A DETAILED GENOMIC ANALYSIS OF A HIGHER PLANT WITH SPECIAL REFERENCE TO RICE

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

BY

MADHU SHIVSHANKAR DHAR
M.Sc.

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

1989

DEDICATED TO MY
PARENTS-IN-LAW AND PARENTS

C O N T E N T S

CHAPTER 2: ORGANIZATION OF RICE GENOME AT A DNA FRAGMENT LENGTH OF 20 kbp

CHAPTER 3: ORGANIZATION AND PROPERTIES OF REPEATED DNA SEQAUENCES IN RICE

CHAPTER 4: MOLECULAR CLONING, RESTRICTION ENZYME ANALYSIS AND SEQUENCING OF A SPECIFIC LONG REPETITIVE DNA SEQUENCE IN RICE

A C K N O W L E D G E M E N T

I avail this opportunity to express my deep sense of gratitude to my teacher and guide Dr. P.K. Ranjekar, whose constant help, guidance and encouragement throughout my research work has resulted in the form of this thesis. His advice and critical comments should stand me in good stead in the years to come.

The magnanimous help given by Vidya at all stages of my research work cannot be expressed in words. Her constant help and valuable advice throughout have contributed largely towards the progress and success of my work.

Help rendered by Medha, Vaiju, B.G. Patil and Shubhada in carrying out certain experiments with me was very much needed and I duly acknowledge them for their whole hearted cooperation.

My friends and colleagues, Shirish, Meena, Kalpana, Madhumalti, Yogesh, Vivek, Gagan, Nirmala, Lalitha, Hendre, Prafulla and Suman have all contributed immensely in their own way towards the successful completion of this thesis. This is only a token of appreciation for them.

I gratefully acknowledge the technical assistance given by Mr. S.T. Dhume and Mr. M.V. Kamthe in handling the various instruments.

Sincere thanks are also due to Dr. M.V. Hegde (Poona University) and Joyce and Asha for their help by extending the facilities of their lab to learn few techniques required for my work.

I am also thankful to Drs. M.C. Srinivasan and John Barnabas for their encouragement and help rendered in their capacity as Head, Division of Biochemical Sciences and the Director, NCL

for his kind permission to submit this work in the form of a 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.

I offer my deepest gratitude towards all my family members for their continuous help and encouragement during the course of my work, my brother-in-law, B.S. Dhar for his neat and skilful typing. Finally, a special mention must be made of my husband whose constant encouragement, patience and support throughout were the major factors contributing to the successful completion of my thesis.

Madhu

Madhu

D E C L A R A T I O N

Certified that the work incorporated in the thesis " A detailed genomic analysis of a higher plant with special reference to rice", submitted by Mrs. Madhu S. Dhar 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

(P.K. Ranjekar)

Research Guide

LIST OF ABBREVIATIONS

kbp	Kilo base pairs
R-24	Ratnagiri-24
%	percentage
cm	centimeter
^{o}C	degrees centigrade
AR	analytical reagent
GR	guaranteed reagent
M	molar
rpm	revolutions per minute
min	minute
v/v	volume by volume
ug	microgram
ml	millilitre
mg	milligram
g	gram
nm	nanometer
A	absorbance
M.s	Mole. Second. Lit^{-1}
PB	phosphate buffer
U	units of enzyme
cv	cultivar
mA	milliampere
dATP	deoxyribose adenosine 5'-triphosphate

Cam	Chloramphenicol
Amp	Ampicillin
Tet	Tetracycline
psi	pounds per square inch
E. coli	Escherichia coli
O.D.	optical density
DIGE	Direct gel electrophoresis
DTT	dithiothritol
dCTP	deoxyribose cytidine 5'-triphosphate
dGTP	deoxyribose guanosine 5'-triphosphate
TTP	thymidine 5'-triphosphate
ddATP	dideoxyribose adenosine 5'-triphosphate
ddCTP	dideoxyribose guanosine triphosphate
ddTTP	dideoxyribose thymidine triphosphate
ssDNA	single stranded DNA

LIST OF FIGURES

LIST OF TABLES

SYNOPSIS

SYNOPSIS

The plant molecular biology group at NCL initiated the work on rice genome as early as 1978, mainly because very little data on rice DNA was available at that time. In fact, Deshpande and Ranjekar's report in 1980 [1] was probably the first one in which they showed that the repetitive DNA content in rice was 50%. Later, Gupta et al [2] reported that these DNA sequences remained uninterspersed with single copy DNA sequences at a DNA fragment length of 6.5 kbp. This was considered to be a very unique observation at that time, in view of the widespread occurrence of interspersion of repeat and single copy DNA sequences in all the plant genomes studied till then.

In the present work an indepth analysis of rice repetitive DNA sequences has been carried out with a main objective of gaining more information about their distribution and organization in different rice cultivars and in two different tissue systems mainly embryo and shoot. Some of the important results obtained during the study are summarised below.

1. NOVEL DNA SEQUENCE ORGANIZATION IN RICE GENOME

In order to find out if the unique genome organization pattern of rice (Oryza sativa cv. Basmati 370) observed at a DNA fragment length of 6.5 kbp persisted even at a higher DNA fragment length, reassociation of high molecular weight rice

DNA (20 kbp) was studied at a limit Cot value of 50. It was found that the percentage of reassociation of rice DNA was 56% at this Cot value as compared to 52% at a fragment size of 6.5 kbp. Thus, there was no significant increase in reassociation with an increase in DNA fragment length. The decrease in percent reassociation after S1 nuclease was from 56 to 48% in case of 20 kbp rice DNA as compared to 52 to 48% in case of 6.5 kbp. Hence, it was concluded that repeated and single copy DNA sequences were uninterspersed in rice genome even at a high DNA fragment length of 20 kbp.

DNA sequence organization was also examined in two other rice cultivars namely Ratnagiri-24 (R-24) and Sona and these cultivars again exhibited the lack of interspersion as in case of Basmati 370 at a DNA fragment length of 20 kbp.

2. HIGH THERMAL STABILITY OF RICE REPETITIVE DNA

Thermal stability of rice total DNA and two repetitive DNAs namely Cot 0.1 (highly repetitive) and Cot 50 (total repetitive) from the rice cultivar Basmati 370 was determined. As compared to total rice DNA (Tm=85.1^{o}C), the Cot 0.1 and Cot 50 DNAs had higher Tm values of 95.5^{o}C (G + C content=63%) and 88.9^{o}C (G + C content=47.8%), respectively.

When highly repetitive and total repetitive DNA fragments from rice (Basmati 370) were analyzed on agarose gels, they

revealed the presence of long (4-20 kbp) and short (0.2-0.3 kbp) DNA sequences. The isolated long and short repeats of Cot 0.1 DNA showed a monophasic curve with a Tm of 94.8°C and 92.9°C respectively. The long repeats of Cot 50 had a Tm of 97.5°C while its short repeats showed a biphasic curve with a major low melting fraction (Tm = 83.7°C) and a minor high melting fraction (Tm = 98°C).

3. PRESENCE OF DIFFERENT REPEAT FAMILIES IN RICE

To assess whether short and long repeats in rice consisted of same/similar type of repeat sequences or they were independent of each other, Southern hybridization experiments were done using long and short repeats of Cot 0.1 and Cot 50 DNAs as probes It was found that long and short repeats of Cot 0.1 were not homologous to each other, whereas the long and short repeat sequences reassociating by Cot 50 did show some homology. Based on these data, the following four types of repeat DNA sequences were identified :

(a) Short repeats (0.2-0.3 kbp) reassociating by Cot 0.1.

(b) Long repeats (9-20 kbp) reassociating by Cot 0.1.

(c) Short repeats (0.2-0.3 kbp) reassociating by Cot 50.

(d) Long repeats (9-20 kbp) reassociating by Cot 50.

4. MOLECULAR CLONING, RESTRICTION ENZYME ANALYSIS AND SEQUENCING OF RICE LONG REPETITIVE DNA SEQUENCES

Since a cloned probe of a specific repeat family will give more information about its organization and properties in the genome, the long repetitive sequences reassociating by Cot 50 were cloned in pBR325 at BamH1 site. The $Amp^{r}Cam^{r}Tet^{s}$ colonies were selected after transformation and out of several recombinants three (pRL1, pRL7 and pRL10) were selected randomly for further detailed characterization. The plasmids carrying the repetitive DNA sequences were digested with BamH1 and size of the insert was 3.4 kbp, 3.6 kbp and 3.2 kbp in pRL1, pRL7 and pRL10, respectively. Further restriction enzyme analysis of the cloned DNA sequences revealed that there were no sites for BclI and EcoRI, single site for PvuII and PstI and abundance of AluI cutting sites.

The BamH1-PstI fragment (about 0.4 kbp) in pRL7 was subcloned in M13mp18 and was sequenced. The partial sequence obtained showed an appreciable amount of homology with the nucleotide sequence of 25S.rDNA fragment.

5. PREDOMINANCE AND TISSUE SPECIFICITY OF ADENINE METHYLATION IN RICE

Analysis of total rice (Oryza sativa cv. Basmati 370) DNA, repetitive DNA and a cloned repeat sequence using the 'A' and

and 'C' methylation specific enzymes namely MboI, Sau3AI, DpnI, MspI and HpaII indicated the abundance of adenine methylation in all the three systems. Studies on cytosine methylation in 5'-CCGG-3' sequences suggested more of CpC methylation than CpG; but in general, the 'C' methylation 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 and CpC dinucleotides in these two tissue systems.

Restriction enzyme digestion patterns of the three cultivars under study indicated differences in specific repeat DNA sequences. For instance, MboI and Sau3AI repeat families were present in Sona shoot DNA while they were absent in Basmati 370 and R-24. Similarly, MspI and HpaII repeat families were observed in Basmati 370 embryo DNA only.

REFERENCES

1. Deshpande, V.G. and Ranjekar, P.K. (1980) Hoppe-Seyler's Z. Physiol. Chem., 361 : 1223-1233.

2. Gupta, V.S., Gadre, S.R. and Ranjekar, P.K. (1981) Biochim. Biophys. Acta, 656 : 147-154.

CHAPTER 1

REPEATED DNA SEQUENCES IN PLANTS GENOME ORGANIZATION, TYPES FUNCTION AND EVOLUTION.

1.1 TYPES OF DNA SEQUENCES

In the mid 1960's, the pioneering work of the Carnegie group has established that all eukaryotes, including plants, have a large proportion of their DNA as repeated sequences [1]. This discovery has led to extensive investigation on various aspects such as structure, function, organization, divergence and evolution of these sequences in animals and plants.

On the basis of reassociation kinetic experiments, the eukaryotic DNA can be divided into three broad classes [2].

(i) Unique or single copy or a few copy sequences : These are present approximately as one copy per haploid genome. Certain genes that code for proteins such as seed storage proteins [3] and wheat histone H_4 genes [4] as well as genes for rDNA, 5S RNA and tRNA are present in multiple copies.

(ii) Moderately repeated sequences : These are present in approximately 1000 to 100,000 copies in all the cell and can be resolved into two main types, long and short, depending upon their size. Some of these sequences have been shown to be transcribed [2].

(iii) Highly repeated sequences : These sequences consist of clustered repetitions of relatively short sequence units. They are located mainly at centromeric and telomeric

regions [2] and are considered to be involved in the structural organization of chromosomes.

1.2 GENOME ORGANIZATION OF REPEAT SEQUENCES

Earlier studies on several animal DNAs have revealed two general patterns on genome organization namely (i) 'short period' interspersion pattern in which 100-300 bps long repeated sequences are interspersed with longer, single-copy sequences of approximately 1000-2000 bps [5] and (ii) 'long period' interspersion pattern in which approximately 5000 bp long repeated sequences are interspersed with single copy sequences that may be as long as 35 kbp [6-11]. As compared to the interspersion data in animals, less information is available in plants. Studies from our laboratory and elsewhere have revealed that the size of interspersed single-copy DNA sequences in plants varies in a narrow range of 1000-2000 bps with the exception of Vigna radiata, Gossypium hirsutum and Pisum sativum, where other size classes are present. The interspersed repeat sequences, on the other hand, have a wide variety of size classes such as 50-100, 300-1000, 1250-2000 and 4000-4500 bps. Oryza sativa, Lilium usitatissimum, Cucumis sativus and Arabidopsis thaliana are rather unique in exhibiting a lack of interspersion at a DNA fragment length of 20 kbp, 10 kbp, 5 kbp and 65 kbp respectively and show the presence of very long repetitive stretches [12-16]. This suggests that the classification of interspersion as simply 'short period' or 'long period'

or 'intermediate' is inaccurate in case of plants and that it (1C < 5 pg) would be more appropriate to state that plant genomes essentially have a 'mixed' type of interspersion pattern [17] (Table - 1.1).

Another unique feature of genome organization in plants is the relationship between nuclear DNA content and the type of interspersion pattern that is present in the plant. Plants with a genome size (1C) smaller than 5 pg show predominance of long period and mixed interspersion pattern, while genomes having 1C nuclear DNA amount greater than 5 pg show the predominance of short period interspersion pattern. This suggests that nuclear DNA content plays an important role in determining the overall pattern of DNA sequence arrangement in plants and interspersion may thus be considered as a nucleotypic effect [41,42].

A regulatory importance was assigned to the short interspersion pattern [43-48]. In the light of the occurrence of diversity of genome organization patterns in plants, it has been suggested that there is a pre-programming in each plant which plays an important role in gene regulation [17]. Pre-programming refers to a specific arrangement or distribution of different types of DNA sequences in a given genome from the functional point of view.

1.3 Homogeneous and heterogeneous families

To understand chromosomal functions at the molecular level,

TABLE 1.1 : GENOME ORGANIZATION IN PLANTS

Species	Haploid DNA content (pg)	Total repetitive DNA %	Length of interspersed sequences (bps)		Pattern of interspersion	References
			Repetitive	Single copy		
1	2	3	4	5	6	7
Arabidopsis thaliana	0.60	5.0-6.0	-	-	Absent (extremely long)	15
Vigna radiata	0.50	35.0	300-1200;1200-6700	750; 3950	Intermediate	18
Linum usitatissimum	0.60	56.0	10,000	-	Long	13
Oryza sativa	0.60	52.0	50-100;6000-64000	-	Absent	12
Setaria italica	0.80	30.0	1500	2700	Long	19
Gossypium hirsutum	0.80	38.0	1250	1800; 4000	Intermediate	20
Cajanus cajan	0.90	26.0	500;500-1500 1500	Not determined	Mixed	21
Panicum miliare	1.00	38.0	150-250;1500	1700	Mixed	19
Cucumis sativus	1.10	37.0	-	-	Absent	14
Tricosanthes anguina	1.10	65.0	300;1500	2000; 2000-5000	Mixed	22

1	2	3	4	5	6	7
Cucurbita maxima	Not determined	53.0	300;1500	1800; 1800-5000	Mixed	22
Matthiola incana	1.50	60.0	350-400	1050	Short	23
Eleusine coracana	1.60	49.0	150-200;4000-4200	1900	Mixed	24
Luffa cylindrica	1.75	51.0	300;300-1500; 2000-4000	2100	Mixed	25
Benincasa hispida	1.75	48.0	300;300-1500; 2000-4000	2100	Mixed	25
Vigna mungo	1.76	40.0	500;500-1500;1500	Not determined	Mixed	26
Phaseolus aureus	1.80	47.0	1900	1300	Intermediate	27
Phaseolus vulgaris	1.85	40.0	1545	1300	Intermediate	27
Petroselium sativum	2.00	70.0	300-400	1000	Mixed	28

1	2	3	4	5	6	7
Vigna aconitifolia	2.05	38.0	500;500-1500;1500	Not determined	Mixed	29
Pennisetum americanum	2.5	54.0	4300-4500	1900	Long	30,31
Echinochloa fumentacea	2.70	42.0	100-1000;1500	2500	Mixed	19
Luffa acutangula	2.75	59.0	300;300-1500;2000-4000	2900	Mixed	25
Coccinia indica	2.75	25.0	300;300-1500; 2000-4000	2300	Mixed	25
Glycine max	3.25	67.0	300-400;1500	1235	Short	32,33
Nicotiana tabacum	3.90	70.0	300;1500	1400	Short	34
Sorghum vulgare	4.60	52.0	150-300;500-1000; 1500	2800	Mixed	35
Pisum sativum	4.95	70.0	300-400	300-1000	Short	36
Zea mays	7.70	58.0	500-1000;5000	2100	Short	37

1	2	3	4	5	6	7
Secale cereale	9.00	73.5	300-800	1500	Short	38
Allium cepa	16.75	84.8	300	1000	Short	39
Triticum aestivum	18.10	75.0	300-400	1500	Short	40

we need to know the characteristics of the various types of nucleotide sequences found in DNA. Individual sequences whether unique (single copy) or repetitive may play an important role only when interacting with the other sequences in the genome or by influencing the activity of other sequences in the cell. Thus in order to draw biologically meaningful conclusions about the sequence organization in the DNA molecules, it is essential to have biologically meaningful assays for distinguishing nucleotide sequences. One such assay developed by Bendich and Anderson [49] and Priesler and Thompson [50] classifies repeat sequences into homogeneous and heterogeneous families by comparison of reassociation rates of DNAs at different temperatures. In their terms, a 'family' is defined as a set of copies of a sequence which are similar enough to reassociate under a given set of conditions. According to Bendich and Anderson, a homogeneous family contains member sequences of the same similarity, e.g. in one family all members are related by say 70% homology, in a second by 85% homology and so on. A heterogeneous family, on the other hand, contains member sequences of varying similarity i.e. ranging from nearly perfect replicas to sequences barely similar enough to reassociate. The homogeneous or heterogeneous families may, however, contain several hundreds or thousands of repeat sequences exhibiting similar reassociation kinetics. This assay cannot distinguish each family or sequences from the other but would give an idea about the evolutionary state of the repeat DNA sequences and in turn the 'genome turn over' rate.

Using this assay, these authors have concluded that the repeat families in parsley, fern, barley, daffodil and deer fern are homogeneous [49]. Similar studies conducted by Priesler and Thompson on pea and mung bean have shown that the repeat families in mung bean are mostly homogeneous while those in pea are heterogeneous [50]. Murray et al, have also determined the percentage of 'true' single copy and fossil repeats in the pea and mung bean genomes and have concluded that the genome of mung bean is turning over at a rate slower than that of pea. These authors have further correlated the rates of genome turn over with the pattern of interspersion observed and have suggested that a rapid turnover would produce extensive short period interspersion as is seen in pea, while a slow turnover rate would result in an increase in the average interspersion difference as observed in mung bean [51].

Higher plant genomes undergo structural changes during plant development as well as over the evolutionary time periods. In the past few years, a few examples are available wherein any one type of nucleotide sequence/family of repeat sequences has been studied in sufficient depth to understand its role or function. Classifying DNA sequences by copy number has certain drawbacks because in renaturation kinetics experiments, the sequences in reaction are heterogeneous mixtures and contain sequences from different repeat families. Hence these type of experiments do not provide adequate information on the fine structure, organization and evolution of nuclear DNA.

1.4 TANDEM AND DISPERSED REPEAT FAMILIES

In recent years restriction enzymes have served as extremely powerful tools in (i) identification of repeated sequence families (ii) detection of sequence-specific methylation, (iii) localization of specific gene fragments and (iv) analyzing the overall distribution of base composition in total DNAs as well as in the vicinities of specific DNA sequences.

Based on restriction enzyme analysis, the following two types of repeatfamilies have been identified [2].

(i) Tandem repeats

These repeats appear as a ladder of bands on the gel after partial digestion with restriction endonuclease. The various bands in the ladder have molecular weights which are multiples of a basic monomer unit. Tandem repeats include satellite DNA sequences and repeat units coding for 25 S + 18 S rDNA or 5S rRNA. The highly repeated sequences are most often arranged in tandem arrays and are sometimes referred to as satellite DNA. Such sequences have been studied in a number of animal and plant species [52,53], but their function if any, is not yet knowm and remains a subject of speculation. Cryptic satellite DNAs have also been identified in a number of plant species [54-65], and some of these satellite sequences have been characterised and sequenced (Table 1.2).

TABLE 1.2 : TANDEM REPEAT FAMILIES IN HIGHER PLANTS

Plant species	Enzyme	Remarks	Reference
Arabidopsis thaliana	*HindIII* or *AluI*	Tandemly arranged long arrays of 180 kbp element. About 4000 to 6000 copies per haploid genome.	66
Citrus limon	*BspI*	Many bands representing high periodicity in the satellite DNA.	57
Cucumis melo	*HindIII* or *AluI*	Family with basic unit of 380 bps occurring in high copy number. All copies however, are not clustered.	62
Cucumis sativus	*AluI*	Family of elements of 180 bps occurring in high copy number. All copies however are not clustered.	62
Cucurbita pepo	*HindIII*	Tandemly arranged family of 351 bps elements present in about 1.3×10^5 copies. constituting 4-8% of the genome.	67
Raphanus sativus	*HindIII*	Tandemly repeated array of 177 bps unit from satellite DNA which has structural feature common to alphoid satellite Present in 0.6 million copies.	68
Secale cereale	*HaeIII*	Four different families with repeat units 120, 480, 610 and 630 bps in length. First three are exclusively present in telomeric hetero-chromatin.	56

Plant species	Enzymes	Remarks	Reference
Sinapsis alba	AluI HaeIII and HinfI	Basic unit of 172 bps arranged tandemly to form a satellite.	64
Vicia faba	FokI	Tandem array of elements 59 bps in length. Estimated copy number upto $5x10^6$ to 10^7 per diploid genome	65
	MluI	Tandem array of 325 bps fragments showing considerable homology among members.	69
Zea mays	HaeIII	Basic unit of 185 bps tandemly arranged present upto 10^6 copies. Limited only to knob heterochromatin.	59
Brassica deracea	HindIII	Basic unit of 177 bps. Tandemly arranged which show upto 80% homology with other satellite DNA sequences of Crucifereae and also shows more than 50% homology with many tRNA genes.	70
17 species of family Brassicaceae	HindIII	Basic units of 175 bps are tandemly arranged. The Brassica napus monomer hybridized (i) under stringent conditions with all the seven Brassica species as well as related genera such as Eruca and Capsella (ii) under moderate conditions with the genera Raphanus Sinapsis and Crambe (iii) under relaxed conditions with Raphanus, Sinapsis and Crambe.	70

Eukaryotic ribosomal RNA genes are organized as families of tandemly repeated genes, some or all of which may comprise the nucleolar organiser regions of chromosomes. Each repeat unit of rDNA contains a single transcription unit and a non-transcribed intergenic spacer (IGS). The IGS region of each repeat unit consists of an array of tandemly repeated sequences referred to as subrepeats which are 100-300 bp in length. These studies show that the 25S, 18S ribosomal RNA genes have been studied with respect to their general organization, variation in copy number and variation between different species in many plants like wheat, rice, maize, flax, petunia, pumpkin, radish, broad bean, and soyabean [65, 71-87]. Variation in the number and sequence of these repeats seems to be responsible for the rDNA variability. The size of all these rDNA units varies between 8-12 kbp depending on the species and their copy number is also variable (700-15000 copies/diploid genome).

(ii) Dispersed repeats

These repeats do not appear as a ladder of bands on partial digestion of DNA with restriction enzymes but are interspersed either with single copy sequences or with other repetitive sequences in heterogeneous and non-tandem clusters. Compared to the information available on tandem repeats in plants, relatively less number of dispersed repeat families have been studied with respect to their detailed organization in the genome. Studies

on animal genomes distinguish short interspersed elements (SINES) less than 500 bp from long interspersed elements (LINES) over 5 kbp in length [53]. In higher plants, dispersed repeat families such as AluI, BamH1, HindIII, HaeIII and FokI have been identified (Table 1.3), but so far very few have been characterised in detail [95]. The function of these sequences in plants is not known, although it was postulated earlier that they might be involved in regulatory processes [96]. In animals, however, they have defined roles [97].

