next studied in the embryo and shoot system of cv Basmati 370 using a cloned pRL 7 sequence as a probe. For this purpose, the high-molecular-weight rice embryo and shoot DNAs were digested with methylation-specific restriction enzymes MboI and Sau3AI and a few enzymes like AluI and TaqI (Fig. 4A and B), and were hybridized to ³²P-labelled plasmid pRL7 DNA. It can be seen from Fig. 5A that a single strong band is present per lane with each enzyme, with the absence of any background smear in the case of embryonic DNA. The molecular weight of this band varies in a very small range of 2.3 to 2.5 kbp. In the case of shoot DNA, the hybridization pattern is observed to be different with each enzyme (Fig. 5B). With MboI (lane a), for example, two bands of molecular weights of about 15 and 5.5 kbp are prominent. With Sau3AI (lane b), four bands in the region 0.6–1.36 kbp and one band at 0.31 kbp are observed. In the case of AluI and TaqI (lanes c and d), however, many bands are seen in the range of 0.5 to 6.6 kbp. From these results, therefore, it is clear that there is some reorganization in this specific repeat unit in embryo and shoot systems.

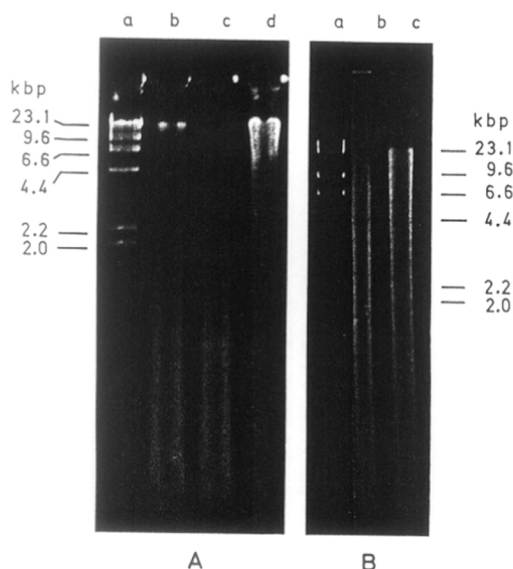


Fig. 3A and B. Digestion of rice cv Basmati 370 embryo DNA with MboI, Sau3AI, DpnI (A) and MspI, HpaII (B). **A** Lambda HindIII digest lane a; MboI digest lane b; Sau3AI digest lane c; DpnI digest lane d. **B** Lambda HindIII digest lane a; MspI digest lane b; HpaII digest lane c. Electrophoresis was carried out on 1.4% (A) and 1.0% (B) neutral, agarose slab gels in TAE buffer, pH 8.1, at a constant current of 30 mA.

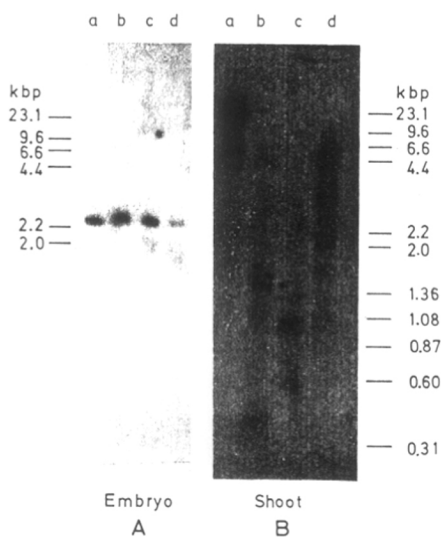


Fig. 5A and B. Autoradiograms showing the hybridization of pRL7 cloned DNA with rice cv Basmati 370 embryo (A) and shoot DNAs (B) digested with different restriction enzymes. **A and B** MboI digest lane a; Sau3AI digest lane b; AluI digest lane c; TaqI digest lane d.

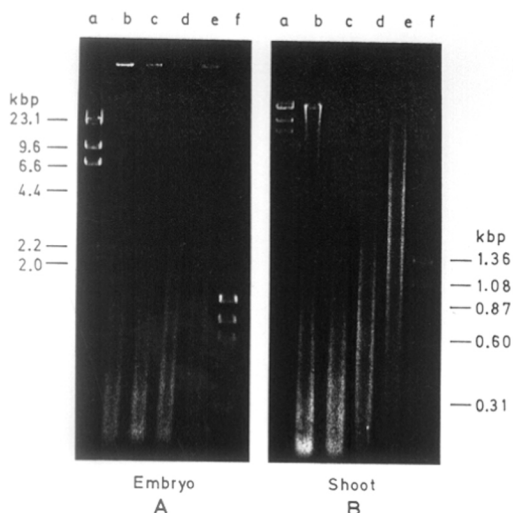


Fig. 4A and B. Digestion of total rice embryo (A) and shoot DNA (B) with different restriction enzymes. **A and B** Lambda HindIII digest lane a; MboI digest lane b; Sau3AI digest lane c; AluI digest lane d; TaqI digest lane e; Φ × 174 HaeIII digest lane f. Electrophoresis was carried out on 1.4% agarose gels in TAE buffer, pH 8.1, at a constant current of 30 mA.