Recently, there has been a report on the characterization of the EcoRI repeat family (1072 bps) in Lupinus luteus L. [90]. Some interesting observations have been reported. Sequence analysis showed that 33.4% of its total length was represented by scattered simple direct repeats. Computer studies showed the presence of three promoter regions for RNA polymerase III. This polymerase is believed to be directed by two internal control regions termed Box A and B (consensus sequence : PuPuPyNNAPuPyGG and GATCPuANNC) located 31 to 74 nucleotides away from each other. Control regions A and B found on this EcoRI insert deviated from the consensus sequences only by one nucleotide. Such internal promoters for RNA polymerase III have been shown to be derived from ancestral tRNA genes [98] and are found to be present in the Alu sequences in primates and rodents [99].

Transposable elements have also been shown to belong to dispersed repetitive sequences but their copy number is usually

TABLE 1.3 : DISPERSEDREPEAT FAMILIES IN HIGHER PLANTS

Plant Species	Enzyme	Remarks	Reference
Allium cepa	BamH1	Dispersed family with a basic repeat unit of 1000 bps. 8.0×10^4 to 10^6 copies per haploid genome.	88
Lilium henryi	BamH1	Dispersed family with two repeats of 2000 bps and 5000 bps lying adjacent and sharing a common BamH1 site Present in all the 12 chromosomes.	89
Lupinus luteus	EcoRI	Dispersed family of highly repeated elements about 1071 bp in length.	90
Vicia faba	BamH1	Several families ranging from 250 bps to 2400 bps. None of them organized in a single tandem array. Show high sequence heterogeneity among members.	91
Zea mays	HpaII	Family of dispersed elements with a basic repeat unit of 300-1300 bps	92
Cucurbita maxima Trichosanthes anguina Cucumis sativus Cajanus cajan Phaseolus vulgaris	MboI	Families of dispersed elements with sizes in the range of 0.5 to 5 kbp.	93
Vigna radiata	AluI	The members are extensively interspersed with some clustering. They also occur extrachromosomally. There is no tissue specificity but this family is highly species specific. The cloned repeat shows a number of direct and inverted repeats and some short palindromic sequences.	94

much lower. Actually it is now assumed that many of the dispersed repetitive sequences have spread in the genome by transposition-like events [100].

1.5 TRANSPOSABLE ELEMENTS

Early evidences for transposable elements produced by maize geneticists [101] have now been confirmed by molecular analyses in bacteria, yeast, Drosophila, maize and snapdragon. Transposable elements (mobile genetic elements, insertional elements or transposons) are specific DNA fragments normally present in a genome which are able to move from one site to another, either on the same chromosome or a different one.

Only a few plant transposable elements have been isolated and characterized (Table 1.4). A characteristic feature of transposons is the presence of terminal inverted repeats which are supposed to be the substrate for transposase activity, flanked by direct duplication of the target site. Transposition of mobile genetic elements generates mutations or chromosomal rearrangements and thus affects gene expression. According to the Saedler and Never's model [120], transposition in higher plants involves excision of a transposable element from one position in the genome and reintegration of the same element at another site. Most often, excision is not precise and the target sequence is replaced by an imperfect duplication generating polymorphism [120].

TABLE 1.4 : TRANSPOSABLE ELEMENTS IN PLANTS

Plants	Transposon	Size in kbp	Reference
Maize	Ac/Ds	4.563	102,103
(*Zea mays*)		4.500	104
		4.369	104-107
	Ds	2.000	105
	Ds1	0.405	108
	Ds2	2.040	104,109
			110
	Ds	1.300	111
	Mu1	1.367	112
	Mu1	1.400	113
	En(Spm)	8.400	114,115
	En	8.400	113
	I	2.242	116
	GnI	0.700	117
	BsI	3.500	100
Snapdragon	Tam1	17.000	100,118
(*Antirrhimum majus*)	Tam2	5.000	100,118
	Tam3	3.500	100,118
Soybean	Tgm1	3.550	119
(*Glycine max*)			

1.6 FUNCTIONS OF REPEATED DNA SEQUENCES

Based on the extensive work on repeated DNA sequences, several roles have been assigned to them and some of these are described in the following paragraphs.

(i) Role in variation in nuclear DNA content

Variation in the nuclear DNA content between species is immense [121]. Even within certain genera, the DNA content can vary upto 10 fold [122,123]. It cannot, however, be stated precisely whether the changes in the DNA content are restricted to repetitive or unique DNA sequences alone. In order to ascertain the nature of these changes, Hutchinson *et al* [124] determined correlations between the amounts of repetitive and nonrepetitive DNAs and total DNA contents and calculated the ratio of increase in repetitive DNA to the increase in non-repetitive DNA. This ratio was 5.53 and 3.91 in species belonging to the genera *Lolium* and *Lathyrus* respectively [121]. Flavell *et al* [125] have shown that plant species with DNA content above 5 pg have more repetitive DNA than those with DNA content less than 5 pg. Reports on the DNA content of millets from our laboratory [126] have also shown that the proportion of repetitive DNA varies linearly with the DNA content. This is true for the single copy DNA also but the slopes of the two lines are different. Above a nuclear DNA amount of 6 pg, the content of single copy DNA remains more or less constant while that of repetitive DNA continues to incre-

ase. Similar observations have also been made in species belonging to families Compositae [124,127], genera like Anemone [124,128] and Allium [124,129] and a wide range of higher plants [124,130]. Hence it appears that repeated DNA contributes significantly to the genome size in plants.

(ii) Role in speciation

Mechanisms that initiate speciation may involve many different kinds of mutations or genomic changes in repetitive DNA sequences. Such mutations may create pre-mating or post-mating barriers [131]. In repeat sequences, mutations are known that prevent the formation of fertile hybrids between species such as between wheat and rye. However, these mutations are not present in all the isolates of the species. It is likely that they could have arisen after speciation or they could have caused speciation but were lost following the accumulation of other pre- or post-mating barriers that have kept the species apart. Differences in repeated sequences may result in

(a) reduced fitness of hybrids due to failure in chromosome pairing at meiosis

(b) variation in gene expression due to changes in the families of repeated sequences, though such examples are yet to be established.

(c) variation in chromosome behaviour. For example in a wheat x hybrid, it has been suggested that the heterochromatin

on rye chromosome prevents it from completing the replication cycle rapidly during endosperm development to permit normal endosperm development [132]. Such species specific repeat families and their probable role in mitotic pairing have been identified and characterized in S. cereale and wheat [56,133].

The quantitative changes observed in the nuclear DNA amounts have also been implicated to have a role in speciation [124].

(iii) Role as transposable elements

In mammals, the short interspersed AluI and AluI equivalent repeats and long interspersed repeat elements are the two most abundant and best characterized middle repetitive sequence families [134-137]. Recent studies have revealed a highly significant (about 80%) sequence homology between the longer unit of the human AluI consensus sequence and the 5' and 3' portions of 7SL RNA. The latter is the abundant cytoplasmic RNA, 300 bases in length and forms a part of the signal recognition particle [138]. It has been suggested that the AluI repeats are processed 7SL RNA transripts containing 3' poly(A) segments which may have provided the template for their reverse transcription into DNA prior to genomic integration.

Data on the LI (long interspersed repeat family) of dispersed repeated sequences indicates that this family consists of largely of incomplete, probably non-functional, processed pseudogene-like copies together with a small number of transcriptionally

active elements [137,139,140]. Further studies have shown that these elements form part of RNA polymerase II directed transcripts of heterogeneous nuclear RNA in a number of somatic cell lines [141-142]. Similarities between LI elements and retro-transposons sequences include their association with DNA rearrangements and their occassional presence as extra chromosomal DNA [143]. Thus both the AluI repeats and the LI elements appear to be mobile. The LI elements possess potential internal open reading frames which may code for transposition functions and evidence points to their specific expression in early development.

In Drosophila, several families of dispersed repetitive elements have been identified which correspond to families of transposable genetic elements such as Copia-like sequences (Copia 412, 297, 17.6, mdg 1, mdg 3, B 104 [144,145], FB elements [146], Gypsy [147], P-elements [148], hobo [149] and I-factors [150].

(iv) Role in determining chromatin structure

"In situ" hybridization analysis performed in a number of monocot species using labelled satellite DNAs have revealed their association with heterochromatin regions of chromosomes including telomeres [56,61,151], centromeres [152,153], interstitial blocks [152], knob heterochromatin [59] and B chromosomes of maize [152].

The observation that satellite DNA and heterochromatin are drastically reduced in nuclei with somatic functions is consistent

with the idea that satellite DNA has some role in germ line processes i.e. meiotic chromosome pairing, recombination and evolutionary processes [154]. The widespread occurrence of highly repeated DNA sequences coupled with their high concentration in heterochromatin has led to a speculation on their roles in several cellular functions [44, 155-157]. Satellite DNA sequences seem to be transriptionally inactive, as no transcripts have been shown at different developmental stages so far investigated [62, 68]. Different methylation patterns of melon satellite DNA in hypocotyl and callus tissues have been demonstrated [158]. Since satellite DNAshave been generally localised in the constitutive heterochromatin regions of the chromosomes which in turn have been shown to influence the transcriptional activity of a gene transposed in their vicinity (position effect) [159], it is probable that such satellite DNA sequences may play a role in the control of gene expression [160,161].

The presence of satellite DNAs in several plant species is reminescent of the alphoid satellite sequences that occur in many mammals and in Drosophila [52]. Comparative data among plants and animals have shown that such sequences are related in primary structure (upto 50% homology) and seem to be derived from a common monomer ancestor about 60 bps long [52,68,162]. The sizes of repeats (monomers or dimers) are quite similar; 59 bp in V. faba [65], 180 bps in maize [59],&in radish [68], 177 bps in cauliflower [162] and in melon [63], 172 bps in African

green monkey [169], 340 bps in man [163] & 185 bps and 370 bps in rat [164]. The widespread occurrence of these alphoid-like satellite DNAs among such distantly related organisms points to the fact that some regions of the alphoid DNA sequences may have some universal property, otherwise they would not have been conserved to such an extent. For example, they might be recognised at some stage of the cell cycle by a specific protein [165, 166] or they might participate in the organization of interphase nuclei [167]. There is a striking coincidence between the size of these repeat units and the length of the DNA in the mononucleosome, suggesting that such repeats have a role in determining chromatin structure [168-170]. In fact, a sequence dependent phasing of nucleosomes along repetitive DNA has been found in mouse satellite DNA [171] and in the African green monkey satellite [169]. Since these DNAs have a common size range and a related primary structure, it has been suggested that they should thereforehave a common origin. Recently, it has been demonstrated that small interspersed mammalian repeated DNA sequences (SINES) show a marked homology with several tRNA genes [98,172]. In plants also, these sequences have been shown to be able to be folded into a tRNA like sequence [162]. Such results support the idea of a tRNA gene ancestor for some tandemly repeated nuclear DNA sequences of higher plants and may probably be extended to all eukaryotic cells. So far, there has been only one report wherein Ray Wu *et al*, have shown that a moderately repetitive sequence of rice (*Oryza sativa*) which is tandem in nature is

transcribed and also shows homology with both tRNA and 5S RNA [173].

(v) Role of repetitive DNA sequences in differentiation

Molecular studies in tissues during development have clearly shown that the differences among specialized cells are due to a differential gene activity. In plants, variation in gene expression during cell development has been studied only recently [174-178]. Results of studies on tobacco, cotton and soybean have shown that the structural gene expression is correlated with developmental stages [135-137]. The work on DNA characterization to understand a molecular basis of cell differentiation has been carried out mainly in leaf and fruit in Cucumis melo and C. sativus [179], suspensor cells in Tropaeolum [180,181] and epicotyl cells in pea [182]. These studies indicate that most of the cells in higher plants have quantitatively different DNA in their nuclei. This is due to a differential DNA replication in which repetitive DNA is considered to be involved. Underreplication of repetitive DNA has been reported in a variety of tissues among different developmental stages [179-184]. As compared to the early studies, finer analysis have been undertaken using a cloned repetitive DNA sequence to assess genomic rearrangements/alterations occuring during differentiation process. Using such an approach, changes in the highly repeated DNA sequences in flax under tissue culture conditions are studied [185].

Similarly, amplification of nuclear DNA and changes in copy number are observed in rice during dedifferentiation [186,187].

1.7 EVOLUTION OF REPEATED DNA SEQUENCES

The plant genome is considered to be in a state of constant flux. These genomic changes appear to be of common occurrence in higher plants and as such may be of greater importance in plants than in animals. It has been proposed that genomic flexibility is an important aspect of the adaptability of plants to a changing environment [188]. These rapid changes in the genome involve both repetitive and single copy DNA sequences. However, repetitive DNAs are apparently changing more rapidly than single copy sequences. The evolutionary processes appear to tolerate more extensive changes in repetitive DNA than in single copy sequences. Thus, an estimate of changes in repetitive DNA is used as a parameter in understanding their role in speciation and phylogenetic processes [189,190].

The molecular cloning and sequencing of DNA fragments has allowed a detailed characterization of various sequence families belonging to all reiteration classes. Comparison of sequences of a repeat family within the species or with more or less distantly related species allows one to determine the type of changes that have occurred during evolution and gives some clues to the possible mechanisms which have been involved. In repeat sequences, the main changes which have been observed are point mutations,

insertions or deletions, duplications of short sequences, amplifications and variations in copy number. When these changes affect a member of a sequence family, polymorphism and variability are generated which in some cases can be recognised by digestion with restriction enzymes or more frequently remain silent, and can be detected only by sequencing.

Data on rDNA genes in wheat, maize and Triticaceae[72,77,78, 191-193], dispersed repeat sequences in several Allium species [88], satellite DNA in radish [68] and multigene families such as seed storage protein genes of soybean [194-196], genes encoding for the small subunit or ribulose 1,5-bisphosphate/carboxylase/ oxygenase of Lycopersicon esculentrim (tomato)[197] and the histone genes of corn [198,199] have shown that individual memebrs of a family of repeated sequences show a greater similarity within a species than between related species, suggesting thereby that the individual members of a repeated family within a species are not evolving independently but in a concerted manner [200-202]. This concerted evolution implies homogenization of the members of a repeated family. Various homogenization mechanisms such as amplification, unequal crossing over, gene conversion and transposition constitute the molecular drive of concerted evolution.

The existence of multiple copy families implies a genomic mechanism of amplification that may be either saltatory or cummulative through successive duplications. A family of repeated

sequences is probably the result of amplification of a sequence previously present in one or a few copies per genome. When a newly amplified family of repeated sequences is first stabilised in a genome, all the sequences are probably identical. However, accumulation of base changes renders such families heterogeneous. Sometimes a variant sequence is propagated when individual repeats or subsets of repeats are reamplified [56,154,203-206]. During such a reamplification process, the repeat under consideration is sometimes amplified along with a neighbouring single copy or unrelated repeated sequence thereby giving rise to compound repeating units [46,56,207]. Many of the amplification events are never detected because they are deleted and not fixed in the genome.

A few hypothesis such as rolling circle replication, replicative loop, Okazaki fragments, unequal crossing over, transposition and gene conversion have been put forth to explain the mechanisms of amplification of repeated DNA sequences [208,209]. Recently, it has been demonstrated that small interspersed mammalian repeated DNA sequences (SINES) [172] and several tandem families in plants show a marked homology with several tRNA genes. A recent hypothesis (retroposition mechanism) suggests that the transcription initiation boxes A and B from tRNA genes are involved in SINE sequence amplification [98,210]. It has been proposed that a tRNA ancestor gene may have been transcribed, then reverse transcribed and inserted randomly in the nuclear genome. As A

and B boxes lie inside the tRNA gene, the derived tRNA-like DNA sequences may be amplified as long as the A and B boxes are enough conserved. Recent experiments on Xenopus satellite (I), tortoise (Geoclemys reevessi) and newt (Cynops pyrrhogaster) have demonstrated that some monomers of a multigene family can still be used as templates by RNA polymerase III [96,161,172,211]. DNA sequence analysis in Drosophila and SINE families in Galago and in the mouse B families [98] has shown that these repeated DNA transcripts putatively possess a reverse transcription initiation site. All such data strongly suggest a retroposition mechanism for the SINE sequence amplification. Benslimane et al [162] have suggested that the parental tRNA gene monomer from which the tandem repeat family has arisen must have a short directly repeated DNA segment (FDR) at each end. In the first step, both strands are transcribed by RNA polymerase III. The FDR sequences then allow the linear annealed transcripts to be folded into a circular form. Reverse transcription of this circular form then gives the double stranded complementary DNA which in turn is used as a template for a rolling circle amplification process. The resulting tandemly repeated DNA sequence is then integrated into the chromosomal DNA.

Unequal crossing over is another mechanism which accounts for the reported changes in copy number. It occurs when two arrays of tandem repeats pair and recombine out of phase. This results both in recombination between repeats and variation in

copy number. Evidence for unequal crossing over has come from studies on the ribosomal gene non-transcribed spacer (NTS) variants in barley and Vicia faba [212,213], where it is most likely that atleast part of the variation in rDNA copy number arises from unequal crossing overs.

Gene conversion is another way to homogenize repeated DNA sequences and at the same time to increase their variability. This occurs through the formation of a heteroduplex between A and B sequences. In this event, gene B acts as a sequence donor and remains unaffected, while gene A sharing homology with B receives a block of B sequence and undergoes variation [214]. The heteroduplex formed is then resolved by mismatch repair, and as a result, one of the sequences has corrected a portion of the other, possibly creating a new sequence. Evidence for gene conversion has come from sequencing satellite DNA in radish [68] and the non-transcribed spacer region of wheat ribosomal genes [191].

The data obtained by Flavell et al [203] on cereal genomes has shown that repeated sequences of a family are more closely related to each other within a species than they are to the repeated sequences of the corresponding family in another species. They, therefore, suggest that a single amplification event is not sufficient to explain the above observations, but, amplification of an already amplified sequence has been common in cereal genome evolution. If different members of a family of related

sequences were amplified (or deleted) in diverging species, then the repeated sequences of this family would be more closely related within species than between species. An alternative hypothesis to account for the amplification of individual members of a family of diverged repeated sequences has been proposed by Smith [215,216]. By computer simulation studies, he has shown that it is possible by expansion and contraction of a set of diverged repeated sequences, by unequal crossing over, to form families of identical repeated sequences from one diverged member of an older family of repeated sequences. If different diverged sequences of the same family are "selected" in this way during species divergence, then this can explain how repeated sequences of a family are more closely related within a species than between the species.

Translocations, coupled with the processes of amplification, unequal crossing over and gene conversion, have enabled a sequence to spread to other chromosomes and eventually to all chromosomes in a sexually reproducing population [201,217]. It is quite possible that these processes have been instrumental in establishing the spectrum of repeated sequences now observed in plant chromosomes.

REFERENCES

1. Britten, R.J. and Kohne, D.E. (1968)
Science, 161 : 529 - 540.

2. Jelinek, W.R. and Schmid, C.W. (1982)
Ann. Rev. Biochem., 51 : 813 - 844.

3. Burr, R., Burr, F.A., John, T.P.S., Thomas, M. and Davis, R.W. (1982)
J. Mol. Biol., 154 : 813

4. Tabata, T., Sasaki, K. and Iwabuchi, M. (1983)
Nuc. Acids Res., 11 : 5865 - 5875.

5. Davidson, E.H., Hough, B.R., Amenson, C.S. and Britten, R.J. (1973b)
J. Mol. Biol., 77 : 1-23

6. Davidson, E.H., Britten, R.J. (1973a)
Quart. Rev. Biol., 48 : 565-613.

7. Manning, J.E., Schimd, C.W., Davidson, N. (1975)
Cell, 4 : 141-155.

8. Schmid, C.W., Deininger, P.L. (1975)
Cell, 184 : 345-358.

9. Crain, W.R., Eden, F.C., Pearson, W.R., Davidson, E.H., Britten, R.J. (1976)
Chromosoma, 56 : 309-326.

10. Deininger, P.L., Schmid, C.W. (1976)
J. Mol. Biol., 106 : 773-790.

11. Finnegan, D.J., Rubin, G.M., Young, M.W., Hogness, D.S. (1978)
Cold Spring Harbor Symp. Quant. Biol., 42 : 1053-1063.

12. Gupta, V.S., Gadre, S.R. and Ranjekar, P.K. (1981)
Biochim. Biophys. Acta, 656 : 147—154.

13. Cullis, C.A. (1981)
Biochim. Biophys. Acta, 652 : 1-15.

14. Lagu, M.D., Ranjekar, P.K. and Pillay, D.T.N. (1986)
Cell Biology International Reports, 10 : 869-874.

15. Pruitt, R.E. and Meyerowitz, E.M. (1986)
J. Mol. Biol., 187 : 169-183.

16. Dhar, M.S., Dabak, M.M., Gupta, V.S. and Ranjekar, P.K. (1988)
Plant Sci., 55 : 43-52.

17. Ranjekar, P.K., Gupta, V.S. and Ranade, S.A. (1983)
IN : Gene structure and function in higher plants. Chapter 19, p. 231-236. Eds. G.M. Reddy and E.H. Coe.

18. Murray, M.G., Palmer, J.D., Cuellar, R.E. and Thompson, W.F. (1979)
Biochemistry, 18 : 5259-5266.

19. Sivaraman, L., Gupta, V.S. and Ranjekar, P.K. (1986)
Plant Mol. Biol., 6 : 375-388.

20. Walbot, V. and Dure, III, L.S. (1976)
J. Mol. Biol., **101** : 503-536.

21. Dabak, M.M., Ranade, S.A., Dhar, M.S., Gupta, V.S. and Ranjekar, P.K. (1988)
Ind. J. Biochem. Biophys., **25** : 230-236.

22. Lagu, M.D. (1987)
IN : 'Molecular analysis of Plant genomes with special reference to three Cucurbitaceae species, Ph.D. Thesis, University of Poona, Poona, India.

23. Wenzel, W. and Hemleben, V. (1982)
Plant Syst. Evol., **140** : 75-86.

24. Gupta, V.S. and Ranjekar, P.K. (1981)
J. Biosci., 3 : 417-430.

25. Bhave, M.R., Gupta, V.S. and Ranjekar, P.K. (1986)
Plant Syst. Evol., **152** : 133-151.

26. Ranade, S.A. (1986)
IN : Molecular studies in Plant genomes with special reference to Cowpea and Pigeonpea, Ph.D. Thesis, University of Poona, Poona. India.

27. Seshadri, M. and Ranjekar, P.K. (1980)
Biochim. Biophys. Acta, 610 : 211-220.

28. Kiper, M. and Herzfeld, F. (1978)
Chromosoma,(Berl.), 65 : 335-351.

29. Seshadri, M. (1980)
IN : Study of Nucleic acids : DNA comparison of Phaseolus plant species, Ph.D. Thesis, University of Poona. India.