Restriction enzyme analysis of two other cultivars of rice

Restriction enzyme analysis of shoot and embryo DNAs from two more rice cultivars, viz., R-24 and Sona, was next undertaken. The aim of this experiment was to see whether the tissue/stage-specific methylation changes observed in Basmati 370 were variety specific or could also be found in other varieties. Figures 6 and 7 A, B show the digestion patterns of embryo and shoot DNAs of the two rice varieties R-24 and Sona with MboI, Sau3AI, DpnI, and MspI, HpaII. The digestion patterns with MboI, Sau3AI, and DpnI clearly indicate the predominance of 'A' methylation in the shoot DNAs (Fig. 6B) as compared to the embryo DNAs (Fig. 6A) in both varieties. Also, the MspI and HpaII digestion patterns (Fig. 7) in these DNAs are consistent with that of Basmati 370, wherein a quantitative change in 5^mC from CpG to CpC dinucleotides is observed during the transition from embryo to shoot.

Another finding from Figs. 6 and 7 is that the methylation-specific restriction enzymes also show variety as well as tissue-specific band patterns. In Sona shoot DNA, for instance, the MboI (Fig. 6B, lane b) digest shows two bands of 2.6 kbp and 0.9 kbp, while Sau3AI (Fig. 6B, lane c) reveals a single band of 2.6 kbp. This indicates the presence of regularly arranged, unmethylat-

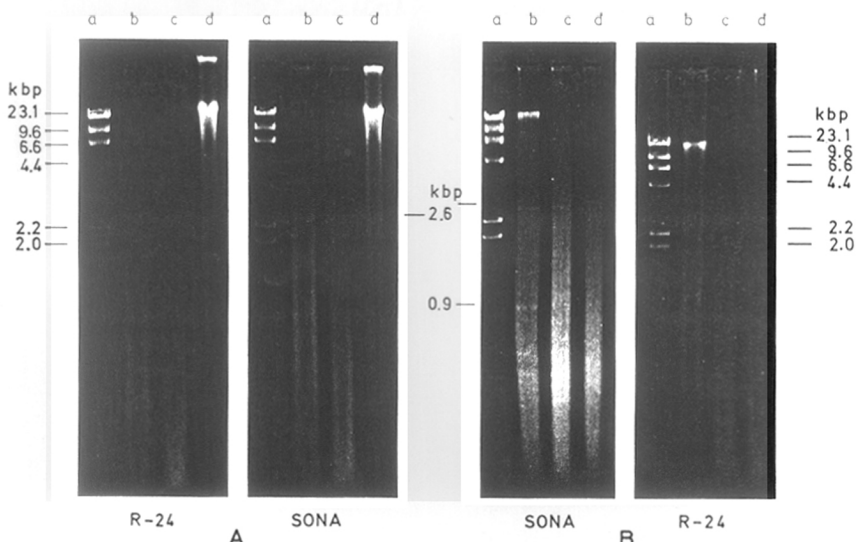


Fig. 6 A and B. Digestion of rice cv R-24 and sona embryo (A) and shoot DNAs (B) with MboI, Sau3AI, and DpnI. A and B Lambda HindIII digest lane a; MboI digest lane b; Sau3AI digest lane c; DpnI digest lane d. Electrophoresis was carried out on 1.4% neutral, agarose slab gels in TAE buffer, pH 8.1, at a constant current of 30 mA.