30. Wimpee, C.F. and Rawson, J.R.Y. (1979)
Biochim. Biophys. Acta, 565 : 192-206.

31. Gupta, V.S. and Ranjekar, P.K. (1982)
Ind. J. Biochem. Biophys., **19** : 167-170.

32. Goldberg, R.B. (1978)
Biochem. Genet., **16** : 45-68.

33. Gurley, W.B., Hepburn, A.G. and Key, J.L. (1979)
Biochim. Biophys. Acta, **561** : 167-183.

34. Zimmerman, J.L. and Goldberg, R.B. (1977)
Chromosoma,(Berl.), 59 : 227-252.

35. Sivaraman, L., Gupta, V.S. and Ranjekar, P.K. (1984)
J. Biosci., 6 : 795-809.

36. Murray, M.G., Cuellar, R.E. and Thompson, W.F. (1978)
Biochemistry, 17 : 5781-5790.

37. Hake, S and Walbot, V. (1980)
Chromosoma,(Berl.), 79 : 251-270.

38. Smith, D.B. and Flavell, R.B. (1977)
Biochim. Biophys. Acta, **474** : 82-97.

39. Stack, S.M. and Comings, D.E. (1979)
Chromosoma,(Berl.), 70 : 161-181.

40. Flavell, R.B. and Smith, D.B. (1976)
Heredity, 37 : 231-252.

41. Bennet, M.D. (1973)
Brookhaven Symp. Biol., 25 : 344-360.

42. Bennet, M.D. (1982)
In : Systematics Association Special Vol. 20. Genome evolution. G.A. Dover and R.B. Flavell, Eds., Academic Press, New York, p. 239-261.

43. Britten, R.J. and Davidson, E.H. (1969)
Science, 165 : 349-357.

44. Davidson, E.H. and Britten, R.J. (1979)
Science, 204 : 1052-1059.

45. Walbot, V. and Goldberg, R.B. (1979)
IN : Nucleic Acids in Plants. Vol. I. T.C. Hall and J.W. Davies (Eds.,), CRC Press, Florida, pp. 3-40.

46. Flavell, R.B. (1980)
Ann. Rev. Plant Physiol., 31 : 569-596.

47. Thompson, W.F., Murray, M.G. (1981)
IN : The Biochemistry of Plants. A comprehensive treatise, Vol. 6. Proteins and Nucleic Acids. A. Marcus (Ed.), Academic Press, New York, pp. 1-82.

48. Davidson, E.H., Jacobs, H.T. and Britten, R.J. (1983)
Nature, 301 : 468-470.

49. Bendich, A.J. and Anderson, R.S. (1977)
Biochemistry, 16 : 4655-4663.

50. Preisler, R.S. and Thompson, W.F. (1981)
J. Mol. Evol., 17 : 78-84.

51. Murray, M.G., Peters, D.L. and Thompson, W.F. (1981)
J. Mol. Evol., 17 : 31-42.

52. Brutlag, D.L. (1980)
Ann. Rev. Genet., **14** : 121-144.

53. Singer, M.F. (1982)
Int. Rev. Cytol., **76** : 67-112.

54. Ranjekar, P.K., Pallota, D. and Lafontaine, J.G. (1978b)
Can. J. Biochem., **56** : 808-815.

55. Bedbrook, J.R., O'Dell, M. and Flavell, R.B. (1980a)
Nature, **288** : 133-137.

56. Bedbrook, J.R., Jones, J., O'Dell, M., Thompson, R.D. and Flavell, R.B. (1980b)
Cell, **19** : 545-560.

57. Fodor, I. and Beridze, T. (1980)
Mol. Biol. Rep., **6** : 17-20.

58. Grisvard, J. and Tuffet-Anghileri, A. (1980)
Nuc. Acids Res., **8** : 2843-2858.

59. Peacock, W.J., Dennis, E.S., Rhoades, M.M. and Pryor, A.J. (1981)
Proc. Natl. Acad. Sci., **78** : 4490-4494.

60. Shmookler-Reis, R., Timmis, J.N. and Ingle, J. (1981)
Biochem. J., **195** : 723-734.

61. Deumling, B. (1981)
Proc. Natl. Acad. Sci., **78** : 338-342.

62. Hemleben, V., Leweke, B., Roth, A and Stadler, J. (1982)
Nuc. Acids Res., **10** : 631-644.

63. Brennick, A. and Hemleben, V. (1983)
Z. Naturforsch, 38C : 1062-1065.

64. Capesius, I. (1983)
Biochim. Biophys. Acta, 739 : 276-280.

65. Kato, A, Yakura, K. and Tanifuji, S. (1984)
Nuc. Acids Res., 13 : 6415-6426.

66. Martinez-Zapater, J.M., Estelle, M.A. and Sommercille, C.R. (1986)
Mol. Gen. Genet., 204 : 417-423.

67. Leclerc, R.F. and Siegel, A. (1987)
Plant Mol. Biol., 8 : 499-507.

68. Grellet, F., Delcasso, D., Panabieres, F. and Delseny, M. (1986)
J. Mol. Biol., 187 : 495-507.

69. Yakura, K, Kato, A. and Tanifuji, S. (1984)
Mol. Gen. Genet., 193 : 400-405.

70. Hallden, C., Bryngelsson, T., Sall, T. and Gustafsson, M. (1987)
J. Mol. Evol., 25 : 318-323.

71. Gerlach, W.L. and Bedbrook, J.R. (1979)
Nuc. Acids Res., 7 : 1869-1885.

72. Appels, R. and Dvorak, J. (1982a)
Theor. Appl. Genet., 63 : 337-348.

73. Flavell, R.B., O'Dell, M. and Thompson, W.F. (1983)
IN : Kew Chromosome, Conference II. P.E. Brandham and M.D. Bennet. (Eds.), George Allen and Unwin, London, pp. 11-17.

74. Flavell, R.B., O'Dell, M., Smith, D.B. and Thompson, W.F. (1985)
IN : Molecular form and function of the plant genome, L. Van Vloten-Doting, G.S.S.P. Groot and T.C. Hall (Eds.), NATO ASI, 83, Plenum Press, New York, pp. 1-14.

75. Olmedilla, A., Delcasso, D. and Delseny, M. (1984)
Plant Sci. Lett., 37 : 123-127.

76. Takaiwa, F., Oono, K., Iida, Y., and Sugiura, M. (1985)
Gene, 37 : 255-259.

77. McMullen, N.D., Hunter, B., Phillips, and Rubenstein, I. (1986)
Nuc. Acids Res., 14 : 4953-4968.

78. Toloczyki, C. and Feix, G. (1986)
Nuc. Acids Res., 14 : 4969-4986.

79. Ellis, T.H.N., GoldsbroughP.B. and Castleton, J.A. (1983)
Nuc. Acids Res., 11 : 3047-3064.

80. Waldron, J., Dunsmuir, P. and Bedbrook, J.R. (1983)
Plant Mol. Biol., 2 : 57-65.

81. Seigel, A. and Kolacz, K. (1983)
Plant Physiol., 72 : 166-171.

82. Delseny, M., Laroche, M. and Penon, P. (1984)
Plant Physiol., 76 : 627-632.

83. Systma, K.J. and Schaal, B.A. (1985)
Evolution, 39 : 534-608.

84. Rogers, S.O., Honda, S. and Bendich, A.J. (1986)
Plant Mol. Biol., 6 : 339-345.

85. Ellis, T.H.N., Davies, D.R., Castleton, J.A. and Bredford, I.D. (1985)
Chromosoma (Berl.), 91 : 74-81.

86. Doyle, J.J. and Beachy, R.N. (1985)
Theor. Appl. Genet., 70 : 369-372.

87. Olmedilla, A., Delcasso, D., Delseny, M. and Cauwet0Marc, A.M. (1985)
Plant Syst. Evol., 150 : 263-274.

88. Evans, I.J., James, A.M. and Barnes, S.R. (1983)
J. Mol. Biol., 170 : 803-826.

89. Sentry, J.W. and Smith, D.R. (1985)
Chromosoma (Berl.), 92 : 149-155.

90. Sakowicz, T., Galazka, G., Konarzeuska, A., Kurinkowski, M. and Klysik, J. (1986)
Planta, 168 : 207-213.

91. Kato, A., Iida, Y., Yakura, K. and Tanifuji, S. (1985)
Plant Mol. Biol., 5 : 41-53.

92. Gupta, M., Shepard, N.S., Bertram, I. and Saedler, H. (1984)
Embo J., 3 : 133-139.

93. Ranade, S.A., Lagu, M.D., Patankar, S.M., Dabak, M.M., Dhar, M.S., Gupta, V.S. and Ranjekar, P.K. (1988)
Biosci. Reports, 8 : 435-441.

94. Roy, P., Bhattacharya, N. and Biswas, B.B. (1988)
Gene, 73 : 57-66.

95. Orgel, L.E. and Crick, F.H.C. (1980)
Nature, 284 : 604-607.

96. Britten, R.J. and Davidson, E.H. (1971)
Quart. Rev. Biol., 46 : 111-138.

97. Herget, T., Reich, M., Stuber, K. and Starzinski-Powitz, A. (1986)
Embo J., 5 : 659-664.

98. Daniels, G.R. and Deininger, P.L. (1985)
Nature, 317 : 819-822.

99. Duncan, C.M., Jagadeeswaran, P., Wang, R.R.C. and Weissman, S.M. (1981)
Gene, 13 : 185-196.

100. Freeling, M. (1984)
Ann. Rev. Plant Physiol., 35 : 277-298.

101. McClintock (1948)
Carnegie Inst. Wash. Yrbk., 47 : 155-169.

102. Behrens, U., Fedoroff, N., Laird, A., Muller-Neumann, M., Starlinger, P. and Yoder, J. (1984)
Mol. Gen. Genet., 194 : 346-347.

103. Muller-Neumann, M., Yoder, I.I. nad Starlinger, P. (1984)
Mol. Gen. Genet., 198 : 19-24.

104. Fedoroff, N., Furtek, D.B. and Nelson, O.E. (1984)
Proc. Natl. Acad. Sci. (USA), 81 : 3825-3829.

105. Fedoroff, N. (1983)
IN : Mobile Genetic Elements, pp. 1-63. J. Shapiro (Ed.,),
Academic Press, New York.

106. Pohlman, R.F., Fedoroff, N. and Messing, J. (1984a)
Cell, 39 : 417

107. Pohlman, R.F., Fedoroff, N. and Messing J. (1984b)
Cell, 37 : 635-643.

108. Sutton, W.D., Gerlach, W.L., Schwartz, D. and Peacock, W.J. (1984)
Science, 223 : 1265-1268.

109. Doring, H.P., Tillman, E. and Starlinger, P. (1984)
Nature, 307 : 127-130.

110. Weck, E., Courage-Tekke, U., Doring, H.F., Fedoroff, N. and Starlinger, P. (1984)
Embo J., 3 : 1713-1716.

111. Courage-Tebbe, U., Doring, H.F., Frommer, W.B., Kunze, R., Laird, A., Merckelbach, A.A., Muller-Neumann, M., Riegel, J., Starlinger, P., Tillman, E., Weck, E., Were, W. and Yoder, J. (1984)
Cold Spring Harbor Symp. Quant. Biol., 49 : 329-338.

112. Barker, R.F., Thompson, D.V., Talbot, D.R., Swanson, J. and Bennetzen, J.L. (1984)
Nuc. Acids Res., 12 : 5955-5967.

113. O'Reilly, C., Shepherd. N.S., Pereira, A., Schwarz-Sommer, Z., Bertram, I., Robertson, D.S., Peterson, P.A. and Saedler, H. (1985)
Embo J., 4 : 877-882.

114. Pereira, A., Schwarz-Sommer, Z., Gierl, A., Bertram, I., Peterson, P.A. and Saedler, H. (1985)
Embo J., 4 : 17-23.

115. Pereira, A., Cuypers, H., Gierl, A., Schwarz-Sommer, Z. and Saedler, H. (1986)
Embo J., 5 : 835-842.

116. Schwarz-Sommer, Z., Gierl, A., Klosgeu, R.B., Wienand, U., Peterson, P.A. and Saedler, H. (1984)
Embo J., 3 : 1021-1028.

117. Shepherd, N.S., Schwarz-Sommer, Z., Wienand, U., Sommer, H. and Deumling, B. (1982)
Mol. Gen. Genet., 188 : 266-271.

118. Sommer, H., Carpenter, R., Harison, B.J. and Saedler, H. (1985)
Mol. Gen. Genet., 199 : 225-231.

119. Vodkin, L.O., Rhodes, P.R. and Goldberg, R.B. (1983)
Cell, 34 : 1023-1031.

120. Saedler, H. and Nevers, P. (1985)
Embo J., 4 : 585-590.

121. Hinegardner, R. (1976)
IN : Molecular Evolution, F.J. Ayala (Ed.), pp. 179-199.
Sinaver Associates Inc., Mass.

122. Rees, H. and Hazarika, M.H. (1969)
Chromosomes Today, 2 : 157-165.

123. Jones, R.N. and Brown, L.M. (1976)
Heredity, 36 : 91-104.

124. Hutchinson, J., Narayan, R.K.J. and Rees, H. (1980)
Chromosoma (Berl.), 78 : 137-145.

125. Flavell, R.B., Bennet, M.D., Smith, J.B., Smith, D.B. (1974)
Biochem. Genet., 12 : 257-269.

126. Lakshmi, S. and Ranjekar, P.K. (1984)
Ind. J. Biochem. Biophys., 21 : 299-303.

127. Bachmann, K. and Price, H.J. (1977)
Chromosoma (Berl.), 61 : 257-267.

128. Cullis, C.A. and Schweizer, D. (1974)
Chromosoma (Berl.), 44 : 417-421.

129. Ranjekar, P.K., Pallota, D. and Lafontaine, J.G. (1978a)
Biochem. Genet., 16 : 957-970.

130. Patankar, S., Joshi, C.P., Ranade, S.A., Bhave, M. and Ranjekar, P.K. (1985)
Proc. Ind. Acad. Sci. (Plant Sciences), 94 : 539-551.

131. White, M.J.D. (1977)
IN : Modes of Speciation. Freeman and Co., San Francisco.

132. Ayonoadu, V. and Rees, H. (1973)
Heredity, 30 : 233-240.

133. Metzlaff, M., Troebner, W., Baldauf, F., Schlegel, R. and Cullum, J. (1986)
Theor. Appl. Genet., 72 : 207-210.

134. Houck, C.M., Rinehart, F.P. and Schmid, C.W. (1979)
J. Mol. Biol., 132 : 289-306.

135. Singer, M.F. (1982a)
Cell, 28 : 433-434.

136. Voliva, C.F., Jahn, C.L., Comer, M.B., Hutchinson, C.A. III and Edgel, M.H. (1983)
Nuc. Acids Res., 11 : 8847-8859.

137. Singer, M.F. and Skowronski, J. (1985)
Trends Biochem. Sci., 10 : 119-122.

138. Walter, P. and Blokel, G. (1980)
Proc. Natl. Acad. Sci. USA , 77 : 7112-7116.

139. Martin, S.L., Voliva, C.F., Burton, F.H., Edgell, M.H. and Hutchinson, C.A. III (1984)
Proc. Natl. Acad. Sci. USA, 81 : 2308-2312.

140. Rogers, J.H. (1985)
Biochim. Biophys. Acta, 824 : 113-120.

141. Lerman, M.I., Thayer, R.E. and Singer, M.F. (1983)
Proc. Natl. Acad. Sci. USA , 80 : 3966-3970.

142. Shafit-Zagardo, B., Brown, F., Zarodny, P. and Maio, J. (1983)
Nature, 304 : 277-280.

143. Schindler, C.W. and Rush, M.G. (1985)
J. Mol. Biol., 181 : 161-173.

144. Rubin, G.M., Brorein, W.J. Jr., Dunsmuir, P., Flavell, A.J., Levis, K., Strobel, E., Toole, J.J. and Young, E. (1981)
Cold Spring Harbor Symp. Quant. Biol., 45 : 619-628.

145. Scherer, G., Tschudi, C., Pereira, J., Delius, H. and Pirrotta, V. (1982)
J. Mol. Biol., 157 : 435-451.

146. Potter, S.S., Truett, M., Phillips, M. and Maher, A. (1980)
Cell, 20 : 639-647.

147. Modolell, J., Bender, W., and Meselson, M. (1983)
Proc. Natl. Acad. Sci. USA , 80 : 1678-1672.

148. Rubin, G.M., Kidwell, M.G. and Bingham, P.M. (1982)
Cell, 30 : 987-994.

149. McGinnis, W., Shemoen, A.W. and Beckendorf, S.K. (1983)
Cell, 34 : 75-84.

150. Bucheton, A., Paro, R., Sang, H.M., Pelisson, A. and Finnegan D.J. (1984)
Cell, 38 : 153-163.

151. Deumling, B. and Greilhuber, J. (1982)
Chromosoma (Berl.), 84 : 535-555.

152. Dennis, E.S., Gerlach, W.L. and Peacock, W.J. (1980b)
Heredity, 44 : 349-366.

153. Hutchinson, J. and Lonsdale, D.M. (1982)
Heredity, 48 : 371-376.

154. Bostock, C. (1980)
Trends in Biochem. Sciences, 5 : 117-119.

155. Appels, R. and Peacock, W.J. (1978)
Int. Rev. Cytology, Suppl. 8, G.H. Bourne and J.F. Danielli (Eds.,), pp. 70-126.

156. Doolittle, W.F. and Sapienza, C. (1980)
Nature, 284 : 601-603.

157. Orgel, L.E. and Crick, F.H.C. (1980)
Nature, 284 : 604-607.

158. Grisvard, J. (1985)
Plant Sci., 39 : 189-193.

159. Spofford, J.P. (1976)
IN : The Genetics and Biology of Drosophila, Ashburner, M. and Novitski, E., (Eds.), Vol. IC, Academic Press, London, 955-1019.

160. Wakefield, L., Ackerman, E.J. and Gurdon, J.B. (1983)
Dev. Biol., 95 : 468-475.

161. Andrews, D.L., Millstein, L., Hamkalo, B.A. and Gottesfeld, J.M. (1984)
Nuc. Acids Res., **12** : 7753-7769.

162. Benslimane, A.A, Dron, M., Hartmann, C. and Rode, A. (1986)
Nuc. Acids Res., **14** : 8111-8119.

163. Wolfe, J., Darling, S.M., Erickson, R.P., Craig, I.W., Buckle V.J., Rigby, P.W.J. (1985)
J. Mol. Biol., **182** : 477-485.

164. Witney, F.R. and Furano, A.V. (1983)
Nuc. Acids Res., **11** : 291-304.

165. Hsieh, T.S. and Brutlag, D.L. (1979)
Proc. Natl. Acad. Sci., **76** : 726-730.

166. Strauss, F. and Varshavsky, A. (1984)
Cell, **37** : 899-901.

167. Manuelidis, L. (1982)
IN : Genome Evolution, p. 263 - 285.

168. Musich, P.R., Brown, F.L. and Maio, J.J. (1982)
Proc. Natl. Acad. Sci. USA , **79** : 118-122.

169. Zhang, X.Y., Fittler, F. and Horz, W. (1983)
Nuc. Acids Res., **11** : 4287-4306.

170. Brock, H., Abler, S., Zhang, X.Y., Fritton, H. and Igo Kemenes, T. (1984)
J. Mol. Biol., **176** : 131-154.

171. Linxweiler, W. and Horz, W. (1985)
Cell, 42 : 281-290.

172. Endoch, H. and Okada, N. (1986)
Proc. Natl. Acad. Sci., 83 : 251-255.

173. Wu, R. and Wu, T. (1987)
Nuc. Acids Res., 15 : 5913-5923.

174. Kamalay, J.C. and Goldberg, R.B. (1980)
Cell, 19 : 935-946.

175. Galau, G.A. and Dure, L. (1981)
Biochemistry, 20 : 4169-4178.

176. Goldberg, R.B., Hoscheck, G., Ditta, G.S. and Breidenbach, R.W. (1981)
Dev. Biol., 83 : 218-231.

177. Peacock, W.J., Dennis, E.S. and Gerlach, W.L. (1982)
IN : Progress in Clinical and Biological Research, Vol. 96, 123-142, C. Barigozzi, G. Montalenti and M.J.D. White, (Eds.), Alan, R. Liss Publishers, New York.

178. Vaillant, V., Buffard, D. and Esnault, R. (1983)
Cell Differentiation, 13 : 201-208.

179. Pearson, G.G., Timmis, J.N. and Ingle, J. (1974)
Chromosoma (Berl.), 45 : 281-294.

180. Nagl, W. (1976)
Plant Sci. Lett., 7 : 1-6.

181. Nagl, W., Peschke, C., VanGyseghem, R. (1976)
Naturwissenschaften, 63 : 198-199.

182. Van Oostveldt, P. and Van Pariys, R. (1976)
Expt. Cell Res., 98 : 210-221.

183. Olszewska, M.J. (1976)
Histochemistry, 49 : 157-175.

184. Schaffner, K.H. and Nagl, W. (1979)
Plant Syst. Evol., Suppl. 2., : 105-110.

185. Cullis, C.A. and Cleary, W. (1986)
Can J. Genet. Cytol., 28 : 247-251.

186. Zheng, K.L., Castiglione, S., Biasini, M.G., Biroli, A., Morandi, C. and Sala, F. (1985)
Theor. Appl. Genet., 74 : 65-70.

187. Kikuchi, S., Takaiwa, F. and Oono, K. (1987)
Mol. Gen. Genet., 210 : 373-380.

188. Walbot, V., Cullis, C.A. (1985)
Ann. Rev. Plant Physiol., 36 : 367-396.

189. Stebbins, G.L. (1950)
IN : Variation and Evolution in Plants, Columbia University Press, New York.

190. Stebbins, G.L. (1982)
IN : Progress in Clinical and Biological Research, Vol. 96, 21-39, C. Barigozzi, G. Montalenti and M.J.D. White, (Eds.), Alan R. Liss Publishers, New York.

191. Appels, R. and Dvorak, J. (1982b)
Theor. Appl. Genet., 63 : 361-365.

192. Olmedilla, A., Delcasso, D., Delseny, M. and Cauwet-Marc, A.M. (1985)
Plant Syst. Evol., 150 : 263-274.

193. Lassner, M. and Dvorak, J. (1986)
Nuc. Acids Res., 14 : 5499-5512.

194. Schuler, M.A., Schmitt, E.S. and Beachy, R.N. (1982a)
Nuc. Acids Res., 10 : 8225-8244.

195. Schuler, M.A., Ladin, B.F., Pollaco, J.C., Freyer, J.G. and Beachy, R.N. (1982b)
Nuc. Acids Res., 10 : 8245-8261.

196. Boulter, D., Evans, J.M., Ellis, J.R., Shirsat, A., Gatehouse J.A. and Croy, R.R.D. (1987)
Plant Physiol. Biochem., 25 : 283-289.

197. Pichersky, E., Bernatzky, R., Tanksley, S.D. and Cashmore, A.R. (1986)
Proc. Natl. Acad Sci. USA, 83 : 3880-3884.

198. Chauket, N., Phillips, G., Chaboute, M.E., Ehling, M. and Gigot, C. (1986)
Plant Mol Biol., 6 : 253-263.

199. Gigot, C., Chauket, N., Chaboutem M.E., Ehling, M. and Phillips, G. (1987)
Plant Physiol. Biochem., 25 .

200. Dover, G.A. (1982)
Nature, 299 : 111-117.

201. Dover, G.A., Brown, S.D.M., Coen, E.S., Stracken, T. and Trick, M. (1982)
IN : Genome Evolution Dover, G.A., and Flavell, R.B. (Eds.), Academic Press, London , pp. 343-372.

202. Flavell, R.B. (1982)
IN : Genome Evolution. (G.A. Dover, and R.B. Flavell, (Eds.), Academic Press, London, pp. 301-323.