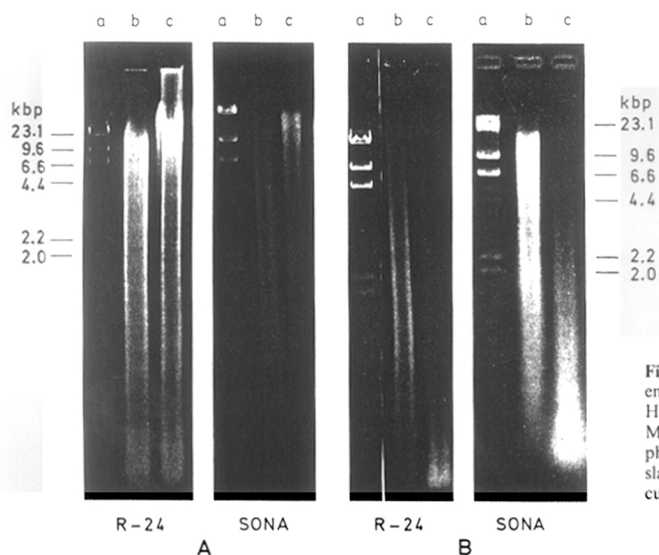


Fig. 7 A and B. Digestion of rice cv R-24 and sona embryo (A) and shoot DNAs (B) with MspI, HpaII. A and B Lambda HindIII digest lane a; MspI digest lane b; HpaII digest lane c; Electrophoresis was carried out on 1% neutral, agarose slab gels in TAE buffer, pH 8.1, at a constant current of 30 mA.

ed 5'-GATC-3' sequences in the Sona shoot DNA. Also, it shows that there are DNA sequences that include methylated 5'-GATC-3' sequences arranged at a regular interval of 2.6 kbp. These methylated as well as unmethylated sequences may or may not belong to same/similar

type of repeat families. In Sona embryo DNA (Fig. 6A), the 2.6-kbp Sau3AI and 0.9-kbp MboI repeat elements are absent, whereas the 2.6-kbp MboI repeat family is present. The R-24 embryo DNA, when digested with MboI, shows many bands with varying molecular

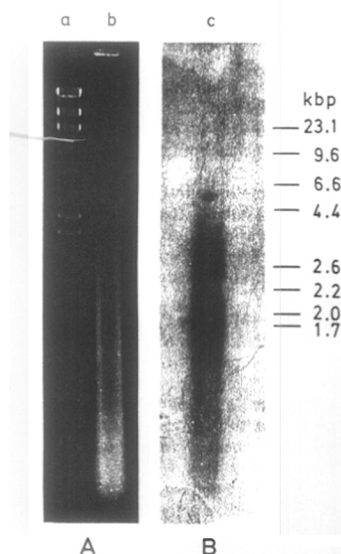


Fig. 8A and B. Digestion pattern and autoradiogram of sona embryo DNA digested with *Sau3AI* and hybridized to Sona shoot *Sau3AI* repeat family (2.6 kbp) as a probe. Lambda *HindIII* digest lane a; Sona embryo *Sau3AI* digest lane b; hybridization of 2.6 kbp Sona shoot *Sau3AI* repeat family to the above digest lane c

weights in the range of 0.5 to 2.2 kbp, while these are absent in the corresponding shoot DNA digest.

When the 2.6-kbp band that is present in the *Sau3AI* digest of Sona shoot DNA (Fig. 6B, lane c) and absent in the embryo DNA (Fig. 6A, lane c) is eluted, labelled, and hybridized to the gel blot of the embryo *Sau3AI* digest, two bands of about 2.6–2.7 kbp and 1.7 kbp are observed (Fig. 8). This suggests that there is a change in the *Sau3AI* repeat element with respect to either its copy number or methylation of its 5'-GATC-3' sequences during the transition from embryo to shoot stage in Sona DNA.

Discussion

An important feature of the present studies is that rice shoot DNA and repetitive DNA have an abundance of 'A' methylation. This is a novel finding in view of the very few reports on adenine methylation in higher plants (Pintor-Toro 1987). It is also observed that this pattern of methylation is tissue specific, i.e., there is a predominance of 'A' methylation in rice shoot DNA as compared to embryo DNA. Since tissue-specific methylation is observed in three varieties of rice, this information may be of structural or functional significance in rice development and may help to assign a role to the repeat se-

quences in this process. The occurrence and role of adenine methylation is still unclear in higher plants. Recently, methylated adenine has been observed in zein genes, but the experimental evidence suggests that it is unrelated to their regulation of expression (Pintor-Toro 1987). In the case of animals and microorganisms, however, methylated adenine has defined roles (Sterglanz and Bugg 1973; Engel and Von Hippel 1978; Hattman 1982; Plasterk et al. 1983; Pukkila et al. 1983; Cheng et al. 1985; Hare and Taylor 1985). For instance, in *E. coli* the rate of adenine methylation within 5'-GATC-3' sequences modulates the activity of certain promoters (Hattman 1982; Plasterk et al. 1983). Engel and Von Hippel (1978) have shown that the presence of 6-methyladenine results in a destabilization of the DNA helix, presumably because the N1-*cis* orientation of the methyl group is favored. Such effects may be significant in protein-DNA interactions (Sterglanz and Bugg 1973; Cheng et al. 1985).