203. Flavell, R.B., Rimpau, J. and Smith, D.B. (1977)
Chromosoma (Berl.), 63 : 205-222.

204. Pech, M., Strreck, R.F. and Zachau, H.G. (1979)
Cell, 18 : 883-893.

205. Dennis, E.S., Dunsmuir, P. and Peacock, W.J. (1980a)
Chromosoma, 79 : 179-198.

206. Anderson, D.M., Scheller, R.H., Pasakony, J.W., McAllister, L.B., Trabert, S.G., Beall, C., Britten, R.J., and Davidson, E.H. (1981)
J. Mol. Biol., 145 : 5-28.

207. Bedbrook, J.R., Jones, J. and Flavell, R.B. (1981)
Cold Spring Harbor Symp. Quant. Biol., 45 : 755-760.

208. Maio, J.J., Brown, F.L. and Musich, P.R. (1977)
J. Mol. Biol., 117 : 637-655.

209. Brison, O. (1983)
Biochimie, 65 : 3-11.

210. Deininger, P.L. and Daniels, G.R. (1986)
Trends in Genetics, 3 : 76-80.

211. Ackerman, E.J. (1983)
Embo, J., 2 : 1417-1422.

212. Saghai-Maroof, M.A., Soliman, K.M., Jorgensen, R.A. and Allard, R.W. (1984)
Proc. Natl. Acad. Sci. (USA), 81 : 8014-8018.

213. Rogers, S.O., Honda, S. and Bendich, A.J. (1986)
Plant Mol. Biol., 6 : 339-345.

214. Kourilsky, P. (1986)
Trends in Genetics, 2 : 60-62.

215. Smith, G.P. (1973)
Cold Spring Harbor Symp. Quantitative Biol., 38 : 507-514.

216. Smith, G.P. (1976)
Science, 191 : 528-535.

217. Dover, G.A (1981)
IN : Mechanisms of Speciation, C. Barigozzi, G. Montalenti, and M.J.D. White (Eds.), Alan, R., Liss Inc., New York.
pp. 435-459.

CHAPTER 2

ORGANIZATION OF RICE GENOME AT A DNA FRAGMENT LENGTH OF 20 Kbp.

ABSTRACT

In order to find out if the unique genome organization pattern of rice observed at a DNA fragment length of 6.5 kbp [6] persisted even at a higher DNA fragment length, reassociation of high molecular weight rice DNA (20 kbp) was studied at a limit Cot value of 50. The percentage of reassociation of rice DNA was 56% at Cot 50 as compared to 52% at a fragment length size of 6.5 kbp indicating that there was no significant increase in reassociation with an increase in DNA fragment length. The decrease in percent reassociation after S1 nuclease was from 56% to 48% in case of 20 kbp rice DNA as compared to 52 to 48% in case of 6.5 kbp DNA. Based on these data, it was concluded that repeated and single copy DNA sequences were uninterspersed in rice genome even at a high DNA fragment length of 20 kbp. DNA sequence organization was also examined in two other rice cultivars namely Ratnagiri-24 (R-24) and Sona at a fragment size of 20 kbp, and these again exhibited the lack of interspersion of repeated and single copy DNA sequences as in case of Basmati 370.

2.1 INTRODUCTION

During the study of genome structure and genome organization in higher plants in our laboratory and elsewhere [1-10], it has been shown that plants exhibit a great diversity in genome organization patterns and that this diversity arises mainly due to variations in the length of interspersed repeat DNA sequences. The work on rice genome was initiated in our laboratory as early as 1978, mainly because very little data on rice DNA was available at that time. Deshpande and Ranjekar [11] and Gupta et al [6] showed it for the first time that the repetitive DNA content in rice was 50% and that these DNA sequences remained uninterspersed with single copy DNA sequences at a DNA fragment length of 6.5 kbp. Absence of interspersion in rice genome at a DNA fragment length of 6.5 kbp was considered to be a very unique observation at that time in view of the widespread occurrence of interspersion of repeat and single copy DNA sequences in plant genomes.

The aim of the present study was to check whether the unique pattern of absence of DNA sequence interspersion in rice persisted even at a higher DNA fragment length of 20 kbp. In addition, I wished to find out if the genomes of two more varieties of rice also showed the absence of interspersion between repeated and single copy DNA sequences.

2.2 MATERIALS AND METHODS

2.2.1 Seed material and germination

Seeds of rice (varieties : Basmati 370, Ratnagiri-24 (R-24) and Sona) were thoroughly washed first with water and then with 50% alcohol and finally soaked overnight in distilled water. They were grown on wet cotton in dark for about 8-10 days. Shoots (10-15 cm height) were cut one inch above the surface and were stored at -70°C till further use.

2.2.2 Chemicals, materials and enzymes

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

Chloroform and ethanol (Swastik Laboratory Pvt. Ltd., Pune) and phenol (SD's Lab-Chem Industry, Bombay) were freshly distilled prior to use. Fine chemicals such as Tris (Trishydroxymethyl amino methane), SDS (Sodium dodecyl sulphate), Pipes (Piperazine-N-N'-bis 2-ethane sulfonic acid) buffer and agarose were from Sigma Chemical Co., USA. Enzymes like RNase A and Protease were also from Sigma Chemical Co., USA. DNA molecular weight markers, λ DNA digested with HindIII, ØX174 RF DNA digested with HaeIII and S1 nuclease were from Bethesda Research Laboratories, USA.

2.2.3 DNA isolation

DNAs were extracted from the frozen tissue by a combination of procedures of Marmur [12] and Ranjekar et al [13]. The frozen plant tissue was homogenised in Tris maleate buffer (0.5 M sucrose, 0.05 M tris, 0.05 M maleic acid, 0.003 M $CaCl_2$, pH 6.0 containing 0.1% Triton X-100) in a Sumeet mixer at a maximum speed for 90 seconds. The tissue homogenate was filtered through two layers of muslin cloth and the filtrate was spun at 9000 rpm in Sorvall Centrifuge (rotor : GSA) for 20 minutes to collect the nuclear pellet. The latter was first washed with saline-EDTA (0.15 M NaCl, 0.1 M sodium salt of EDTA, pH 8.0) and then suspended in a minimum volume of the same buffer. The nuclei were lysed by the addition of 20% SDS to a final concentration of 2% followed by incubation at $62^{o}C$ for 40 min. The suspension was cooled and the proteins were denatured by adding 5 M sodium perchlorate to an effective concentration of 1 M. This was followed by deproteinization with equal volume of chloroform : isoamyl alcohol (24 : 1 v/v) mixture. This step was repeated till the protein interphase was totally absent. The aqueous layer was treated with 50 ug/ml RNase A (made DNase free by heating at $80^{o}C$ for 10 min) at $37^{o}C$ for 1 h and was deproteinised with chloroform : isoamyl alcohol mixture. The DNA from the aqueous layer was precipitated with 2 volumes of chilled ethanol and the DNA fibres were spooled out, dried and dissolved in 10 mM Tris-HCl, pH 7.4. About 1 mg DNA was obtained from 500 g shoot tissue.

All the purified DNA preparations exhibited optical density ratios as A_{280}/A_{260} = 0.55 and A_{230}/A_{260} = 0.45 and had an absorbance of less than 0.1 at 300 nm. Each DNA preparation was also checked by agarose gel electrophoresis using λ HindIII digest as a high molecular weight marker DNA to determine size heterogeneity. The DNAs obtained were generally of an average size 20 kbp.

2.2.4 Reassociation kinetics of 20 kbp long DNA fragments by hydroxyapatite (HAP) column chromatography

DNA reassociation in the Cot range of 5×10^{-2} M.s to 5×10^{3} M.s was measured only for the rice variety Basmati 370. These experiments were carried out by HAP column chromatography [14] where HAP was prepared according to the procedure of Tiselius et al and stored at 4°C with a few drops of chloroform. Prior to use, HAP was checked for DNA recovery which was normally more than 95%.

In HAP column preparation, the HAP slurry was poured into a double jacketed column (15 x 2.5 cm) to a bed volume of about 16 ml and the column was equilibrated at 62°C with 0.12 M PB, pH 6.8. DNA samples (approximately 500 ug) in 0.12 M PB, pH 6.8 were denatured by keeping in a boiling water bath for 10 min. They were then incubated to a specific Cot value at 62°C and loaded on preequilibrated HAP column at the same temperature. The single stranded DNA was eluted with 0.12 M PB, pH 6.8 and double stranded

DNA was eluted with 0.4 M PB, pH 6.8 at 62°C.

The relative amount of single and double stranded DNA was determined by measuring the absorbance at 260 nm in a Shimadzu double beam spectrophotometer (Model UV 210 A) assuming one absorbance unit (A_{260}) is = 50 ug DNA/ml. The extent of DNA reassociation was calculated using the formula :

$$\% R = \frac{\text{ug of DNA eluted with 0.4 M PB, pH 6.8}}{\text{ug of DNA eluted with 0.12 M PB, pH 6.8} + \text{ug of DNA eluted with 0.4 M PB, pH 6.8}} \times 100 \quad [16]$$

The experimental data points were plotted on a semilog scale with Cot values on X-axis and % reassociation on Y-axis. E. coli DNA was used as an internal standard in these experiments.

2.2.5 Computer analysis of reassociation kinetics data

Since the individual DNA components are usually not well separated in a Cot curve, various computer fitting procedures are used to model the reassociation curves into two, three or more ideal second order components. In the present work, the least squares regression analysis of DNA reassociation kinetics data was carried out by using a standard optimization subroutine STEPIT which was written by J.P. Chandler and distributed by QCPE, Indiana University, USA.

Second order reassociation kinetics can be described by the following equation :

$$\frac{C}{C_o} = \frac{1}{1 + K\ Cot}$$

where C = concentration of free single strands after t sec of reassociation.

C_o = initial concentration of DNA in moles of nucleotides per litre.

K = rate constant of reaction.

The error is given by the equation :

$$\text{Error} = \sum \left\{ \left(\frac{C}{C_o}\right)_{expt} - \left(\frac{C}{C_o}\right)_{fit} \right\}$$

and is minimised by a direct search procedure. The form for $\frac{C}{C_o}$ is assumed to be :

$$\frac{C}{C_o} = T + \sum \frac{F_i}{L + K_i\ Cot}$$

where $\frac{C}{C_o}$ = fraction denatured

T = fraction of the genome that fails to reassociate

F_i = fraction of the genome reassociating in the i^{th} component

K_i = rate of reassociation of i^{th} component

To minimise the error, the parameters T, F_i, and K_i are all allowed to free float.

2.2.6 Isolation of S1-nuclease resistant DNA duplexes

The S1-nuclease resistant DNA duplexes reassociating in the Cot range of 5×10^{-2} M.s to 5×10^{3} M.s were isolated according to the procedure of Murray et al [4] with some modifications. Atleast 0.4 mg DNA was used for each experiment. Native, shoot DNA of 20 kbp fragment length was dissolved in 10 mM Tris-HCl, pH 7.4, to which 3 M NaCl and 0.06 M Pipes buffer, pH 6.8 were added to an effective concentration of 0.18 M and 0.006 M respectively. The samples were denatured in a boiling water bath for 10 min and allowed to reassociate at 62°C over a range of Cot 5×10^{-2} M.s to Cot 5×10^{3} M.s. After reassociation, the DNA samples were adjusted to 25 mM sodium acetate (pH 4.5), 0.1 mM $ZnSO_4$, 25 mM 2-mercaptoethanol and were treated with S1 nuclease (0.1 U/ug of DNA) at 37°C for 20 min.[17] At the end of reaction, the proteins were removed by chloroform : isoamyl alcohol treatment. The DNAs were finally precipitated by 2 volumes of chilled ethanol, centrifuged, dried and dissolved in 10 mM Tris-HCl, pH 7.4. The percentage of S1 nuclease resistant DNA at a specific Cot value was again determined by measuring spectrophotometrically the absorbance at 260 nm and then using the relationship that 10.D at 260 nm = 50 ug/ml DNA.

2.3 RESULTS

2.3.1 Reassociation kinetics of 20 kbp long rice DNA (cv. Basmati 370)

Reassociation kinetics of 20 kbp long rice DNA were studied in the Cot range of 5.0×10^{-2} M.s to 5.0×10^{3} M.s (Fig. 2.1 (a)) and the data was analysed using the nonlinear least squares regression analysis. From Fig. 2.1a, it is clear that the reassociation curve of 20 kbp rice DNA cannot be resolved into any kinetically distinct components. As reported by Gupta et al [6], two components are observed in 6.5 kbp DNA curve and three in case of 0.5 kbp DNA curve. The merging of repetitive DNA components with increase in DNA fragment length is considered to be an indication of interspersion of different DNA components with each other. However, when the extent of DNA reassociation at limit Cot of 50 is compared at 20 kbp, 6.5 kbp and 0.5 kbp DNA fragment lengths, only a marginal increase from 52% to 56% is observed indicating that rice DNA apparently shows no sequence interspersion between repeated and single copy DNA sequences even at a high DNA fragment length of 20 kbp. The merging of different DNA components may, however, be due to interspersion of repeat DNA sequences amongst themselves.

FIGURE 2.1 : REASSOCIATION KINETICS OF 20 kbp LONG RICE DNA

(a) Reassociation kinetics curve of 20 kbp long rice DNA without S1 nuclease treatment. The DNA concentration in each experiment was about 0.4 mg/ml in 0.12 phosphate buffer, pH 6.8. The solid line indicates the least squares fit, allowing all the parameters to free float. The lower dashed curve represents the predicted reassociation kinetics of a pure component.

o——o Reassociation kinetics curve of standard E. coli DNA.

(b) Reassociation kinetics of 20 kbp long rice DNA after S1 nuclease treatment. The DNA concentration in each experiment was about 0.4 mg/ml in 0.12 M phosphate buffer, pH 6.8.

The solid line indicates the least squares fit, allowing all the parameters to free float. The lower dashed curves represent the predicted reassociation kinetics of pure components.

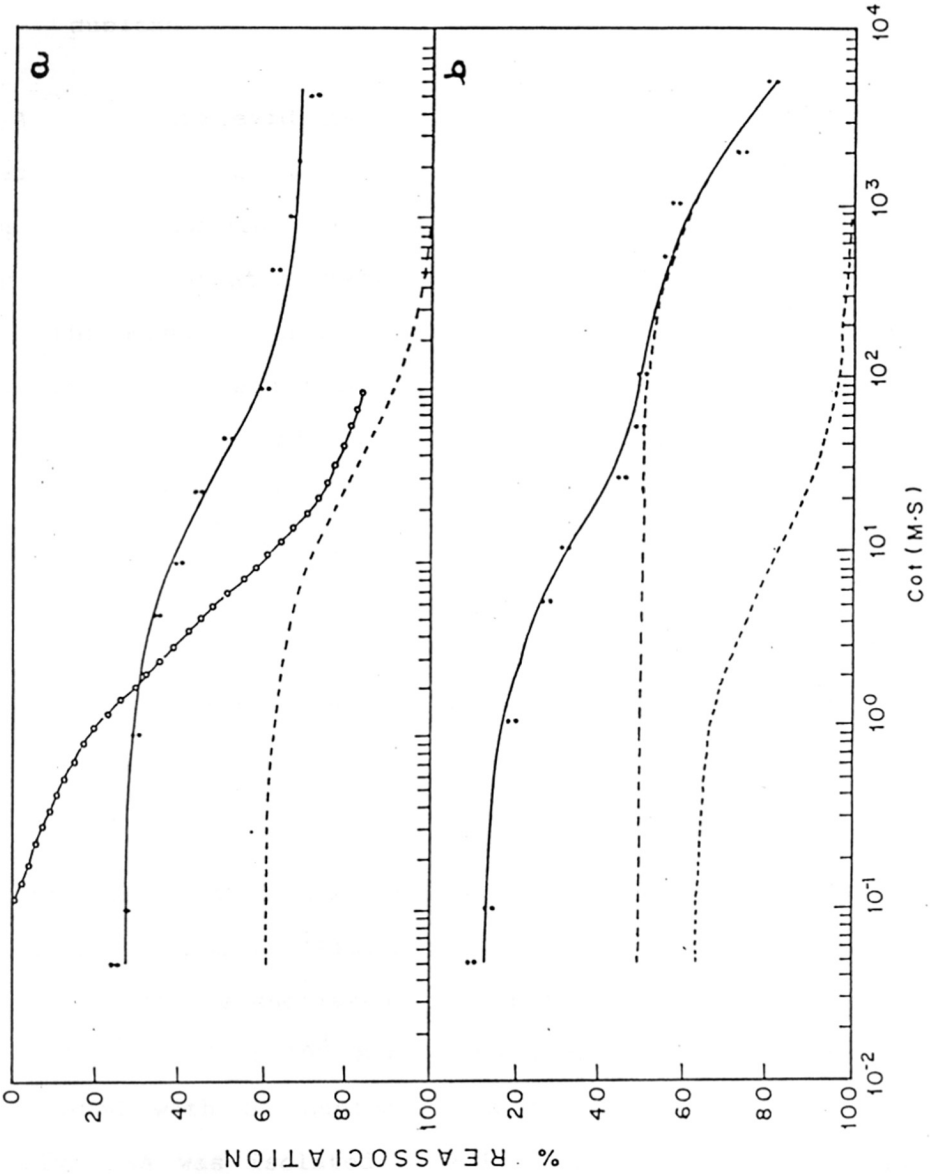

a
b
% REASSOCIATION
0
20
40
60
80
100
10^{-2}
10^{-1}
10^{0}
10^{1}
10^{2}
10^{3}
10^{4}
Cot (M·S)

FIG. 2·1.

2.3.2 S1 nuclease resistance of reassociated repetitive DNA duplexes

An alternate evidence for the occurrence of interspersion of repetitive and single copy DNA sequences in a genome is to assess the change in % DNA reassociation before and after S1 nuclease treatment at a specific Cot value. During HAP fractionation, the single stranded tails which presumably represent single copy DNA sequences attached to reassociated duplexes are also eluted with high salt buffer leading to an overestimation of DNA reassociation values. Since S1 nuclease is known to digest the single strand tails, it is supposed to give a fairly accurate estimation of percentage of reassociated DNA. Thus, a decrease in % reassociation after S1 nuclease indicates interspersion of repetitive and single copy DNA sequences, while absence of interspersion can be inferred if no significant change is observed.

For the construction of the Cot curve of rice DNA with S1 nuclease treatment (Fig. 2.1b), rice DNA of 20 kbp fragment length was first reassociated to various Cot values in the range of 5×10^{-2} to 5×10^{3} M.s. The reassociated duplexes were then treated with S1 nuclease and the S1 nuclease resistant repetitive DNA was isolated by HAP column chromatography [14]. The extent of reassociation after S1 nuclease treatment at each Cot value was determined as described earlier. Table 2.1 gives

TABLE 2.1 : COMPARISON OF THE % REASSOCIATION OF 20 kbp RICE DNA IN THE COT RANGE 5 X 10^{-2} M.s - 5 X 10^{3} M.s

Cot values (M.s)	% Reassociation[a]	
	-S1 nuclease	+S1 nuclease
5 x 10^{-2}	27	13
1 x 10^{-1}	28	14
1	29	17
5	33	26
10	38	32
25	46	43
50	56 ←	48 ←
100	58	50
500	65	57
1000	66	61
2000	67	69

a : values obtained from Fig. 2.1a,b.

→ : indicates the value at limit Cot.

the comparison of % reassociation with and without S1 nuclease at each Cot value.

From Fig. 2.1b, it is observed that the 20 kbp (S1 nuclease resistant) Cot curve reveals the presence of two distinct components. This is comparable with the 6.5 kbp data as reported by Gupta et al [6]. Also, the percent reassociation after S1 nuclease treatment is considerably less at low Cot values (< Cot 5). At a limit Cot of 50, there is only a slight decrease from 56% to 48%.

To confirm the absence of interspersion of repetitive and single copy DNA sequences at 20 kbp and its presence amongst different repeat families, the extent of reassociation of S1 nuclease resistant DNAs of 0.5 kbp and 20 kbp fragment length was compared at Cot 0.1 M.s and Cot 50 M.s where highly and total repetitive sequences, respectively, form duplexes. From Table 2.2, it can be seen that there is about two-fold decrease in reassociation at Cot 0.1 at the higher fragment length after S1 nuclease treatment whereas, no comparable difference is found at Cot 50. This suggests extensive interspersion among highly repeated DNA sequences and of sequences reassociating between Cot 0.1 to 50. A comparison of percent reassociation with and without S1 nuclease at 20 kbp, 6.5 kbp and 0.5 kbp DNA fragment lengths, tabulated in Table 2.3, indicates an absence of interspersion of total repetitive and single copy DNA sequences.

TABLE 2.2 : COMPARISON OF DNA REASSOCIATION KINETICS DATA AT COT 0.1 M.s AND COT 50 M.s OF 0.5 kbp AND 20 kbp RICE DNA

No.	DNA	% reassociation		Magnitude of increase in reassociation of 20 kbp DNA[a]	
		Cot 0.1	Cot 50	Cot 0.1	Cot 50
1.	0.5 kbp (+S1)	8	48	-	-
2.	0.5 kbp (-S1)	9	52	-	-
3.	20 kbp (+S1)	14	48	1.75	1.00
4.	20 kbp (-S1)	28	56	3.20	1.07

a : Calculated using the formula :

$$\text{Magnitude of increase in reassociation of 20 kbp DNA} = \frac{\text{\% reassociation of 20 kbp DNA at Cot 0.1/Cot 50}}{\text{\% reassociation of 0.5 kbp DNA at Cot 0.1/Cot 50}}$$

TABLE 2.3 : S1 NUCLEASE RESISTANCE OF REPETITIVE DNA DUPLEXES AT DIFFERENT FRAGMENT LENGTHS OF RICE DNA

Fragment size in bp	% Reassociation[a] -S1 nuclease	+S1 nuclease	Reference
550 (7)	52 ± 0.15	48 ± 0.40	6
6,500 (5)	55 ± 0.10	48 ± 0.52	6
20,000 (5)	56 ± 0.26	46 ± 0.57	Present work

a : Reassociation is measured at limit Cot value of 50 i.e. the Cot value at which most of the repeated sequences reassociate.

Figures in parenthesis indicate the number of independent experiments carried out.

2.3.3 DNA sequence organization in two more varieties of rice

In order to assess the existence of any variety specific variations in the DNA sequence organization pattern, a few reassociation experiments were carried out with and without S1 nuclease treatment at the limit repetitive Cot value of 50 in two more rice varieties namely Ratnagiri-24 (R-24) and Sona which are locally available hybrid varieties of rice. The results obtained are summarised in Table 2.4 from which two observations can be made. First, the extent of DNA reassociation at Cot 50 in R-24 and Sona is similar to that of Basmati 370. Second, there is very little decrease in percent reassociation after S1 nuclease treatment at Cot 50. These data thus indicate the probable absence of interspersion of repeat and single copy DNA sequences at 20 kbp fragment length in these two varieties also.

2.4 DISCUSSION

Earlier, Britten and Davidson [18] had correlated the biological significance of DNA sequence organization with gene regulation in eukaryotes. However, according to the present concepts [5,19,20], it appears that variations in interspersion patterns need not be correlated with gene regulation as one would not expect the basic mechanism to differ in various taxa. Instead, interspersion patterns should be more related to the dynamics

TABLE 2.4 : % REASSOCIATION OF DNA WITH AND WITHOUT S1 NUCLEASE AT LIMIT REPETITIVE COT IN THREE CULTIVARS OF RICE OF 20 kbp.

Variety	% Reassociation[a]	
	-S1 nuclease	+S1 nuclease
Basmati 370 (5)	56 ± 0.26	46 ± 0.57
Ratnagiri-24 (5)	52 ± 0.10	48 ± 0.45
Sona (4)	52 ± 0.25	48 ± 0.50

a : Reassociation is measured at limit Cot value of 50 i.e. the Cot value at which most of the repeated sequences reassociate.

Figures in parenthesis indicate the number of independent experiments carried out.

of amplification and translocation of DNA sequences [21].

Much of the repetitive DNA in eukaryotes is considered to have evolved from periodic amplification events with members of the resulting sequence families being subjected to translocation, base substitution and deletion. When both amplification and deletion events occur during evolution, the portion of the genome subjected to these processes can be viewed as 'turning over' over a geological time scale. The genome size remains constant under 'steady state' conditions and the rates of amplification and deletions are equal for any one species, although different states may exist for different species. High turnover rates would lead to short period interspersion patterns, while low turnover rates would lead to long period interspersion patterns [21].

The absence of interspersion of repetitive and single copy DNA sequences at a DNA fragment length of 20 kbp in rice genome indicates that repeats and single copy DNA sequences are clustered. This is indeed a unique situation as the repetitive DNA has a general tendency to be dispersed throughout the genome rather than remaining in very long stretches without interruption by single copy DNA sequences. Absence of interspersion of repetitive DNA and single copy DNA sequences has been shown earlier in case of flax and cucumber upto fragment lengths of 10 and 5 kbp, respectively [5,10]. Recently, it has been found in

Arabidopsis that the sequence interspersion pattern is extremely long, where the estimated average length of single copy DNA sequences in about 65 kbp [22]. Considering the data in cucumber, falx and Arabidopsis, rice genome appears to be unique in its DNA sequence organization pattern. A question is raised about the nature of forces which lead to clustering of repeat and single copy DNA sequences and the probable significance of this unique arrangement. Does this situation reflect a very slow turnover of a genome? In that case, why only repeat families undergo a selective interspersion among themselves in rice?

REFERENCES

1. Flavell, R.B. and Smith, D.B. (1976)
Heredity,37 : 231-252.

2. Zimmermann, J.L. and Goldberg, R.B. (1977)
Chromosoma,(Berl.), 59 : 227-252.

3. Smith, D.B. and Flavell, R.B. (1977)
Biochim. Biophys. Acta, 474 : 82-97.

4. Murray, M.G., Cuellar, R.E. and Thompson, W.F. (1978)
Biochemistry, 17 : 5781-5790.

5. Cullis, C.A. (1981)
Biochim. Biophys. Acta, 652 : 1-15.

6. Gupta, V.S., Gadre, S.R. and Ranjekar, P.K. (1981)
Biochim. Biophys. Acta, 656 : 147-154.

7. Gupta, V.S. and Ranjekar, P.K. (1981)
J. Biosci., 3 : 417-430.

8. Bhave, M.R., Gupta, V.S. and Ranjekar, P.K. (1986)
Plant Syst. Evol., 152 : 133-151.

9. Sivaraman, L., Gupta, V.S. and Ranjekar, P.K. (1986)
Plant Mol. Biol., 6 : 375-388.

10. Lagu, M., Ranjekar, P.K. and Pillay, D.T.N. (1986)
Cell Biol. Int. Rep., 10 : 869-874.

11. Deshpande, V.G. and Ranjekar, P.K. (1980)
Hoppe-Seyler's Z. Physiol. Chem., 361 : 1223-1233.

12. Marmur, J. (1961)
J. Mol. Biol., 3 : 208-218.

13. Ranjekar, P.K., Pallotta, D. and Lafontaine, J.G. (1976).
Biochim. Biophys. Acta., 425 : 30-40.

14. Kohne, D.E. and Britten, R.J. (1971)
In : Procedures in Nucleic Acids Research, G.L. Cantoni and D.R. Davis, (Eds.), Vol. 2, pp. 500- 512.

15. Tiselius, A., Hjerten, S. and Lewin, O. (1956)
Arch. Biochem. Biophys., 65 : 132-155.

16. Ranjekar, P.K. and Murthy, M.R.V. (1973)
J. Neurochem., 20 : 1257-1264.

17. Goldberg, R.B., Crain, W.R., Ruderman, J.V., Moore, G.P., Barett, T.R., Higgins, R.C., Gelfond, R.A., Galau, G.A., Britten, R.J. and Davidson, E.H. (1975).
Chromosoma (Berl.), 51 : 225-251.

18. Britten, R.J. and Davidson, E.H. (1969)
Science, 165 : 349-357.

19. Ullrich, R.C., Kohorn, B.P. and Specht, C.A. (1980)
Chromosoma (Berl.), 81 : 371-378.

20. Wagemann, M., Epplen, J.T, Bachmann, K, Engel, J.Dand Schmidtke, J. (1981)
Experientia, 37 : 1274-1276.

21. Thompson, W.F. and Murray, M.G. (1980)
In : Proceedings of the 4th John Innes Symposium. The Plant Genome, pp. 31-45, D.R. Davis and D.A. Hopwood (Eds.), Norwich, England.

22. Pruitt, R.E. and Meyerowitz, E.M. (1986)
J. Mol. Biol., 187 : 169-183.

CHAPTER 3

ORGANIZATION AND PROPERTIES OF REPEATED DNA SEQUENCES IN RICE

ABSTRACT

Thermal stability experiments revealed that as compared to total rice DNA (Tm 85.1°C), the Cot 0.1 and Cot 50 DNAs had higher Tm values of 95.5°C (G + C 63.9%) and 88.9°C (G + C 47.8%) respectively, indicating that rice repetitive DNA sequences were highly G + C rich. Analysis by agarose gel electrophoresis showed the presence of long (4-20 kbp) and short (0.2-0.3 kbp) DNA sequences in both the repetitive DNA fractions. The long and short repeats of Cot 0.1 DNA showed a monophasic curve with a Tm of 94.8°C and 92.9°C, respectively. The long repeats of Cot 50 had a Tm of 97.5°C while the short repeats showed a biphasic curve with a major low melting fraction (Tm 83.7°C) and a minor high melting fraction (Tm 98°C). Restriction enzyme analysis of long repeats of Cot 0.1 and Cot 50 also confirmed the presence of G + C rich sequences.

To assess whether the short and long repeats in rice consisted of same/similar type of repeat sequences or they were independent of each other, Southern hybridisation experiments were done using long and short repeats as probes. It was found that the long and short repeats of Cot 0.1 were not homologous to each other, whereas the long and short repeat sequences reassociating by Cot 50 did show some homology. Based on these data four types of repeat DNA sequences namely long repeats (9-20 kbp) reassociating by Cot 0.1 and Cot 50 and short repeats (0.2-0.3 kbp) reassociating by Cot 0.1 and Cot 50 were identified.

3.1 INTRODUCTION

The arrangement of repeated and single copy DNA has been studied in a number of plant species [1-8]. Based on these studies, repetitive DNA sequences have been categorised into long and short type of sequences with fragment lengths in the range of 1.5-10 kbp and 0.2-1.0 kbp respectively. In animals, repeat sequences have been characterized extensively in Drosophila and Sea urchin with respect to their genomic distribution, sequence organization and evolution [9-11]. Though, the genome organization patterns in plants have been reviewed extensively [12-14], very little experimental data is available about their repetitive DNA sequences. In the earlier chapter (Chapter 2), I have described the novel DNA sequence organization of rice genome at a DNA fragment length of 20 kbp. In this Chapter, I include the data on thermal stability, restriction enzyme analysis and homology studies of long and short repeat sequences in rice (cv. Basmati 370) repetitive DNA.

3.2 MATERIALS AND METHODS

3.2.1 Chemicals, materials and enzymes

Calf thymus DNA and all the four deoxyribonucleotide triphosphates were from Sigma Chemical Co., USA.

Radiolabelled α-^{32}P-dATP was obtained from Radiochemical Centre, Amersham, UK.

Nitrocellulose papers from Advanced Microdevices, Ambala, India (MDI, type SCN, 0.45 microns) were used for Southern blotting.

Black and white ORWO or ILFORD photographic film (35 mm, 125 ASA) and X-ray film AGFA Curix RPI were used for gel photography and autoradiography, respectively.

All the restriction endonucleases were either from Bethesda Research Laboratories (BRL), USA or New England Biolabs (NEB), USA.

3.2.2 Isolation of highly repetitive and total repetitive DNA

The highly repetitive (Cot 0.1) and total repetitive (Cot 50) DNA fractions were isolated from native, unsheared rice DNA (20 kbp). The DNA was dissolved in 0.18 M NaCl containing 0.006 M Pipes buffer (pH 6.8) and was denatured for 10 min followed by reassociation at 62°C to Cot 1.0×10^{-1} M.s and 5.0×10^{1} M.s. After reassociation, DNA samples were adjusted to 25 mM sodium acetate (pH 4.5), 0.1 mM $ZnSO_4$ and 25 mM 2-mercapto ethanol [15]. S1 nuclease (0.1 U/ug of DNA) was then added and the samples were incubated for 15 min at 37°C to digest single-stranded DNA. The reaction was terminated by treatment with chloroform : isoamyl alcohol (24:1 v/v mixture) and S1 nuclease resistant repetitive DNA duplexes were precipitated with 2 volumes of chilled ethanol.

The S1 nuclease resistant DNA samples were sized by neutral agarose slab gel electrophoresis using 0.7% agarose gels in TAE buffer (40 mM Tris, 20 mM acetate, 2 mM EDTA), pH 8.1 at a constant current of 40 mA. The gels were stained with ethidium bromide (1 ug/ml) in dark for 30 min and visualised on a long wavelength 302 nm UV transilluminator (UV products, San Gabriel, California) and photographed with a 35 mm SLR camera (Minolta X700 with macrophotography and zoom lens system) using a red filter. λ DNA digested with HindIII and ϕX174 RF DNA digested with HaeIII were used as molecular weight markers. Both the Cot 0.1 and Cot 50 DNAs showed the presence of long repeats (4-20 kbp) and short repeats (< 0.5 kbp). To narrow down the wide range of long repeats, the DNA fragments in the size range of 9-20 kbp were used for all the work. The fragments of size 4-9 kbp were termed as the intermediate repeats.

Two regions corresponding to long (9-20 kbp) and short (< 0.5 kbp) repeats were cut and the agarose pieces were submerged in phenol equilibrated with tris buffer and frozen at -70°C for 2 h. The frozen gel pieces and phenol were thawed out gradually to 4-10°C and aqueous layer was collected by centrifugation (10,000 rpm, 4°C, 10 min). This was followed by chloroform : isoamyl alcohol (24:1 v/v) treatment to remove traces of phenol. The ethidium bromide from the aqueous layer was removed by several extractions with water-saturated butanol. The DNA was precipitated by adding 3 M sodium acetate, pH 5.3 to a

final concentration of 0.3 M and two volumes of chilled ethanol. The DNA precipitate was allowed to aggregate at -20°C for 14-16 h or at -70°C for 2 h. It was then collected after centrifugation, dried and dissolved in 10 mM tris-HCl, pH 7.4.

3.2.3 Thermal denaturation of DNA

To determine the DNA heterogeneity in base composition and G + C content, thermal denaturation analysis of DNA was carried out in Beckmann DU-8B UV-Visible spectrophotometer equipped with a Tm programme cassette. The melting temperatures (Tm) and G + C contents were calculated according to Mandel and Marmur [16].

3.2.4 Restriction endonuclease digestions of DNA and gel electrophoresis

Digestions of long repeat DNA sequences by restriction enzymes and analysis of DNA digests by agarose gel electrophoresis were according to Maniatis et al [17].

For the analysis of short repeat digests, samples were eletrophoresed on 8% or 10% polyacrylamide vertical slab gels in TBE(X) (0.089 M tris borate, 0.089 M boric acid, 0.002 M EDTA), pH 8.0 buffer at a constant current of 20 mA for 3-4 h [17]. The gels were stained and photographed as described earlier.

3.2.5 Southern hybridization and autoradiography

The long and short repetitive sequences were labelled using α-^{32}P-dATP by nick translation according to Rigby et al [18]. Transfer of DNA from agarose gels to nitrocellulose membrane, prehybridization, hybridization and washing of the blots to remove unhybridized radioactive material were according to Maniatis et al [17].

The dry filters were mounted on a thin plastic paper and fixed in a X-ray cassette with Agfa Curix X-ray film and intensifying screens. After exposure for about 24 h at -70^{o}C, the film was developed with Agfa X-ray developer, washed, dried and then photographed.

3.3 RESULTS

3.3.1 Occurrence of long and short repeats in rice repetitive DNA

When rice Cot 50 DNA is analysed by agarose gel electrophoresis, it reveals the presence of both long (4-20 kbp) and short (0.2-0.3 kbp) DNA sequences which account for 60% and 40% respectively of the total Cot 50 DNA (Fig. 3.1 A,C). The Cot 0.1 fraction which is considered to consist of mainly highly repetitive DNA sequences also shows the occurrence of the above two types of repeats except that the short repeats represent 60% of the total Cot 0.1 DNA (Fig. 3.1 A,B).

FIGURE 3.1 : ELECTROPHORESIS OF HIGH MOLECULAR WEIGHT RICE DNA WITH HIGHLY REPETITIVE (COT 0.1) AND TOTAL REPETITIVE (COT 50) DNAs AND THEIR LONG AND SHORT REPEATS

(A) High molecular weight rice DNA (lane a)

λ HindIII digest (lane b)
Cot 0.1 DNA (lane c)
Cot 50 DNA (lane d)
ϕX174 RF DNA HaeIII digest (lane e)

(B) Long and short repeats isolated from Cot 0.1 DNA

λ HindIII digest (lane a)
Long repeats (lane b)
Intermediate repeats (lane c)
Short repeats (lane d)

(C) Long and short repeats isolated from Cot 50 DNA

Long repeats (lane a)
λ HindIII digest (lane b)
Intermediate repeats (lane c)
Short repeats (lane d)

Electrophoresis was carried out on 1.0% (A) and 1.4% (B and C) neutral agarose slab gels in TAE buffer (pH 8.1) at a constant current of 30 mA.

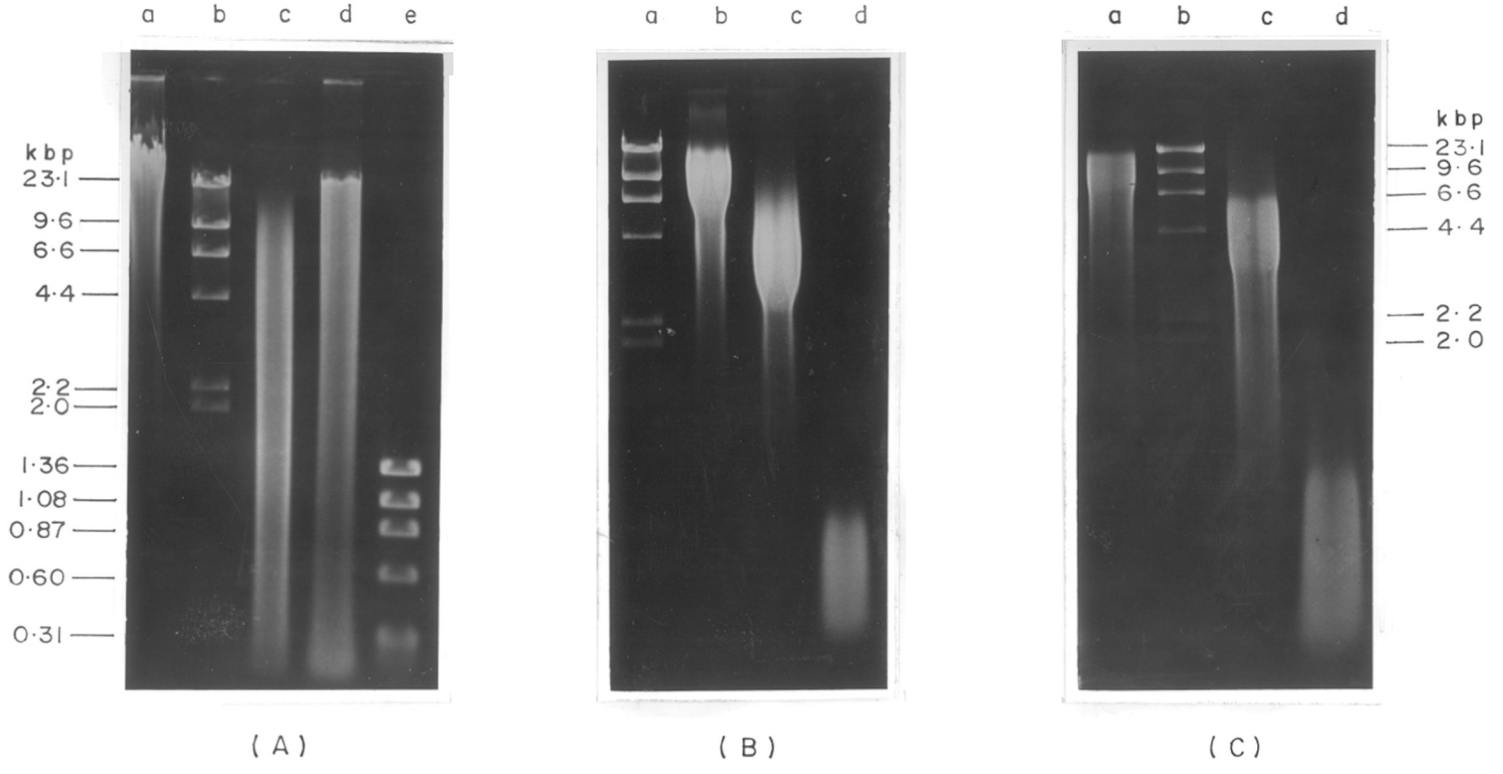

a b c d e
kbp
23·1
9·6
6·6
4·4
2·2
2·0
1·36
1·08
0·87
0·60
0·31
(A)
a b c d
(B)
a b c d
kbp
23·1
9·6
6·6
4·4
2·2
2·0
(C)

FIG. 3·1

3.3.2 Thermal stability of long and short repeats

As an initial step towards characterization of long and short repeats, their thermal stability was first determined.

The 1°C/min melting curves of the total repetitive DNAs and their repeat fractions are shown in Fig. 3.2 and data are tabulated in Table 3.1. The high molecular weight rice DNA has a Tm of 85.1°C. The Cot 0.1 and Cot 50 DNAs on the other hand have higher Tm values of the order of 95.5°C (G + C 63.9%, $\sigma^{2/3}$ 13.0) and 88.9°C (G + C 47.8%, $\sigma^{2/3}$ 16.7) respectively. This indicates that repetitive DNA sequences in rice are G + C rich. It is interesting to note that highly repetitive (Cot 0.1) DNA in particular is very thermostable.

From Fig. 3.2 it is also clear that the long and short repeats of Cot 0.1 DNA show a monophasic curve with a Tm 0f 94.8°C (G + C 62.2%, $\sigma^{2/3}$ 5.0) and 92.9°C (G + C 57.5%, $\sigma^{2/3}$ 14.5) respectively. The long repeats of Cot 50 DNA have a Tm of 97.5°C while its short repeats show a biphasic curve with a major low melting fraction of Tm 83.7°C (G + C 35.1%) and a minor high melting fraction of Tm 98°C (G + C 70.0%). These results show that the repeat sequences in rice, in general, are highly thermostable except the short repeats of Cot 50.

FIGURE 3.2 : THERMAL DENATURATION PROFILES OF RICE REPETITIVE DNAs

The DNA melting was carried out in 0.012 M PB, pH 6.8 at a heating rate of $1^{o}C$/min. A correction factor of $14^{o}C$ was added to get the actual Tm value.

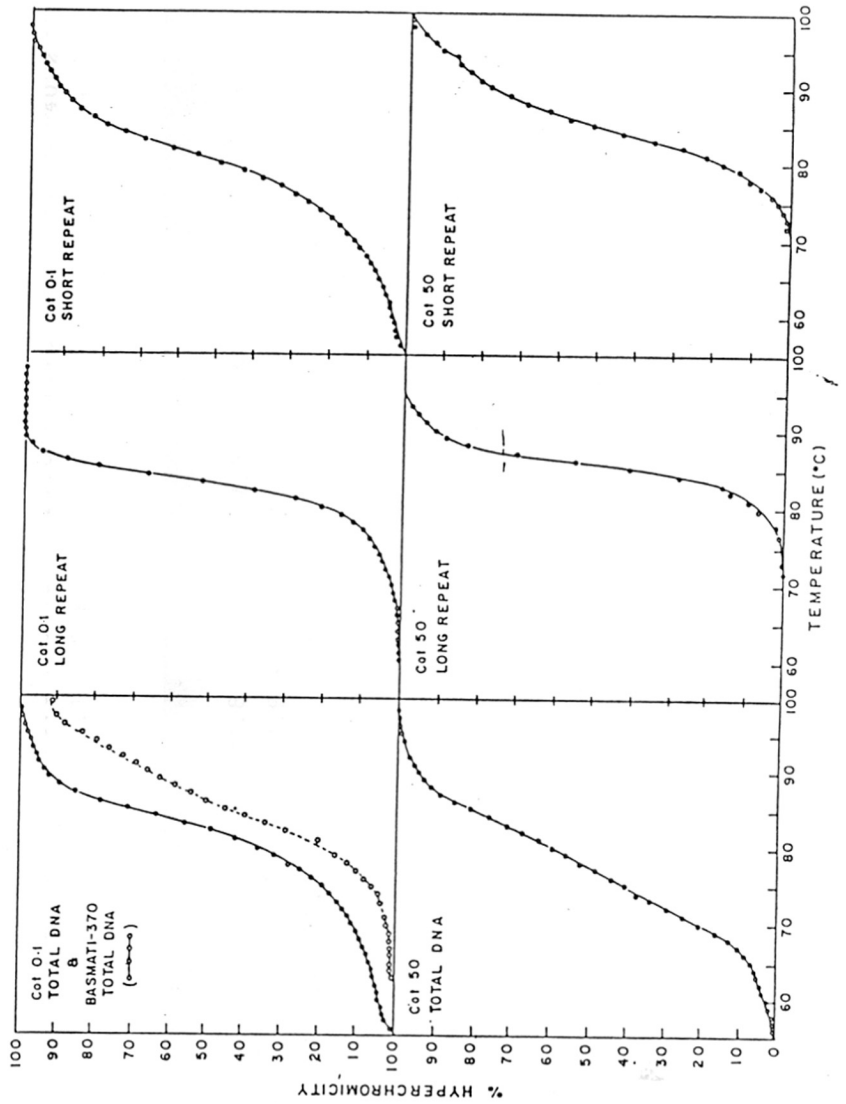

Col 0-1
TOTAL DNA
&
BASMATI-370
TOTAL DNA
(•—•—•)
Col 0-1
LONG REPEAT
Col 0-1
SHORT REPEAT
Col 50
TOTAL DNA
Col 50
LONG REPEAT
Col 50
SHORT REPEAT
% HYPERCHROMICITY
TEMPERATURE (°C)

FIG. 3.2.

TABLE 3.1 ; THERMAL DENATURATION PROPERTIES OF RICE REPETITIVE DNAs

Fraction	Proportion in the Genome	Tm ^{o}C[a]	Hyperchromicity[b]	G + C content[c]
	%		%	%
Calf thymus DNA (4)	-	85.8 ± 0.75	27.8 ± 1.18	40.2
Total rice DNA (5)		85.1 ± 0.8	24.4 ± 1.1	38.5
Highly repetitive DNA (5) (Cot 1.0 x 10^{-1} M.s)	15.0	95.5 ± 1.38	20.0 ± 1.5	63.9
Cot 0.1				
Long repeat (4)	6.0	94.8 ± 0.21	27.0 ± 2.6	62.2
Short repeat (4)	9.0	92.9 ± 1.06	21.5 ± 1.6	57.5
Total repetitive DNA (5) (Cot 5.0 x 10^{1} M.s)	50.0	88.9 ± 1.4	17.7 ± 1.7	47.8
Cot 50				
Long repeat (4)	27.0	97.5 ± 0.5	16.8 ± 0.21	68.8
Short repeat (4)	20.0 LMF	83.7 ± 0.35 LMF	11.4 ± 0.9	35.1 LMF
	3.0 HMF	98 ± 0.35 HMF		70.0 HMF

[a] Temperature corresponding to 50% hyperchromicity

[b] Calculated using the formula $H = \frac{A_{260}(98^{o}C) - A_{260}(62^{o}C)}{A_{260}(98^{o}C)} \times 100$

[c] Calculated using the formula % G + C = (Tm - 69.3) x 2.44

LMF and HMF stands for low melting fraction and high melting fraction respectively.