Another important feature of the present work is the differential methylation of repeat DNA that corresponds to a specific, cloned repeat element pRL 7 during the transition from embryo to shoot stage. This repeat unit is present in a well-organized manner in the embryo DNA, whereas in the case of shoot DNA, many copies of this sequence are present in a varied environment with random distribution, while some are well organized to give a particular band pattern. From this data it appears that there is some form of genomic rearrangement of a specific, repetitive sequence and it may partly be due to the changes in methylation of this repeat unit. In the literature, there are a few reports on genomic alterations in tissue culture conditions. In rice, for example, Zheng et al. (1985) have reported the amplification of a specific, highly repeated sequence in cultured cells of variety Roncarlo during the differentiation process. Kikuchi et al. (1987) have given preliminary results in rice (cultivar Nipponbare) showing variations in the copy number of two repeated DNA sequences, one being amplified and the other being reduced in copy number in the differentiation process under cultured conditions.

Ours is probably the first report that describes the differential methylation status of specific repeat sequences in rice (cv Basmati 370) during the transition from embryo to shoot without any external inducing factor.

References

- Aspart L, Got A, Delseny M, Mocquot MB, Pradet A (1983) Adaptation of ribonucleic acid metabolism to anoxia in rice embryos. *Plant Physiol* 72:115–121
- Bassi P, Cionini PG, Cremonini R, Seghizzi P (1984) Underrepresentation of nuclear DNA sequences in differentiating root cells of *Vicia faba*. *Protoplasma* 123:70–77

- Cheng S, Gail H, Modrich P (1985) Extent of equilibrium perturbation of the DNA helix upon enzymatic methylation of adenine residues. *J Biol Chem* 260:191–194
- Comper SJ, Palmer RD (1981) DNA methylation controls the inducibility of the mouse metallothionein-1 gene in lymphoid cells. *Cell* 25:233–240
- Cremonini R, Lo Iero M, Durante M (1981) Differentiation of metaxylem cell line in the root of *Allium cepa*. II. Changes in nucleotide sequence redundancy. *Protoplasma* 108:1–7
- Delseny M, Laroche M, Penon P (1984) Methylation pattern of radish, *Raphanus sativus*, nuclear ribosomal RNA genes. *Plant Physiol* 76:627–634
- Dhar MS, Dabak MM, Gupta VS, Ranjekar PK (1988) Organization and properties of repeated DNA sequences in rice genome. *Plant Sci* 55:43–52
- Durante M, Cremonini R, Brunori A, Avanzi A, Innocenti M (1977) Differentiation of metaxylem cell line in the root of *Allium cepa*. I. DNA heterogeneity and ribosomal cistrons of two different stages of differentiation. *Protoplasma* 93:289–304
- Ellis THN, Goldsbrough PB, Castleton JA (1983) Transcription and methylation of flax rDNA. *Nucleic Acids Res* 11:3041–3046
- Engel JD, Von Hippel PH (1978) Effects of methylation on the stability of nucleic acid conformation. *J Biol Chem* 253:927–934
- Goldsbrough PB, Ellis THN (1981) The analysis of the distribution of restriction endonuclease sites in repetitive DNAs. *Nucleic Acids Res* 9:1551–1558
- Hare JT, Taylor JH (1985) One role for DNA methylation in vertebrate cells is strand discrimination in mismatch repair. *Proc Natl Acad Sci USA* 82:7350–7354
- Hattman S (1982) DNA methyltransferase-dependent transcription of the phage Mu mom gene. *Proc Natl Acad Sci USA* 79:5518–5521
- Kikuchi S, Takaiwa F, Oono K (1987) Variable copy number DNA sequences in rice. *Mol Gen Genet* 210:373–380
- Mandel JL, Chambon P (1979) DNA methylation: organ specific variations in the methylation patterns within and around ovalbumin and other chicken genes. *Nucleic Acids Res* 7:2081–2103
- Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor/NY
- Marmur J (1961) A procedure for the isolation of deoxyribonucleic acid from microorganisms. *J Mol Biol* 3:208–218
- Olmedilla A, Delcasso D, Delseny M (1984) Methylation pattern of nuclear ribosomal RNA genes from rice (*Oryza sativa*). *Plant Sci Lett* 37:123–127
- Pintor-Toro JA (1987) Adenine methylation in zein genes. *Biochem Biophys Res Comm* 147:1082–1087
- Plasterk RHA, Vrieling H, Van de Putte P (1983) Transcription initiation of Mu mom depends on methylation of the promoter region and a phage-coded transactivator. *Nature London* 301:344–347
- Pukkila PJ, Peterson J, Herman G, Modrich P, Messelson M (1983) Effects of high levels of DNA adenine methylation on methylation-directed mismatch repair of *Escherichia coli*. *Genetics* 104:571–582
- Ranjekar PK, Pallotta D, Lafontaine J (1976) Analysis of the genome of plants. II. Characterization of repetitive DNA in barley *Hordeum vulgare* and wheat *Triticum aestivum*. *Biochem Biophys Acta* 425:30–40
- Rigby PWJ, Dieckman M, Rhodes C, Berg O (1977) Labeling deoxyribonucleic acid to high specific activity in vitro by nick-translation with DNA polymerase I. *J Mol Biol* 113:237–251
- Segel A, Kolacz K (1983) Heterogeneity of pumpkin ribosomal DNA. *Plant Physiol* 72:166–171
- Spena A, Viotti A, Pirrotta V (1983) Two adjacent genomic zein sequences: structure, organization, and tissue-specific restriction pattern. *J Mol Biol* 169:799–811
- Sterglanz H, Bugg CE (1973) Conformation of N⁶-methylation, a base involved in DNA modification: restriction process. *Science* 182:833–834
- Uchimiya H, Kato H, Ohgawara T, Harada H, Sugiura M (1982) Sequence-specific methylation of ribosomal RNA genes contained in the nuclear DNA of tobacco. *Plant Cell Physiol* 23:1129–1131
- Zheng KL, Castiglione S, Biasini MG, Biroli A, Morandi C, Sale F (1985) Nuclear DNA amplification in cultured cells of *Oryza sativa* L. *Theor Appl Genet* 74:65–70

Restriction fragment length polymorphism: A recent approach in plant breeding

Vaijayanti Pethe, Meena Lagu, P K Chitnis, Vidya Gupta & P K Ranjekar*
Division of Biochemical Sciences, National Chemical Laboratory, Pune 411 008
Received 9 February 1989; revised 27 June 1989

Introduction

A plant genetic improvement cycle is generally initiated by the perception that an existing cultivar is deficient in one or more traits. Identification of source of alleles producing the desired expression of the trait and introduction of these alleles into the cultivar are the important steps in plant breeding procedure. These classical breeding procedures of crossing, backcrossing, selfing and selection entail considerable time-lag and often involve a few generations before a new variety is released. Recent molecular approaches such as plant genetic engineering and Restriction Fragment Length Polymorphism (RFLP) have proved useful in plant improvement/breeding and are also expected to shorten the time-lag. The RFLPs in particular have a special and virtually unique potential for the manipulation of a genetic variation which is of a phylogenetic quantitative nature, and can also play a major role in increasing the effectiveness of current and future breeding practices¹.

Nucleotide variations in DNA sequence occur as a result of a number of natural processes, viz. mutations, deletions, insertions and effects of environmental factors. Some of these variations are detected at the level of populations of the same species, while others are identifiable between species. Similarly, very few variations are expressed in the phenotype, while many others are silent. It has been estimated in the case of human genome that nucleotide polymorphism in a normal population occurs at a frequency of one in hundred². This type of nucleotide length variation (intraspecific as well as interspecific) when characterized using restriction endonucleases is described as RFLP¹. It is detected and quantitated for specific fragment(s) of DNA after digestion with restriction endonucleases. In case of simple unique DNA sequences, like those of bacteria, viruses and organellar DNAs in eukaryotes

(e.g. mitochondria or chloroplast DNAs), a fragment length polymorphism can often be detected simply by conventional gel electrophoresis and staining with ethidium bromide. However, in case of complex eukaryotic DNAs, this polymorphism can be identified only after using specific probes followed by autoradiography.

RFLPs were first used as a tool for the genetic analysis in 1974, when a temperature sensitive mutation of adenovirus was associated with a specific RFLP³. The use of RFLPs as markers of genetic disease was first made by Kan and Dozy⁴ in 1978 for the analysis of sickle cell anaemia. Botstein *et al.*⁵ described the theoretical basis of this method for mapping genes associated with disease in humans. Since then this technique has been routinely used for direct detection and mapping of a nucleotide sequence within the human genome. It also serves as a powerful tool for assaying the genetic inheritance of a diseased gene and developing detailed genetic linkage maps. Though a majority of methods and applications of RFLP analysis have been developed for human genetics, Burr *et al.*⁶ described, for the first time, a possible use of this methodology in the improvement of crop plants. Later, the methods for mapping and applications of RFLP in plant breeding have been reviewed by Beckmann and Soller¹.

Criteria for selection of RFLP probes/markers

The RFLP patterns for a given source DNA and probe DNA are highly specific. Consequently, a large number of empirical trials are required with different probe enzyme combinations and source DNAs before a meaningful RFLP pattern can emerge. Once such pattern is identified, it has to be correlated to known phenotypic, isoenzyme and other genetic loci. Establishment of such a correlation finally results in the use of RFLP as a marker.