Figures in parenthesis indicate the number of experiments carried out.

3.3.3 Restriction enzyme analysis of rice repetitive DNA

After getting an idea about the thermal stability of long and short repetitive sequences, they were next characterized by restriction endonuclease analysis. Figure 3.3 A,B includes the digestion patterns of long repeats derived from Cot 0.1 and Cot 50 with mainly G + C rich sequence specific enzymes which include MspI, HpaII, HaeIII and HhaI. From this figure, it can be seen that these enzymes show digestion of repetitive DNAs, thus confirming the presence of GC rich sequences. Comments regarding the occurrence of discrete bands and variations in the extent of DNA digestion in case of MspI and HpaII are included in Chapter 5.

When long repeat sequences reassociating by Cot 0.1 and Cot 50 are digested with AvaI and EcoRV, band patterns are observed with both the fractions (Fig. 3.4). AvaI shows two bands of 4.3 kbp and 4.9 kbp while EcoRV shows a single band of 4.6 kbp in both the cases. This indicates that 5'-GPyCGPuG-3' and 5'-GATATC-3' type of sequences are present at regularly spaced intervals in rice genome.

Figure 3.5 A,B shows the restriction enzyme digestion patterns of short repeat sequences of both Cot 0.1 and Cot 50. By polyacrylamide gel electrophoresis, it can be seen that the short repeats are too small in size to reveal any distinct differences.

FIGURE 3.3 : DIGESTION OF LONG REPEATS OF COT 0.1 (A) AND COT 50 (B) DNAs DIGESTED WITH DIFFERENT RESTRICTION ENZYMES

λ HindIII digest (lane a)
AluI digest (lane b)
MspI digest (lane c)
HpaII digest (lane d)
HaeIII digest (lane e)
HhaI digest (lane f)

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

In Fig. 3.3 (A) the occurrence of a single band in HaeIII digest and 2 bands in HhaI digest indicate the presence of discrete repeat families in rice genome.

FIGURE 3.4 : DIGESTION OF LONG REPEATS OF COT 0.1 AND COT 50 DNAs WITH AvaI AND EcoRV

λ HindIII digest (lane a)
Cot 0.1 AvaI digest (lane b)
Cot 0.1 EcoRV digest (lane c)
Cot 50 AvaI digest (lane d)
Cot 50 EcoRV digest (lane e)

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

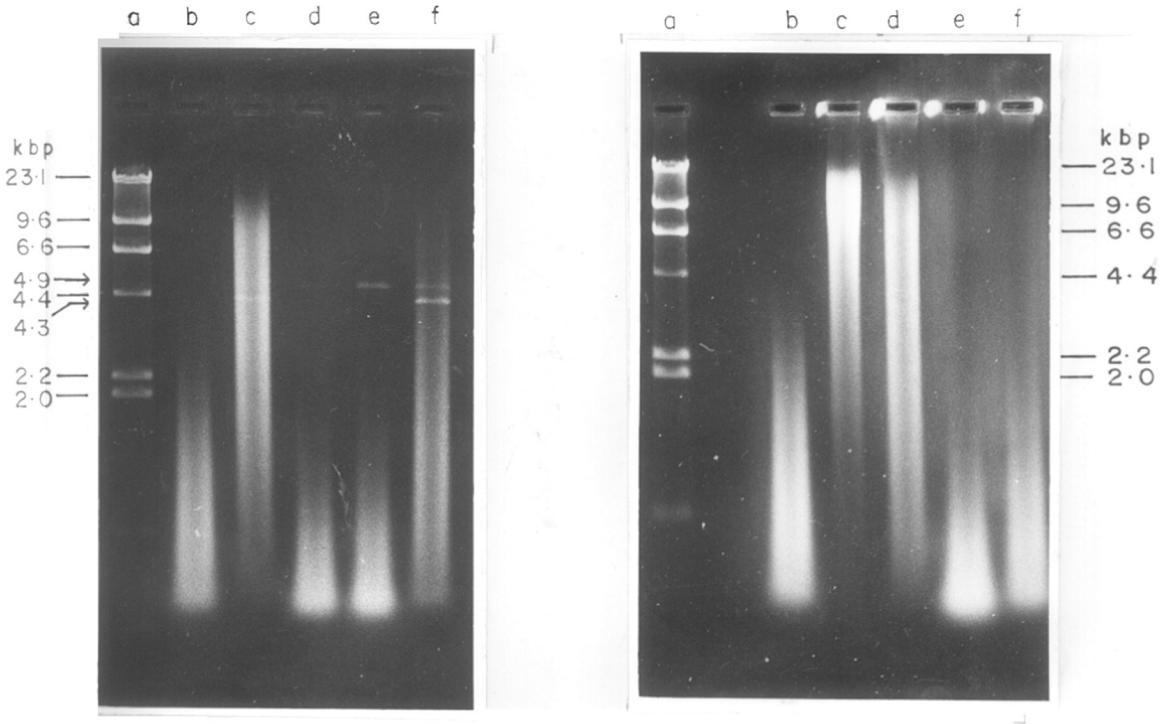

a b c d e f
kbp
23·1
9·6
6·6
4·9
4·4
4·3
2·2
2·0
(A)
a b c d e f
kbp
23·1
9·6
6·6
4·4
2·2
2·0
(B)

FIG. 3·3

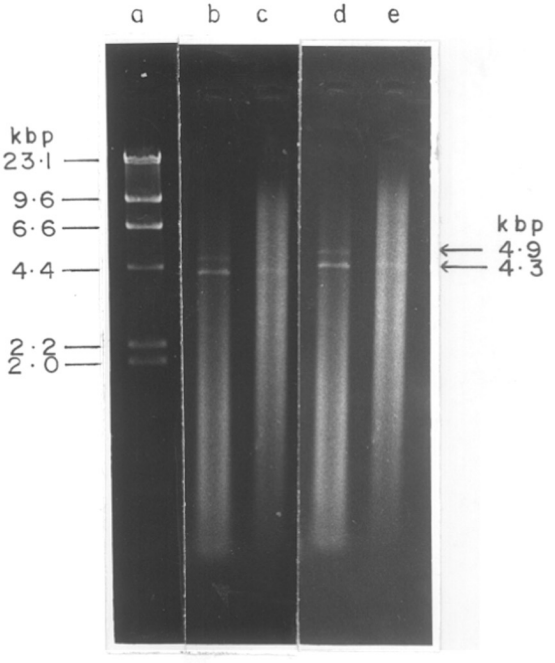

a b c d e
kbp
23·1
9·6
6·6
4·4
2·2
2·0
kbp
4·9
4·3

FIG. 3·4

FIGURE 3.5 : DIGESTION OF SHORT REPEATS OF COT 0.1 (A) AND COT 50 (B) DNAs WITH DIFFERENT RESTRICTION ENZYMES

(A)
- ϕX174 HaeIII digest (lane a)
- KpnI digest (lane b)
- SalI digest (lane c)
- AluI digest (lane d)
- TaqI digest (lane e)
- EcoRI digest (lane f)
- HindIII digest (lane g)
- HaeIII digest (lane h)
- HhaI digest (lane i)
- XbaI digest (lane j)
- XhoI digest (lane k)
- Control Cot 0.1 short repeat (lane l)

(B)
- KpnI digest (lane a)
- SalI digest (lane b)
- AluI digest (lane c)
- TaqI digest (lane d)
- EcoRI digest(lane e)
- HindIII digest (lane f)
- HaeIII digest (lane g)
- HhaI digest (lane h)
- XbaI digest (lane i)
- XhoI digest (lane j)
- ϕX174 HaeIII digest (lane k)

Electrophoresis was carried out on 8% neutral polyacrylamide slab gels in TBE buffer (pH 8.0) at a constant current of 20 mA.

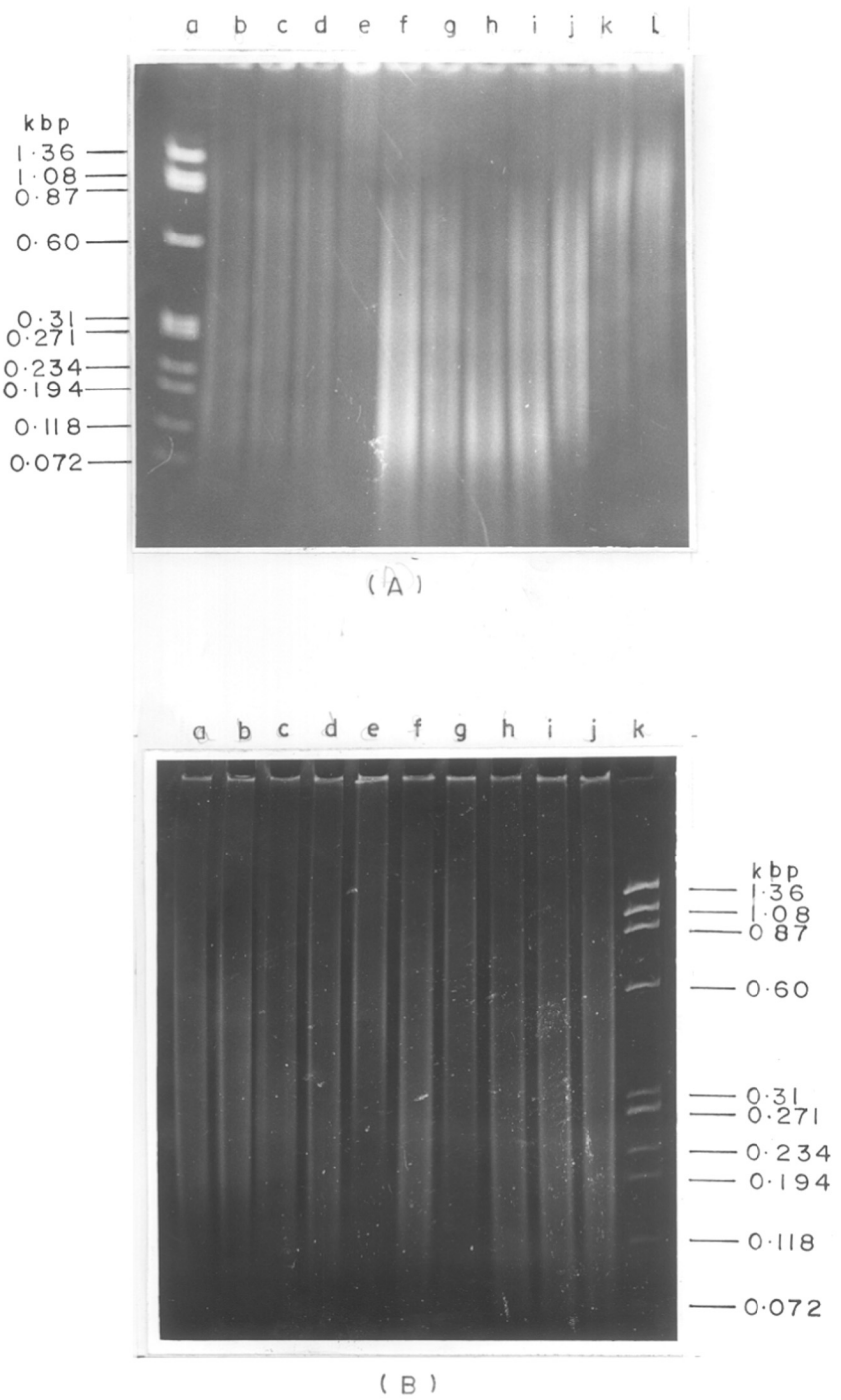

a b c d e f g h i j k l
kbp
1·36
1·08
0·87
0·60
0·31
0·271
0·234
0·194
0·118
0·072
(A)
a b c d e f g h i j k
kbp
1·36
1·08
0·87
0·60
0·31
0·271
0·234
0·194
0·118
0·072
(B)

FIG. 3·5.

3.3.4 DNA sequence homology between long and short repeats

From the thermal stability data and restriction enzyme analysis of Cot 0.1 and Cot 50 DNA fractions, it is clear that rice repetitive DNA consists of different stretches of repeat families showing variations in their length, thermal stability and sequence heterogeneity. To assess whether these repeat families belong to the same/similar type of repeat sequences or they are independent of each other, Southern hybridization experiments were undertaken using short as well as long repeats of Cot 0.1 and Cot 50 DNAs as probes.

From Fig. 3.6 A,B it is observed that short repeats of Cot 0.1 show very little hybridization to its long repeats whereas strong hybridization is seen with themselves. However, the same probe shows strong hybridization to long repeats of Cot 50 DNA, while weak hybridization is seen with its short repeats. These results are confirmed when the long repeats of Cot 0.1 are used as probe (Fig. 3.7 A,B). This suggests that the long and short repeats reassociating by Cot 0.1 are independent and belong to different repeat families.

Figure 3.8 A,B shows the hybridization pattern when short repeats of Cot 50 are used as a probe. Here hybridization is seen with its own long repeats, while very little hybridization is seen with short as well as long repeats of Cot 0.1. This indicates that the long and short repeats reassociating

FIGURE 3.6 : AUTORADIOGRAMS SHOWING THE HOMOLOGY OF COT 0.1 SHORT REPEATS WITH THE LONG AND SHORT REPEATS OF COT 0.1 (A) AND COT 50 (B)

(A)
- Cot 0.1 short repeats (lane a)
- Cot 0.1 intermediate repeats (lane b)
- λ HindIII digest (lane c)
- Cot 0.1 long repeats (lane d)

(B)
- Cot 50 long repeats (lane a)
- Cot 50 intermediate repeats (lane b)
- Cot 50 short repeats (lane c)

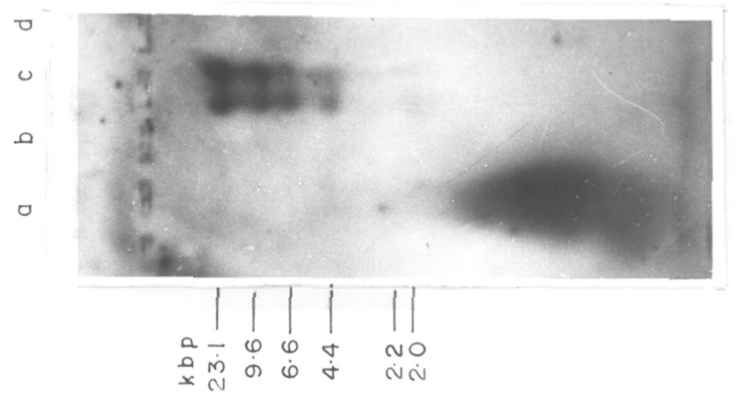


FIG. 3·6

FIGURE 3.7 : AUTORADIOGRAMS SHOWING THE HOMOLOGY OF COT 0.1 LONG REPEATS WITH THE LONG AND SHORT REPEATS OF COT 0.1 AND COT 50 DNAs

(A) (B)

(A) Cot 0.1 short repeats (lane a)
Cot 0.1 intermediate repeats (lane b)
Cot 0.1 long repeats (lane c)
λ HindIII digest (lane d)

(B) Cot 50 short repeats (lane a)
Cot 50 intermediate repeats (lane b)
Cot 50 long repeats (lane c)
λ HindIII digest (lane d)

FIGURE 3.8 : AUTORADIOGRAMS SHOWING THE HOMOLOGY OF COT 50 SHORT REPEATS WITH THE LONG AND SHORT REPEATS OF COT 0.1 AND COT 50 DNAs

(A) (B)

(A) Cot 0.1 long repeats (lane a)
Cot 0.1 intermediate repeats (lane b)
Cot 0.1 short repeats (lane c)

(B) Cot 50 long repeats (lane a)
Cot 50 intermediate repeats (lane b)
Cot 50 short repeats (lane c)

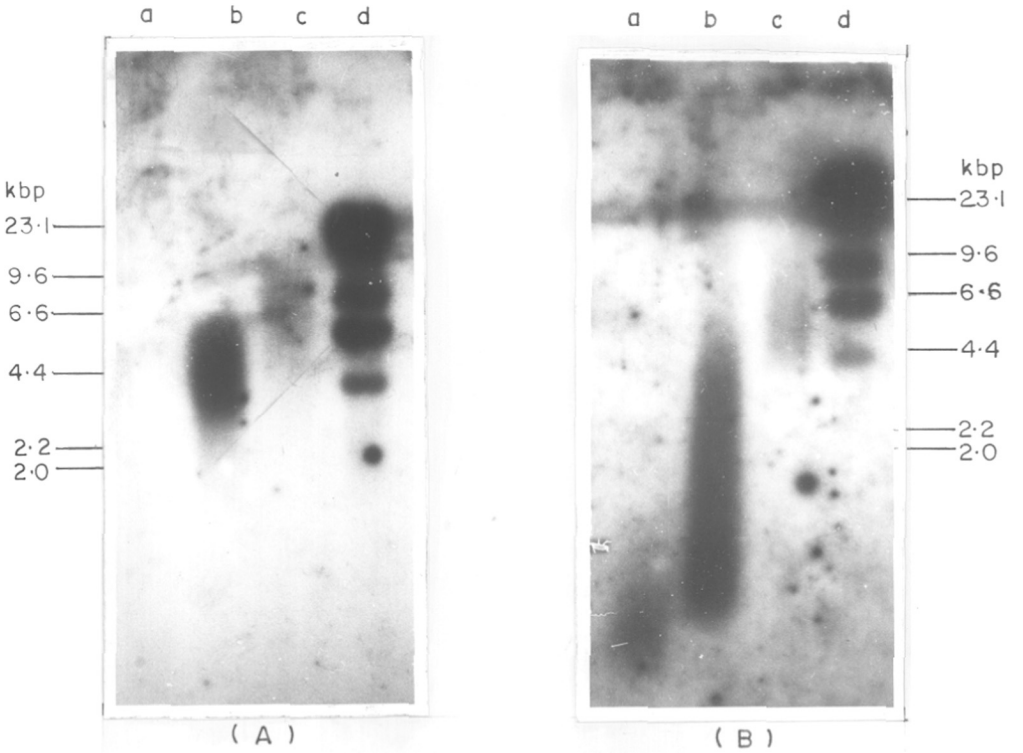


FIG. 3·7.

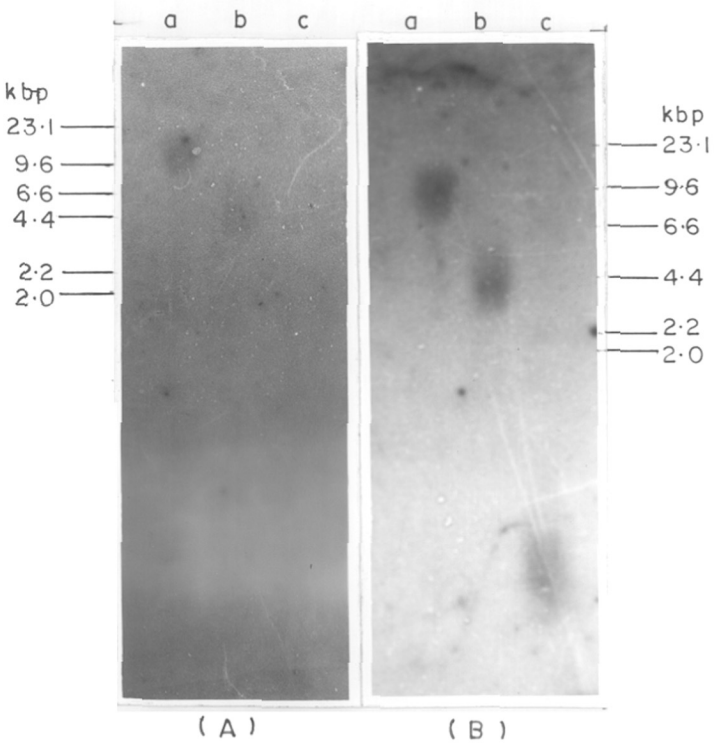


FIG. 3·8.

by Cot 0.1 to 50 may belong to similar/same type of repeat families which have undergone very little sequence divergence during the course of evolution. These results are summarised in Table 3.2.

3.4 DISCUSSION

Characterization of rice repetitive DNA by restriction enzymes has been reported earlier [19] and a basic repeat unit of about 300-800 bp is detected after EcoRI digestion. The rRNA genes have also been analyzed in rice and it is shown that the size of the rice rDNA unit is about 7800-8000 bp [20]. The present work describes the occurrence of different long and short repeat sequences in rice genome and attempts have been made to assess their homology with each other.

The first important feature of rice repetitive DNA is its high thermal stability. As shown in Table 3.1, most of the repetitive classes have Tm values much higher than the native rice DNA. Although high melting DNA fractions have been noticed in a few other plant species [21], the occurrence of a substantial amount of high melting fraction (about 30 % of the total DNA) in rice is particularly very striking. The high thermal stability of rice repetitive DNA can be interpreted to indicate a very low rate of sequence divergence or a recent origin of repeats.

TABLE 3.2 : HOMOLOGY BETWEEN LONG AND SHORT REPEATS OF RICE REPETITIVE DNAs

(A) Probe : Cot 1.0×10^{-1} M.s (Short repeats)

Repetitive DNA	Long repeat	Short repeat
Cot 0.1	-	+
Cot 50	+ (Strong)	- (Weak)

(B) Probe : Cot 5.0×10^{1} M.s (Short repeat)

Repetitive DNA	Long repeat	Short repeat
Cot 0.1	- (Weak)	-
Cot 50	+	+

Secondly, based on the hybridization and melting data, rice repetitive DNA consists of the following classes of repeat sequences :

1. Short repeats reassociating by Cot 0.1 (Tm= 92.9^{o}C) with no homology to the long repeats reassociating by Cot 0.1.
2. Long repeats reassociating by Cot 0.1 (Tm=94.8^{o}C).
3. Short repeats reassociating by Cot 50 with homology to long repeats reassociating by Cot 50; and consisting of a major low melting fraction (Tm=83.7^{o}C).
4. Long repeats reassociating between Cot 0.1 to 50 (Tm=97.5^{o}C)

From the above discussion, it is clear that long and short repeats reassociating by Cot 50 show homology to each other, indicating that they constitute same or similar type of repeat families. Moreover, these repeats appear to have undergone a very low degree of sequence divergence. As reported in a few plant species [22,23], it is found that short repeats are more diverged than long repeats. Hence, it is generally assumed that short repeats evolve from long repeats. If this is so, then the long repeats should exhibit higher thermal stability because they are relatively recent in origin and have got little time to diverge. This is exactly the case in the repeats reassociating by Cot 50 in the rice genome. As against this, there are evidences which show that short repeats evolve from shorter sequences such as tRNA or 5S RNA [24,25].

The short and long repeats reassociating by Cot 0.1 do not show any hybridization with each other, suggesting that they belong to different repeat families or have independently diverged.