Since RFLP results from changes in parameters such as nucleotide pairs, sequence translocations, transpositions, deletions and inversions, these very

*To whom all correspondence should be addressed.

parameters can be used to detect RFLPs. In the case of eukaryotic genomes, detection of RFLPs becomes slightly difficult on account of the presence of large proportions of repetitive DNAs. In such cases, generally single copy or unique DNA sequences are preferred as probes for RFLP. These sequences can originate either from genomic libraries by selection or from cDNA libraries. Chromosome sorting can be of great help in generating probes. With the help of this technique chromosomes can be sorted out and their respective libraries can be constructed. Screening with probes derived from chromosome specific libraries will help in loci identification and subsequent mapping of the same⁷.

Irrespective of nature and type of probe, there are some specific requirements for detection of RFLP. These are availability of maximum possible probes, well established homozygous genomes/genotypes, genetic linkage map for the known phenotypic and/or isoenzyme and other markers/loci and/or various aneuploids for confirmation and/or localization of RFLP loci.

From the technical point of view, the probes that are selected must also exhibit the following criteria:

- (i) They should yield a strong hybridization signal
- (ii) They should reveal small number of fragments
- (iii) They should reveal polymorphisms among different homozygous lines

Types of probes/markers used in RFLP

For genetic mapping with RFLP, various kinds of probes have generally been used. These include single copy (single genes), low copy (gene families) and even tandemly repeated DNA sequences⁸. Thus, for the RFLP analysis of human DNAs, a variety of probes from unique and repeated DNA sequences were found to be useful⁹. However, little is known about the comparative qualities of single copy and repeated DNA sequences as probes/markers for plant material. The other class of probes that can be used is retrotransposons or pseudogenes. These probes have greater resolving power particularly in cultivar to cultivar comparison. Retrotransposons include processed mRNA or pseudogenes and also highly reiterated sequences such as human *Alu I* family. Retrotransposition is believed to occur via a process involving reverse transcription of RNA sequences and subsequent integration of DNA copies into chromosomal DNA. Other promising probes for RFLP work include non-transcribed spacer regions¹⁰, probes derived from mobile elements¹¹ and DNA sequences which shuttle between nuclear and organelle DNAs¹².

Practical approaches for the detection of RFLP

With reference to eukaryotic genomes in general and plant genomes in particular, a number of methodologies are now available for the detection and analysis of RFLPs. Southern blotting and hybridization is a method of choice and routine application¹³. This basically involves separation of restriction enzyme fragments by gel electrophoresis, blotting and hybridization of the blots to suitable labelled probes. Dot blots are efficient alternative to Southern blots, when large populations are to be screened for the presence or absence of specific DNA sequences. Partial digest method is often used to detect genomic DNA variations with infrequent polymorphism. Here, genomic DNA is partially digested (60-80%) and probed with the sequence of interest after Southern blotting.

State of art of RFLP work in plants

To date, RFLP analysis has been used for the genetic analysis of a number of plant species such as maize¹⁴⁻¹⁶, tomato^{14,17,18}, pea¹⁹, barley¹¹ and very recently rice²⁰. These studies have resulted in interesting correlations between polymorphic restriction fragments and inheritable characters such as solid content of tomato fruit¹⁷. The RFLPs in case of plants have also been applied to the understanding of species relationships in the genus *Brassica* and related genera like *Diploaxis* and *Eruca*²¹. In this context, it is interesting to note that one of the plant biotechnology companies in the US (NPI, Salt Lake City, Utah) has applied RFLP technology on a commercial scale. The work dealing with RFLP analysis in a few important plant species is described below.

(i) *Maize—Maize (Zea mays L.)* is the first plant whose genetic linkage maps were constructed using RFLPs as the source of genetic markers¹⁴. Simultaneously, Helentjaris *et al.*¹⁵ have described the use of monosomics to map cloned DNA fragments in maize. They used RFLP to attribute these cloned DNAs to definite loci and subsequently arranged these loci into linkage groups by segregation analysis. In this study, these linkage groups were assigned to specific chromosomes facilitating the construction of a linkage map for maize containing 145 RFLP loci. Recently, the group of same scientists was successful in identifying the genome locations of duplicate nucleotide sequences in maize by analysis of RFLPs¹⁶. While preparing a linkage map for maize, based upon loci detected through the use of RFLPs, they found that 62 out of the 217 cloned maize sequences tested detected more than one fragment on genomic Southern blots. From the genomic locations of these duplicate sequences, they

found that these sequences usually originated from different chromosomes.

(ii) *Tomato*—Helentjaris *et al.*¹⁴ have also constructed the genetic linkage maps for tomato (*Lycopersicon* sp.) using RFLPs as source of genetic markers. In order to detect these RFLPs, they used unique DNA sequence clones for hybridization with Southern blots containing restriction enzyme digested genomic DNA from different homozygous lines. A subsequent comparison of RFLP inheritance patterns in F_2 populations from tomato was used to study the arrangement of loci detected by these clones into genetic linkage groups for the *Lycopersicon* sp. Simultaneously Bernatzky and Tanksley¹⁸ have developed a saturated linkage map in tomato based on isozymes and random cDNA sequences. Interspecific backcross and F_2 populations of *Lycopersicon esculentum* and *L. pennellii* were employed in the linkage analysis. Allelic differences in cDNA markers were based on RFLPs detected through Southern analysis. A total of 57 unique cDNA clones were analyzed and majority of these markers were found to correspond to single loci and dispersed throughout the genome. Presently the tomato linkage map is marked by more than 200 loci. Recently, in an interspecific backcross of tomato QTL mapping was done. At least six QTLs controlling fruit mass, four QTLs for the concentrations of soluble solids and five QTLs for fruit pH have been mapped²².

(iii) *Rice*—Tanksley *et al.*²⁰ have developed a linkage map of rice. Approximately 200 clones were evaluated for the copy number based on probing of genomic Southern. The Cornell RFLP map of rice now consists of 123 markers which are grouped into linkage groups and positioned on specific chromosomes. All 12 rice chromosomes are now marked by more than one RFLP marker.

Application of RFLP

It is clear that RFLP analysis is being employed extensively for some of the cultivated plants. The RFLP analysis differs with respect to cross pollinated and self pollinated species. In the outcrossed species, linkage equilibrium is seen, i.e. they carry among them all possible combinations of the marker and alleles segregating in the population, in proportion to the relative frequencies of these alleles in the population as a whole. As a result in outcrossed species, alternative marker genotypes will indifferently include both positive and negative alleles at any linked loci, and the mean quantitative value of the alternative marker genotypes will not differ even when a linked segregating locus is present in the population. Thus, for marker-allele analysis in

cross-pollinated species, each individual must be evaluated separately by raising numerous offsprings out of a single individual, as against the self pollinated species.

The RFLP technology can be effectively used in the mapping and manipulation of the quantitative trait loci (QTLs)²¹. Many agriculturally important traits like yield, quality and horizontal resistances to various stresses are quantitatively inherited. Manipulation of QTLs is a major limitation of both classical breeding and genetic engineering. RFLP methodology provides three pathways to the identification and manipulation of QTLs. These are (i) determination of RFLP-QTL linkage, subsequent mapping and evaluation of QTL effects, (ii) direct identification of allelic variation at a QTL and (iii) identification and cloning of QTL by means of insertional mutagenesis.

RFLP technology undoubtedly helps in detecting genes underlying quantitative variation but as in conventional programs the danger of inadvertently selecting undesirable characters is seen.

Earlier reports reveal that portions of the genome of *Lycopersicon chmielewskii* introgressed into the cultivated species are reported to increase soluble solids in fruits of certain tomato varieties. Recently, two of the three introgressed segments have been identified with RFLP markers and tested for effects on soluble solids in a single F_2 population²³. Two of these three segments were found to increase soluble solids; however, the effect of one of these was dependent on the genetic background. Both segments were found associated with deleterious characters including increase in fruit pH, lower yield and small fruits²⁴. These results confirm the utility of RFLP markers in detecting QTLs but demonstrate the danger in establishing breeding programs around such linkages until the effects of the quantitative genes have been tested in a variety of genetic backgrounds and for associated effects on other characters of agronomic importance.

Apart from the above applications, RFLP analysis provides new insights into organization of genome and stability of gene order within and between species. Such comparative gene mapping may help to generate information on genetic and chromosomal evolution and on phylogenetic relationships among taxa. RFLP also helps in the analysis of copy number, linkage organization and the complexity of closely related and multicopy genes, which are difficult to assess by conventional genetic techniques. It is possible to foresee greater applications of the RFLP technique in the field of plant breeding, especially in the light of the availability of new tech-

niques such as polymerase chain reaction (PCR)²⁵, field alteration gel electrophoresis²⁶, pulsed field gradient electrophoresis²⁷ and related techniques and the development of stable non-radioactive probes.

The PCR is an *in vitro* method for the primer directed enzymatic amplification of specific DNA sequences. This technique is capable of synthesizing over a million copies of a specific target DNA sequence in a few hours, significantly facilitating all subsequent analytic procedures.