The Cot 0.1 short repeats do not show any hybridization with short repeats of Cot 0.1 to 50. This is rather an unexpected observation. It suggests that the Cot 0.1 short repeats may be interspersed with other repeat families which do not reassociate rapidly at Cot 0.1. This indicates that rapidly reassociating sequences are not interspersed with very slow or single copy sequences. Similar observation of interspersion of repetitive sequences belonging to different kinetic classes amongst themselves have been reported earlier [26,27].

Finally, the rice repetitive DNA does not show any organized structure with most of the enzymes used except AvaI and EcoRV. As compared to the available data on tandem and dispersed repeat families in plants, it is for the first time that the occurrence of families is shown in rice with these two enzymes.

Thus, from the above discussion, it is clear that we have got some information about a gross interrelationship of long and short repeat sequences in the rice genome. Finer analysis can be achieved by cloning and sequence analysis of a specific DNA fragment from these repetitive classes. These data are described in the next Chapter.

REFERENCES

1. Flavell, R.B. and Smith, D.B. (1976)
Heredity, 37 : 231-252.

2. Zimmerman, J.L. and Goldberg, R.B. (1977)
Chromosoma, (Berl.), 59 : 227-252.

3. Smith, D.B. and Flavell, R.B. (1977)
Biochim. Biophys. Acta, 474 : 82-97.

4. Gupta, V.S., Gadre, S.R. and Ranjekar, P.K. (1981)
Biochim. Biophys. Acta, 656 : 147-154.

5. Gupta, V.S. and Ranjekar, P.K. (1982)
Ind. J. Biochem., Biophys., 19 : 167-170.

6. Wimpee, C.F. and Rawson, J.R.Y. (1979)
Biochim. Biophys. Acta, 562 : 192-206.

7. Murray, M.G., Palmer, J.D., Cuellar, R.E. and Thompson, W.F. (1979)
Biochemistry, 18 : 5259-5266.

8. Cullis, C.A. (1981)
Biochim. Biophys. Acta, 652 : 1-15.

9. Crain, W.R., Eden, F.C., Pearson, W.R., Davidson, E.H. and Britten, R.J. (1976)
Chromosoma (Berl.), 56 : 309-326.

10. Klein, W.H., Thomas, T.L., Lai, C., Schellar, R.H., Britten, R.J. and Davidson, E.H. (1978)
Cell, 14 : 889-900.

11. Moore, G.P., Schellar, R.H., Davidson, E.H. and Britten, R.J. (1978)
Cell, 15 : 649-660.

12. Walbot, V. and Goldberg, R. (1979)
In : Nucleic Acids in Plants, T.C. Hall and J.W. Davies (Eds.), CRC Press, Baco Raton, Florida, 3-40.

13. Thompson, W.F. and Murray, M.G. (1981)
In : The nuclear genome : Structure and Function, Marcus A. (Ed.), The Biochemistry of Plants. A Comprehensive Treatise, Vol. 6, Proteins and nucleic acids, Marcus Academic Press, New York, pp. 1-18.

14. Sorenson, J.C. (1984)
In : The structure and expression of nuclear genes in higher plants, Advances in Genetics, 22, 109-144.

15. Goldberg, R.B., Grain, W.R., Ruderman, J.V., Moore, G.P., Barette, T.R., Higgins, R.C., Gelfond, R.A., Galau, G.A., Britten, R.J. and Davidson, E.H. (1975)
Chromosoma (Berl.), 51 : 225-251.

16. Mandel, M. and Marmur, J. (1968)
In : Methods in Enzymology, 12B, Academic Press, New York, pp. 195-206.

17. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982)
In : Molecular Cloning : A laboratory Manual, Cold Spring Harbor University, New York.

18. Rigby, P.W.J., Dieckman, M., Rhodes, C. and Berg, O. (1977)
J. Mol. Biol., 113 : 237-251.

19. Pental, D. and Barnes, S.R. (1985)
Theor. Appl. Genet., 70 : 185-191.

20. Olmedilla, A., Delcasso, D. and Delseny, M. (1984)
Plant Sci. Lett., 37 : 123-127.

21. Sivaraman, L. and Ranjekar, P.K. (1984)
Ind. J. Biochem. Biophys., 21 : 299-303.

22. Goldberg, R.B. (1978)
Biochem. Genet., 16 : 45-68.

23. Kiper, M. and Herzfeld, H. (1978)
Chromosoma (Berl.), 65 : 335-351.

24. Grellet, F., Delcasso, D., Panabieres, F. and Delseny, M. (1986)
J. Mol. Biol., 187 : 495-507.

25. Benslimane, A.A., Dron, M., Hartman, C. and Rode, A. (1986)
Nuc. Acids Res., 14 : 8111-8119.

26. Sivaraman, L., Gupta, V.S. and Ranjekar, P.K. (1986)
Plant Mol. Biol., 6 : 375-388.

27. Bhave, M.R., Gupta, V.S. and Ranjekar, P.K. (1986)
Plant Syst. Evol., 152 : 133-151.

CHAPTER 4

MOLECULAR CLONING, RESTRICTION ENZYME ANALYSIS AND SEQUENCING OF A SPECIFIC LONG REPETITIVE DNA SEQUENCE IN RICE

ABSTRACT

Rice long repetitive DNA (9-20 kbp) reassociating by Cot 50 M.s was cloned in pBR325. Out of several recombinants (Cam^r $Amp^r Tet^s$), only a few were selected randomly for further characterization. The insert size in all these clones was 3-4 kbp. Restriction enzyme analysis showed absence of EcoRI and Bc1I sites, presence of a single PstI and PvuII site and multiple sites for AluI in three clones namely pRL1, pRL7 and pRL10.

The BamH1-PstI fragment of about 0. 4 kbp in the pRL7 insert DNA (pRL7-0.4 kbp) was subcloned in M13mp18 and partially sequenced using Sanger's dideoxynucleotide chain termination method. Dot matrix comparison of this sequence with rice rDNA sequences revealed appreciable homology with the 25S.rDNA sequence of rice

4.1 INTRODUCTION

In recent years, a large number of repetitive DNA sequences have been cloned and characterised from a number of plant species. Their species-specific nature has been clearly demonstrated in closely as well as distantly related plants [1-6]. Also, specific cloned repetitive sequences have been used as markers to study phylogenetic relationships [7]. Besides cloning, a few repeat families have also been sequenced [8-10] and these studies have revealed conservation/divergence of specific repeat elements.

In the earlier chapters, I have described a unique pattern of genome organization in rice wherein repetitive DNA sequences are arranged as tandem arrays of length as high as 20 kbp. Furthermore, I have also shown the presence of four different types of repeat families in rice repetitive DNA. Out of these four families, the long repeats reassociating by Cot 50 showed some homology to the long and short repeat sequences reassociating by Cot 0.1 and to the short repeats of Cot 50 itself. As a continuation of this study, I have carried out cloning and sequence analysis of long repeats reassociating by Cot 50 and these data are included in this chapter.

4.2 MATERIALS AND METHODS

Vector pBR325 and host E. coli DH1 were used for plasmid cloning and M13mp18 E. coli JM101 system was used for subcloning in M13. The details regarding genotypes, selection markers, size and restriction enzyme sites are described below. The restriction endonuclease maps of vector pBR325 and M13mp18 are also given.

4.2.1 Vector-host systems used for cloning

Hosts name	Genotype	Comments
1. E. coli DH1	F^-, recA1, endA1, gyr A96, thi-1 hsdR17, supE44	An MM294 derivative that is nalidixic acid resistant (gyrA), recA and suppressor containing
2. E. coli JM101	supE, thi, (lacproA,B)/ F1', tra D36, proA,B lac I^q, M15, (r_k^+, m_k^+)	Yannisch-Perron, C., et al [11]

Vector name	Selection	Size	Unique Cloning sites	Refs
1. pBR325 (Fig.4.1A)	Cam^r, Amp^r Tet^s	5.9 kbp	EcoRI, PvuI, PstI, BamH1, SalI, EcoRV SphI, NruI, AvaI	Bolivar et al [12]

FIGURE 4.1 : RESTRICTION ENDONUCLEASE MAP OF

(A) pBR325 DNA

(B) M13mp18 DNA

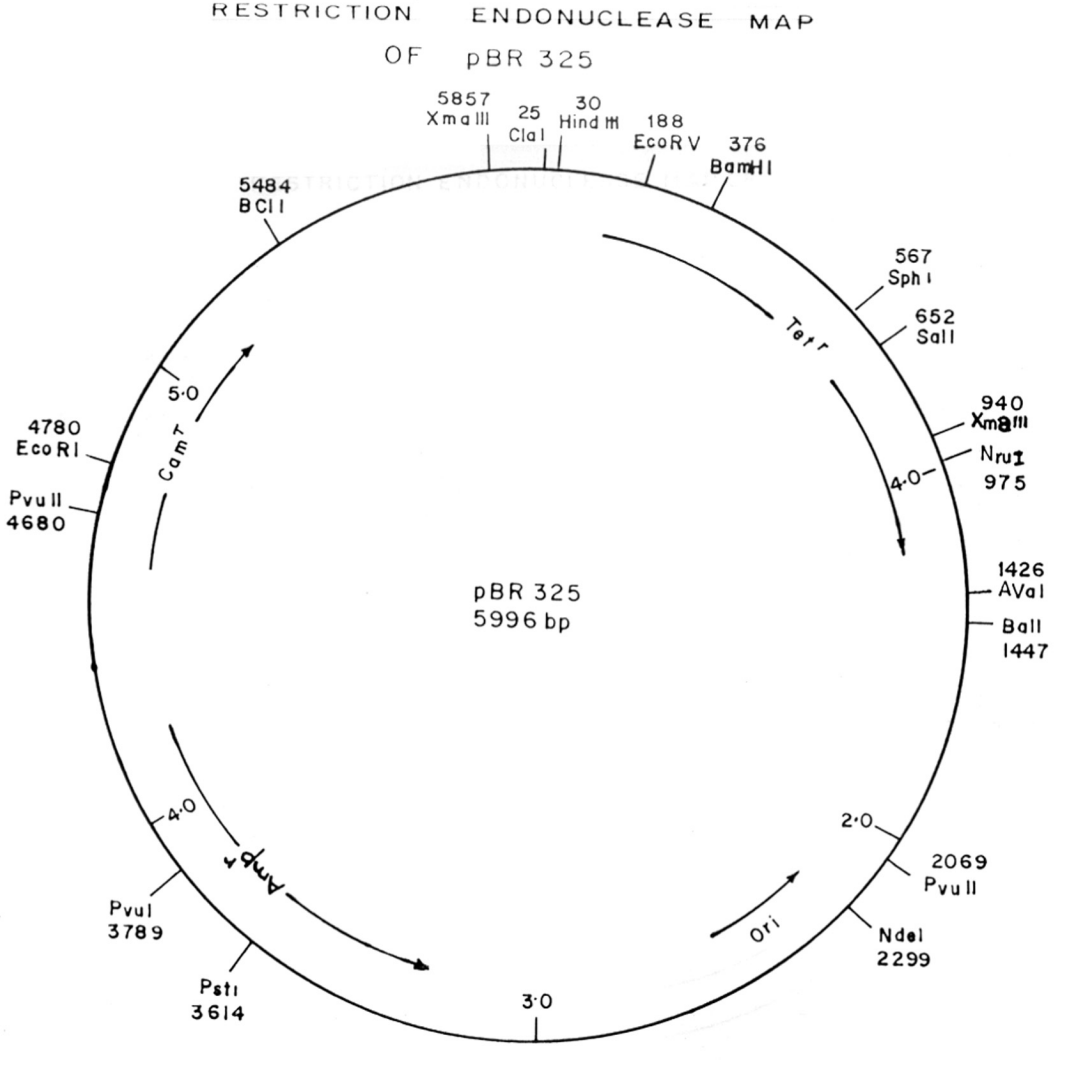

RESTRICTION ENDONUCLEASE MAP
OF pBR 325
5857
Xma III
25
ClaI
30
Hind III
188
EcoRV
376
BamHI
5484
BCII
567
Sph I
652
Sal I
940
Xma III
Nru I
975
1426
AVa I
Ball
1447
2·0
2069
Pvu II
Nde I
2299
3·0
Psti
3614
Pvu I
3789
4·0
Amp r
4780
EcoRI
Pvu II
4680
5·0
Cam r
4·0
Tet r
pBR 325
5996 bp
Ori

FIG. 4·1 A

RESTRICTION ENDONUCLEASE MAP OF M13mp18

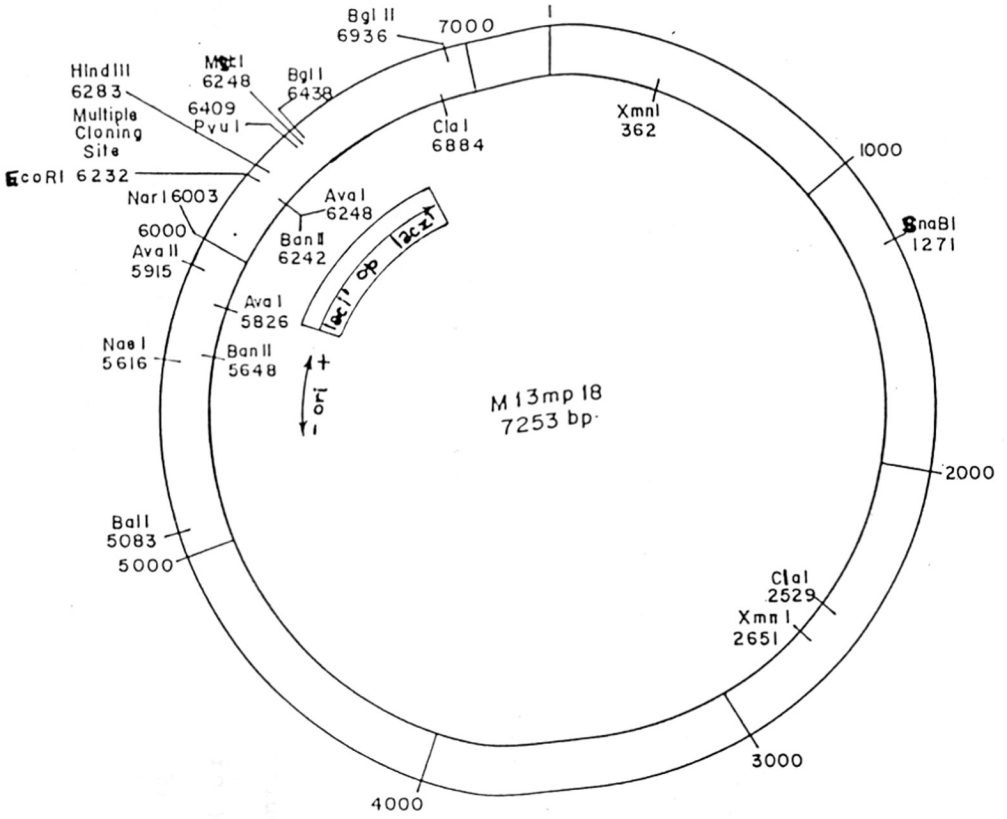


FIG. 4.1 B

2. M13mp18 (Fig. 4.1B)	Lac z gene, colourless (white) plaques in presence of X-gal and IPTG	7.25 kbp	EcoRI, SstI, KpnI, SmaI, XmaI, BamH1, XbaI, SalI, AccI, HincII, PstI, SphI, HindIII	Norrander, J. et al [12] and Yannisch-Perron, C. et al [11]

4.2.2 Prerequirements for cloning

Media : These were prepared in single distilled water and were autoclaved at 15 psi (pounds per square inch) for 20 mins.

Heat labile components like antibiotics were filter sterilized through autoclaved millipore HAWP membranes and substances like IPTG, X-gal and antibiotics were added freshly to the cooled media (45^{o}C).

For the preparation of solid media, agar was included before autoclaving and plates were poured after the mixture was cooled to 45^{o}C. The concentration of agar used was 1.5% and 2% throughout.

Composition of the medium used and the concentrations of antibiotics and other reagents used are given below :

LURIA BROTH (LB)

Tryptone	10.0 g
Yeast extract	5.0 g
Sodium chloride	10.0 g
pH	7.2-7.4
Volume	1000 ml

YT MEDIUM [14]

Tryptone	8.0 g
Yeast extract	5.0 g
Sodium chloride	5.0 g
pH	7.2-7.4
Volume	1000 ml

MINIMAL MEDIUM AGAR [14]

M9 salts	
Na_2HPO_4	6.0 g
KH_2PO_4	3.0 g
NaCl	0.5 g
NH_4Cl	1.0 g
Volume	500 ml

Adjust pH to 7.4 and autoclave. Also autoclave separately 500 ml of 3.0% agar. Cool to 55°C, combine the M9 salts and the agar.

Then add :

1 ml 1.0 M $MgSO_4$

1 ml 0.1 M $CaCl_2$

1 ml 1.0 M Thiamine-HCl

5 ml 40% glucose

PLATING AGAR

YT medium containing 6 g agar/litre was autoclaved and cooled to 45°C-55°C & then the following additions were made per 3 ml.

Plating agar : 50 ul. of 2% X-gal (5-bromo-4-chloro 3-indolylbeta-D-galactoside) in dimethylformamide (avoid plastics with concentrated dimethylformamide)

10 ul of 100 mM IPTG (isopropylbeta D-thiogalacto pyranoside).

AMPICILLIN : Ampicillin trihydrate was dissolved at a concentration of 100 mg/ml in sterile, autoclaved water. Effective concentration of 50 ug/ml was used in the medium.

TETRACYCLINE : Tetracycline hydrochloride was dissolved in sterile autoclaved water at a concentration of 20 mg/ml and was added to the medium at a final concentration of 20 ug/ml.

CHLORAMPHENICOL : It was dissolved in ethanol (3 mg/ml) and was used directly without autoclaving or filter sterilization. It was added at a final concentration of 30 ug/ml to the medium.

4.2.3 Growth, culturing and storing of E. coli DH1

The strain E. coli DH1 was grown in LB. The cells were inoculated from a solid medium (single colony on a plate) or a 15% glycerol suspension in 10 ml medium and shaken at 37°C overnight.

Cells were stored as stabs at 0-4°C or as 15% glycerol suspensions at -70°C. Stabs had to be subcultured every month, whilst glycerol cultures could be stored for atleast 6 months.

4.2.4 Growth and maintenance of E. coli JM101

To ensure that the F episome necessary for transformation present on JM101 has not been lost, JM101 should be first grown and maintained on a minimal medium.

The E. coli K12 strain JM101 was supplied as a liquid culture in 50% glycerol and was stored at -20°C. For recovery, the cells were resuspended by vortexing and then a loopful was streaked

on a minimal agar plate. The plates were incubated at 37°C, overnight and single colonies were picked for further work.

4.2.5 Steps involved in plasmid cloning

Preparation of target DNA

The long repetitive DNA sequences reassociating by Cot 50 and of size 9-20 kbp were isolated from rice cv. Basmati 370 as outlined in Section 3.2.2 and were used for cloning. The digestion and analysis of digestion mixture was carried out as stated in Section 3.2.4. In the BamH1 digest, DNA in the range 2-9 kbp was eluted by freeze-thaw method (Section 3.2.2) and was stored in 4°C till further ligation.

Preparation of vector DNA

Preparation of vector DNA involves both the digestion with a suitable enzyme and its dephosphorylation. Commercial plasmid pBR325 DNA, supplied by Bethesda Research Laboratory, was checked on 0.7% agarose gel prior to its digestion. After confirming the presence of a supercoiled band in the preparation, it (2 ug) was digested with BamH1 (3U/ug) as per the conditions described in Section 3.2.4. A complete digestion of the plasmid was confirmed by observing a single linear band of about 6.0 kbp on the gel. The DNA was deproteinised with chloroform : isoamyl alcohol mixture and aqueous layer was stored at 4°C for dephosphorylation.

In order to hydrolyze 3' and 5' phosphates from the linearised pBR325, to prevent self-ligation during ligation, treatment with bacterial alkaline phosphatase (BAP) is essential. To 100 ng of pBR325, bacterial alkaline phosphatase (70 U/ug DNA) was added in a total volume of 100 ul. The reaction was carried out at 65°C for 1 h and then was terminated by two extractions with equal volume of phenol : chloroform : isoamyl alcohol (50 : 48 : 2) mixture. This was followed by just chloroform : isoamyl alcohol mixture to remove traces of phenol. Finally, the aqueous layer was stored at 4°C tillfurther use.

Ligation of restricted plant DNA to dephosphorylated vector DNA

The BamH1 restricted rice repetitive DNA and the baptised pBR325 DNA were mixed in a reaction volume of 100 ul containing 20 ul of 5X ligation buffer (250 mM Tris-HCl, pH 7.6, 50 mM $MgCl_2$, 5.0 mM DTT, 5.0 mM dATP) and 5 U/ug of T4 DNA ligase. The reaction was carried out at 15°C for 15-16 h and then at 4°C for 48 h. After ligation, the DNA was directly used for transformation.

Preparation of competent cells and transformation of E. coli DH1

Transformation of E. coli DH1 was carried out with 40 ng and 50 ng of ligated DNA. Along with the ligated samples, uncut plasmid pBR325 DNA (25 ng) and host DH1 cells were used as positive and negative controls, respectively. The transformation procedure was essentially according to Mandel and Higa, 1970 [15] with slight

modifications. In brief, a loopful of the culture stored in -70°C was innoculated in 10 ml LB and was grown overnight at 37°C. Next 25 ml of LB in a 250 ml flask was innoculated with 0.25 ml of the overnight culture. The cells were allowed to grow with vigorous shaking at 37°C, till the absorbance at 550 nm was 0.5 (cell density of 5 x 10^7 cells/ml). Subsequently, the culture was chilled on ice for 10 min and the cell pellet was collected by centrifugation at 6000 rpm for 5 min at 4°C. The pellet was suspended in half of the original culture volume of ice cold, sterile solution of 50 mM $CaCl_2$ and 10 mM Tris-HCl, pH 8.0 to make the cells permeable for DNA uptake. The cell suspension was incubated in an ice bath for 20-25 min and then was centrifuged at 8000 rpm, 4°C for 5 min. The cell pellet was resuspended gently in 1/15th of the original culture volume in the same $CaCl_2$-Tris HCl buffer and 0.2 ml aliquots were dispensed in prechilled transformation tubes. The DNA was added to the aliquots and the reaction mixtures were incubated in ice-bath for 40 min. Later, the host cells were transformed by heat shock at 42°C for 3 min. Immediately after this step, 2 ml of luria broth was added to each tube and the tubes were incubated at 37°C for 1 h without shaking to allow the bacteria to recover and to begin to express antibiotic resistance. Again the cell pellet was collected by centrifugation and was suspended in 200 ul LB; out of which half was spread on the prepoured selective medium i.e. Ampicillin and normal plates which were incubated at 37°C for 18 h.

4.2.6 Selection of recombinants and their storage

The isolated, single colonies on selection medium (ampicillin) after transformation were picked up with sterile toothpicks and patched on normal, tetracycline, ampicillin and chloramphenicol plates in this order. Each transformant was numbered serially. Plates were incubated at 37°C for 16-18 h and recombinants were selected by their sensitivity towards tetracycline and resistance for ampicillin and chloramphenicol. They were picked and patched on ampicillin plates and incubated at 37°C, 16-18 h to prepare the master plates which were finally numbered serially. The master plates serve as a source of the desired recombinants for future studies.

The recombinant colonies were maintained for a few weeks on the surface of agar media as master plates (Section 4.2.6). The plates were tightly wrapped in parafilm and stored inverted at 4°C.