In genetic analysis after a specific locus is amplified, a variety of techniques can be used to distinguish allelic variants. Short allele specific oligonucleotide (ASO) probes that hybridize differentially to normal and mutant or polymorphic sequences represent a more general approach to genetic diagnosis or typing²⁸. Allelic variation has been determined by using primers specific to conserved sequences to amplify polymorphic regions and then cloning the PCR products into M13 vectors for sequencing.

The development of techniques for pulsed field gel electrophoresis (PFGE) makes it possible to separate DNA fragments which are 20 kb to several thousand kb in length. The DNA is cut with a rarely cutting enzyme and subjected to PFGE to separate large fragments. In addition, this technique can be used to refine the genetic map and allow rapid determination of the map location of new genes or mutations. The PFGE can also be effectively utilized to determine the position of genes within a restriction fragment, the size of chromosomal deletions and map locations of cloned genes.

References

- Beckmann J S & Soller M, (1986), *Oxford Surveys of Plant Molecular and Cell Biology*, (B J Meflin, ed), Vol. 3, Oxford University Press, Oxford.
- Murray J C, Mills K A, Demopoulos C M, Hornung S & Motulsky A G, (1984), *Proc Nat Acad Sci (USA)*, 81, 3486-3490.
- Grodzicker T, Williams J, Sharp P & Sambrook J, (1974), *Cold Spring Harbor Symp Quant Biol*, 39, 439-446.
- Kan Y W & Dozy A M, (1978), *Lancet*, 2, 910-912.
- Botstein D, White R L, Skolnick M & Davis R W, (1980), *Am J Hum Genet*, 32, 314-331.
- Burr B, Evola S V, Burr F A & Beckmann J S, (1983), *Genetic Engineering: Principles and Methods*, (J K Setlow & A Hollaender, eds), Vol. 5, pp. 45-59, Plenum Press, New York.
- Davies K E, Young B D, Elles R G, Hill M E & Williamson R, (1981), *Nature, London*, 293, 374-376.
- Landry B S & Maicheltmore R W, (1987), *Tailoring Genes for Crop Improvement*, (G Breuning, J Harada, T Kosuge & A Hollaender, eds) pp. 25-44, Plenum Publishing Corporation.
- White R, Leppert M, Bishop T, Barker D, Berkowitz J, Brown C, Callahan P, Holm T & Jerominski L, (1985), *Nature*, 313, 101-105.
- Saghai-Maroo, Soliman M A, Jorgensen R A & Allard R W, (1984), *Proc Natl Acad Sci (USA)*, 81, 8014-8018.
- McClintock B, (1984), *Science*, 226, 792-801.
- Scott N S & Timmis J N, (1984), *Theor Appl Genet*, 67, 279-288.
- Southern E M, (1975), *J Mol Biol*, 98, 503-517.
- Helentjaris T, Slocum M, Wright S, Schaefer S & Nienhuis J, (1986), *Theor Appl Genet*, 72, 761-769.
- Helentjaris T, Weber D F & Wright S, (1986), *Proc Natl Acad Sci (USA)*, 83, 6035-6039.
- Helentjaris T, Weber D & Wright S, (1988), *Genetics*, 118, 353-363.
- Osborn T C, Alexander D C & Fobes J F, (1987), *Theor Appl Genet*, 73, 350-356.
- Bernatzky R & Tanksley S D, (1986), *Genetics*, 112, 887-898.
- Polans N O, Weeden N F & Thompson W S, (1985), *Proc Natl Acad Sci (USA)*, 82, 5083-5087.
- McCouch S R, Kochert G, Yu Z H, Wang Z Y, Khush G S, Coffman W R & Tanksley S D, (1988), *Theor Appl Genet*, 76, 815-829.
- Song K M, Osborn T C & Williams P H, (1988), *Theor Appl Genet*, 75, 784-794.
- Andrew H P, Lander E S, Hewitt D J, Peterson S, Lincoln E S & Tanksley S D, (1988), *Nature, Lond*, 335, 721-726.
- Osborn T C, Alexander D C & Fobes F J, (1987), *Theor Appl Genet*, 73, 350-356.
- Tanksley S D & Hewitt J, (1988), *Theor Appl Genet*, 75, 811-823.
- Saiki R K, Scharf S, Faloona F, Mullis K B, Horn G T, Erlich H A & Arnheim N, (1985), *Science*, 230, 1350-1354.
- Carle G F, Frank M & Oslon M V, (1986), *Science*, 232, 65-68.
- Smith C L, Econome J G, Schutt A, Kleo S & Cantor C R, (1987) *Science*, 236, 1448-1453.
- Erlich H A, Gelfand D H & Saiki R K, (1988), *Nature, Lond*, 331, 461-462.