The bacterial cultures were also stored for longer period in medium containing 15% glycerol at low temperatures (-70°C) without a significant loss of viability. For this purpose, a single recombinant colony was innoculated in 5 ml LB with ampicillin and allowed to grow overnight at 37°C. Then, 0.85 ml of this overnight culture was transferred to a sterile tube containing 0.15 ml of sterile 100% glycerol. The contents were mixed thoroughly by vortexing and the glycerinated suspensions were stored in -70°C.

Viable bacteria could be removed by simply scratching the surface of the frozen stock with a sterile platinum loop or wire.

4.2.7 Isolation of recombinant plasmid DNA

Rapid isolation method [16] : The recombinants were innoculated in 1 ml LB containing ampicillin (50 ug/ml) from glycerol stocks or from master plates, (1 colony/tube) and grown at 37^{o}C, overnight. The well-grown culture was spun and the pellet was collected in an Eppendorf vial. The pellet was suspended in 20-30 ul of cracking buffer (50 mM NaOH, 0.5% SDS, 5 mM EDTA, pH 8.0 0.025% Bromophenol blue). As a result of NaOH-SDS, the cells were lysed and the DNA was released in the supernatant. The solution was then centrifuged and 20-25 ul of supernatant was loaded on 0.7% agarose gel (as mentioned in Section 3.2.4) to check the plasmid DNA.

Isolation of DNA by alkali lysis method

This method gives a much cleaner DNA preparation which can be used further for restriction mapping and analysis. It was carried out according to Birnhoim and Doly, 1979 [17] with some modifications. A single colony from the master plate or glycerol stock was innoculated in 5 ml LB containing ampicillin and grown at 37^{o}C overnight. The cell pellet was collected by centrifugation. It was suspended in 500 ul of ice cold GET solution (50 mM Glucose, 10 mM EDTA, 25 mM Tris-HCl, pH 8.0) by vortexing and was incubated for 5 min at room temperature. Subsequently, 1 ml of a freshly

prepared solution of 0.2 M NaOH and 1% SDS was added to lyse the cells and the DNA was released in the aqueous layer. The contents were mixed well by inverting the tube rapidly two or three times and then it was kept on ice for 5 min. Ice cold solution of 5M potassium acetate, pH 4.8 (750 ul) was added to the above DNA solution, the contents were mixed thoroughly and stored on ice for 5 min for protein precipitation. The aqueous layer containing the DNA was collected by centrifugation and was deproteinised successively by equal volume of tris-saturated phenol : chloroform : isoamyl alcohol and by chloroform : isoamyl alcohol. The aqueous layer was treated with DNase free RNase (20 ug/ml) at 37°C for 1 h and then with chloroform : isoamyl alcohol followed by precipitation of DNA in presence of 0.3 M sodium acetate, pH 5.3 and two volumes of chilled ethanol. The DNA was allowed to precipitate at -20°C overnight or at -70°C for 2 h. It was collected by centrifugation, dried, dissolved in 10 mM Tris-HCl, pH 7.4 and was checked by agarose gel electrophoresis as outlined in Section 3.2.4.

4.2.8 Colony hybridization and autoradiography

This is a rapid method of detecting the presence of insert (plant) DNA in the recombinant bacterial colonies [18]. The nitrocellulose filter was layered onto the master plate (shiny surface in contact with the colonies) and was pressed to ensure intimate contact with colonies. Then the filter was placed on a 3 mm What-

mann paper and soaked in the following solutions serially. Care was taken not to shake or swirl the filters.

0.5 N NaOH - once, 7 min

1 M Tris-HCl, pH 7.4 - twice, 1-4 min

1.5 M NaCl + 0.5 M Tris-HCl, pH 7.4 - 5 times, 5 min.

Following these steps, the filter was dried to whiteness at 60°C for 5-10 min to fix the released DNA on the nitrocellulose paper. It was washed quickly with chloroform to remove proteins and again dried at room temperature. Finally, the filter was soaked well in 0.3 M NaCl and baked at 80°C under vacuum.

The procedures for prehybridization, hybridization and washing of the blots were as described in Section 3.2.5.

The total rice repetitive DNA was labelled with α-^{32}P-dATP as in Section 3.2.5 and used as probe in the above experiment.

After exposure to X-ray film at -70°C for 16-18 h, the film was developed as detailed in Section 3.2.5.

4.2.9 Characterization of recombinant DNAs

The recombinant plasmid DNAs were isolated as in Section 4.2.7. The size of the insert DNAs was determined by digesting the plasmids with BamH1 and analysing the digests by agarose gel electrophoresis (Section 3.2.4).

The insert DNAs, after BamH1 digestion, were isolated by elution (Section 3.2.2) and were characterised by subjecting to single or double digestion with specific restriction enzymes [19].

4.2.10 Steps involved in M13 cloning and sequencing

All the steps involved in these procedures were according to the New England Biolabs' Laboratory Manual.

Preparation of target DNA : The pRL7 plasmid DNA (1 ug) was double digested with BamH1 and PstI and the digestion was checked as described in Section 3.2.4. After ensuring complete digestion, the digest was once extracted with chloroform : IAA mixture and the aqueous layer was stored at 4°C till it was used for ligation.

Preparation of vector DNA : Commercially available M13 vector DNA /i.e. M13mp18 supplied by New England Biolabs, was checked on a 0.7% agarose gel prior to digestion. 50 ng of the vector DNA was double digested with about 3 U/ug BamH1 and PstI according to Maniatis et al [19]. After ensuring complete digestion i.e. the presence of a single band of about 7.0 kbp on the gel, the digested DNA was deproteinised with chloroform : IAA mixture and the aqueous layer was used directly for ligation.

Ligation of restricted target DNA to vector DNA : The restricted plasmid DNA, carrying the rice repetitive DNA (pRL7) was ligated to the linearised M13mp18 with slight modifications. The target

DNA was present at about 5-fold higher molar concentration than the vector DNA. The reaction was carried out at 16°C for 2 h and the ligated DNA was stored at / -20°C till further use.

Preparation of competent cells and transformation of E. coli JM101

Competent cells of E. coli JM101 were prepared and were transformed with the ligated DNA. For transformation, along with the ligated DNA, positive (10 ng M13mp18 RF DNA) and negative (only the host JM101 cells in $CaCl_2$) /were also included. The controls recombinant (white) plaques were purified before they were used for the isolation of template DNA.

4.2.11 Direct gel electrophoresis (DIGE)

This involves sizing of insert DNA which is obtained by direct lysis of cells in vivo.

For this purpose a single colourless plaque was picked by a capillary and innoculated in 2 ml YT medium alongwith 10 ul of a fresh JM101 plating culture. The tube was incubated at 37°C with shaking for about 6-8 h. To 20 ul of this culture, 1 ul of 2% SDS, 2 ul of 0.1 M EDTA and 10 ul of loading buffer (50% glycerol, 0.05 M EDTA, pH 8.0, 0.2% bromophenol blue) were added, the mixture was well shaken and it was allowed to stand at room temperature for about 30 min. Then, the suspension was loaded on a 0.5% agarose gel and electrophoresis was carried out at 8-

10 mA for about 12-14 h using M13mp18 single stranded template DNA as control.

4.2.12 Preparation of template DNA

For the sequencing reaction, it is necessary to isolate the insert DNA from the recombinant plaques in a single stranded form. This is referred to as template DNA and it was isolated according to the method of Schreier and Cortese [21] with some modifications.

1 ml of the 6-8 h grown culture (obtained in Section 4.2.11) of the phage was microfuged and the supernatant containing the phage was collected into another eppendorf tube. To this phage solution, 200 ul of 2.5 M NaCl-20% PEG (6000) [22] was added, the mixture was well shaken and was put on ice for 30 min for phage precipitation. The phage precipitate was collected by centri fugation, dried and dissolved in 100 ul of 10 mM Tris-HCl (pH 7.5)-0.1 mM EDTA. The aqueous layer was extracted once with equal volume of tris-saturated phenol, pH 7.8 and once with chloroform : IAA mixture. Finally, the DNA was precipitated in presence of 10 ul 3.0 M Sodium acetate (pH 6.0) and 200 ul of chilled ethanol. After keeping at -20^{o}C overnight, the precipitate was collected, dried and dissolved in 10 mM Tris-HCl (pH 7.5), 0.1 mM EDTA. The DNA concentration was checked by 0.7% agarose gel electrophoresis using M13mp18 single stranded template DNA as standard.

4.2.13 Dideoxynucleotide chain termination sequencing method

The cloned rice repetitive DNA, isolated from the purified recombinant plaques, was sequenced essentially according to Sanger et al [23].

4.2.14 Casting and electrophoresis of sequencing gels

Cleaning of glass plates

Before pouring the gel, the glass plates were thoroughly cleaned sequentially with tap water, distilled water and 70% ethanol. Then, only the notched plate was siliconised using about 0.5 ml of 2% silane (dichlorodimethyl silane) to ensure that the gel did not stick to it after electrophoresis.

Preparation of denaturing 8% polyacrylamide gels and electrophoresis

The spacers of 0.6 mm thickness were placed on the clean, siliconised notched plates. The smooth plate was aligned on it properly and the bottom and the sides (about 3-4 inches) of both the plates were taped and clamped properly.

The 8% polyacrylamide gel mix was prepared by mixing 20 ml of 40% acrylamide solution (38 g acrylamide and 2 g bis-acrylamide in 100 ml volume), 10 ml 10X TBE buffer (108 g, tris, 55 g boric acid, 9.3 g EDTA/lit.) and 42 g urea. The solution was heated slightly to dissolve urea and then the volume was made upto 100 ml. The solution was filtered and deaerated for 5 min and was

stored at 4°C in a dark-coloured bottle till further use. Before pouring the gel, 80 ul TEMED N,N,N',N',-tetramethylethylene diamine) and 800 ul of 10% ammonium persulphate were added to 100 ml of 8% polyacrylamide gel mix. After pouring, the comb was immediately inserted and the gel was allowed to set for atleast 30 min. Then, the tapes were removed and the gel was clamped to the electrophoresis unit. The unit and the glass plates were maintained at a constant temperature of about 45-50°C using a circulating water bath. After preelectrophoresis at 800 V for 30 min, the samples were loaded and electrophoresed at 2000 V till the bromophenol blue (BPB) ran out. Immediately the sample was loaded again and electrophoresis was continued till BPB just touched the lower edge of the gel. This differential migration helped to read the sequence as proper base positions. 0.5 X TBE buffer was used as a running buffer during electrophoresis.

4.2.15 Autoradiography and sequence analysis

After electrophoresis, the plates were dismantled in such a way that the gel remained evenly stuck to the smooth plate. The gel with the glass plate was wrapped in a Saran wrap and was exposed to X-ray film for about 16 h at -70°C. The film was then developed as described in Section 3.2.5.

Restriction enzyme site analysis, search for open reading frame and dot matrix analysis (comparison) of the BamH1-PstI (pRL7-0.4 kbp) sequence were carried out using SEQAID II program of D. Jones which was a kind gift by Prof. Muthukrishnan, Kansas State University, Kansas, USA.

4.3 RESULTS AND DISCUSSION

4.3.1 Cloning of rice long repetitive DNA in pBR325

The S1 nuclease resistant long repetitive DNA sequences (9-20 kbp) reassociating by Cot 50 were used for cloning in pBR325. To check the digestability of the DNA, it was treated with a few restriction enzymes such as EcoRI, SalI, PvuI, PstI and BamH1 which have unique cloning sites in pBR325 (Fig. 4.2). Out of these enzymes, only BamH1 showed the desired digestion in the range of 2-9 kbp and hence was used during cloning. The region of 2-9 kbp in the BamH1 digest was eluted out, ligated to BamH1 digested and dephosphorylated pBR325 DNA; (Fig. 4.3) and the ligated DNA was used to transform E. coli DH1. The transformed colonies were selected by their sensitivity towards tetracycline and resistance towards chloramphenicol and ampicillin (Fig. 4.4) ($Cam^r Amp^r Tet^s$). The presence of rice repetitive sequence in these clones was confirmed by in situ colony hybridization using α-^{32}P-labelled rice long repetitive DNA (Fig. 4.5). Out of several recombinants,

FIGURE 4.2 : DIGESTION OF LONG REPEATS OF COT 50 WITH DIFFERENT RESTRICTION ENZYMES

Control long repeats (lane a)
λ HindIII digest (lanesb & g)
BamH1 digest (lane c)
EcoRI digest (lane d)
SalI digest (lane e)
HindIII digest (lane f)
PvuI digest (lane h)
PstI digest (lane i)
KpnI digest (lane j)

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

FIGURE 4.3 : DIGESTION OF VECTOR (pBR325) DNA WITH BamH1

λ HindIII digest (lane a)
pBR325 DNA uncut (lane b)
pBR325-BamH1 digest (lane c)

Electrophoresis was carried out on 0.7% neutral agarose slab gel in TAE buffer (pH 8.1) at a constant current of 30 mA.

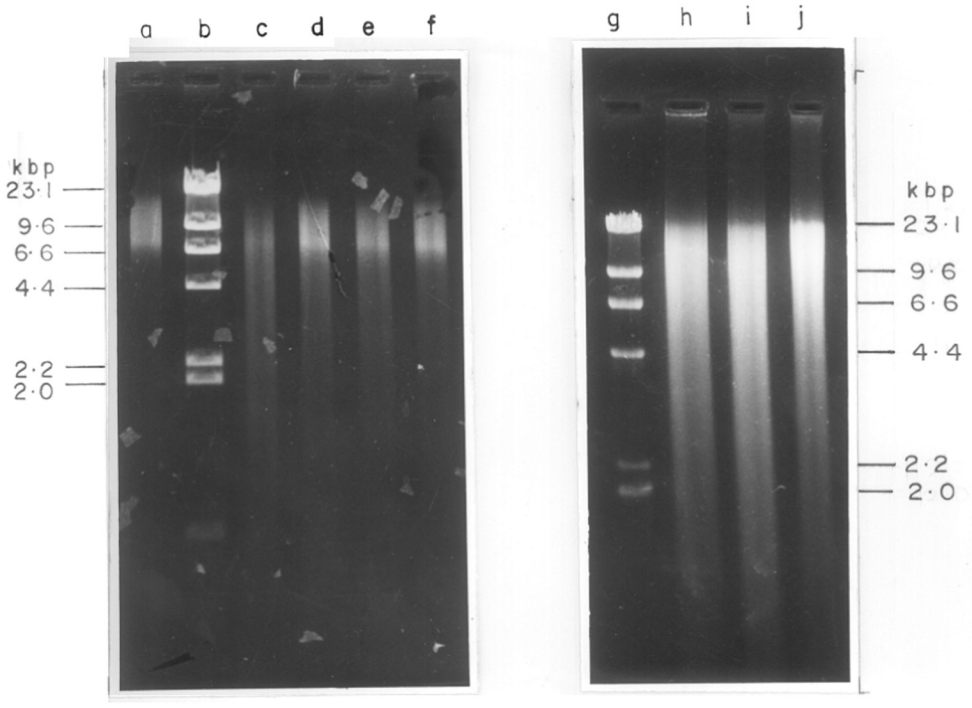

a b c d e f
g h i j
kbp
23·1
9·6
6·6
4·4
2·2
2·0
kbp
23·1
9·6
6·6
4·4
2·2
2·0

FIG. 4·2

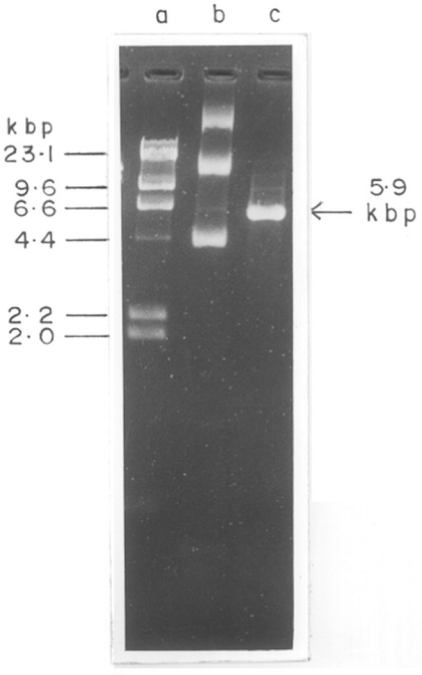

a b c
kbp
23·1
9·6
6·6
4·4
2·2
2·0
5·9
kbp

FIG. 4·3

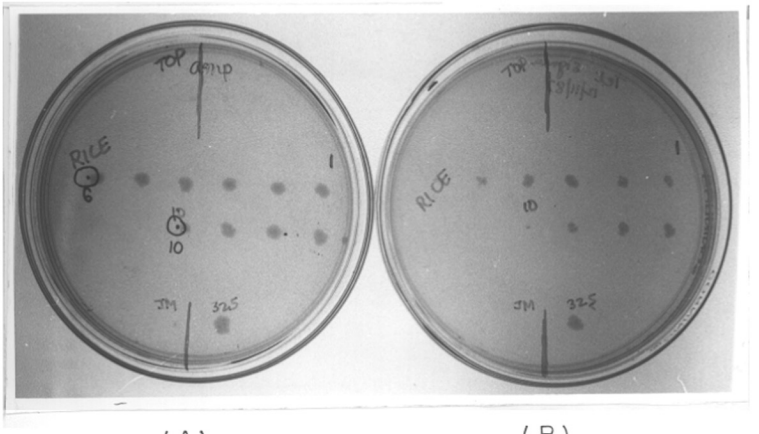
FIGURE 4.4 : SELECTION OF RECOMBINANTS ON AGAR PLATES USING ANTIBIOTIC MARKERS

Luria agar with ampicillin (Plate A)
Luria agar with tetracycline (Plate B)

Host culture DH1 and pBR325 were also patched as negative and positive controls respectively. The circled colonies on ampicillin plates were picked up as recombinants.

FIGURE 4.5 : COLONY HYBRIDIZATION OF THE pRL RECOMBINANTS (GROWN ON A MASTER PLATE)WITH RICE REPETITIVE DNA AS PROBE.

After denaturation, pBR325 DNA was also spotted on the nitrocellulose filter as a negative control.



(A) (B)

FIG. 4·4

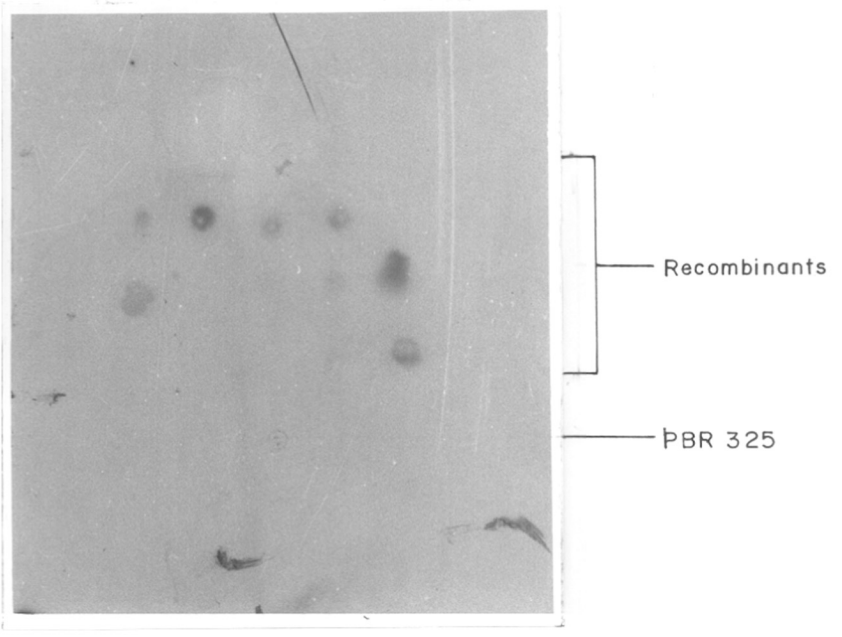


FIG. 4·5

only a few were selected randomly for further characterization.

4.3.2 Size of insert DNAs and their restriction patterns

In order to determine the insert size, plasmids carrying the repetitive sequences were isolated (Fig. 4.6), digested with BamH1 and electrophoresed on a 0.7% agarose gel. As is evident from Fig. 4.7, most of the cloned DNA sequences are of size 3-4 kbp. The insert size, for example, is 3.4 kbp in pRL1, 3.2 kbp in pRL10, 3.6 kbp in pRL2 and pRL7, 4.0 kbp in pRL5 and 3.7 kbp in pRL8.

To assess the presence of sites for a few other restriction enzymes in the inserts, the plasmids were digested with AvaI, HindIII, PvuII and PstI which have either single or double cutting sites in pBR325 (Fig. 4.8). The differences in their digestion patterns with all these enzymes indicate the occurrence of different restriction sites in the cloned DNA sequences.

4.3.3 Presence of pRL7 repeat sequences in rice genome

In order to check the presence of a specific clone repeat sequence in rice genome and to show that it has not undergone any change during the cloning steps, the 3.6 kbp repeat sequence (pRL7) was hybridised to rice total DNA digested with BamH1. The hybridization pattern shows a band of 3.6 kbp and three other bands of high molecular weight which are not the multiples of 3.6 kbp (Fig. 4.9). This indicates that the 3.6 kbp repeat DNA fragment is

FIGURE 4.6 : ELECTROPHORESIS PATTERN OF MINIPREPARATIONS OF pRL1-pRL10 DNAs WITH pBR325 DNA AS CONTROL

λ HindIII digest (lane a)
pBR325 DNA (lane b)
pRL1-pRL10 DNAs (lanes c-l)

Electrophoresis was carried out on 0.7% neutral agarose slab gels in TAE buffer (pH 8.1) at a constant current of 30 mA.

FIGURE 4.7 : DIGESTION OF pRL1-pRL10 DNAs WITH BamH1

λ HindIII digest (lane a)

pBR325 digested with BamH1 (lane b)
pRL1-pRL10 digested with BamH1 (lanes c-l)

The arrows indicate the molecular weights of the linearised pBR325 and the insert DNAs.

Electrophoresis was carried out on 0.7% neutral agarose slab gels in TAE buffer (pH 8.1) at a constant current of 30 mA.

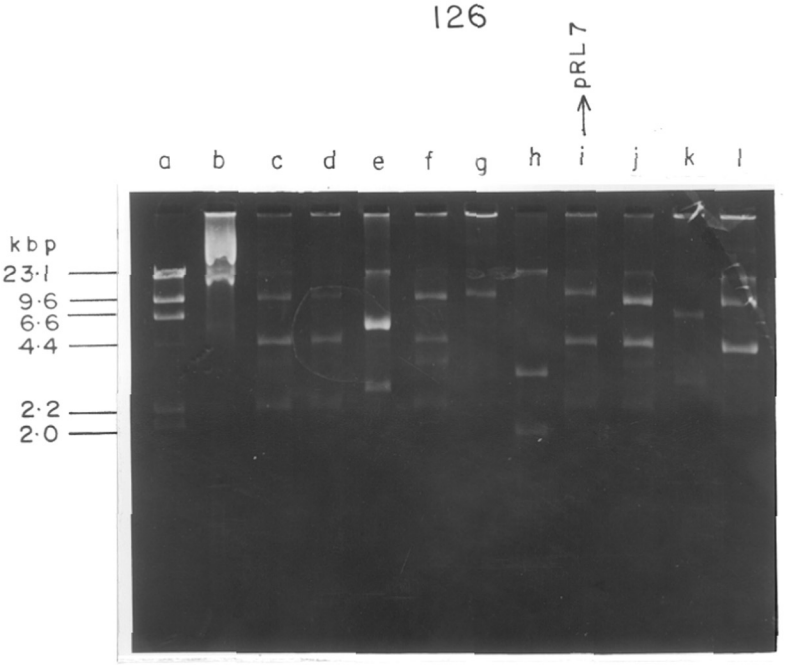
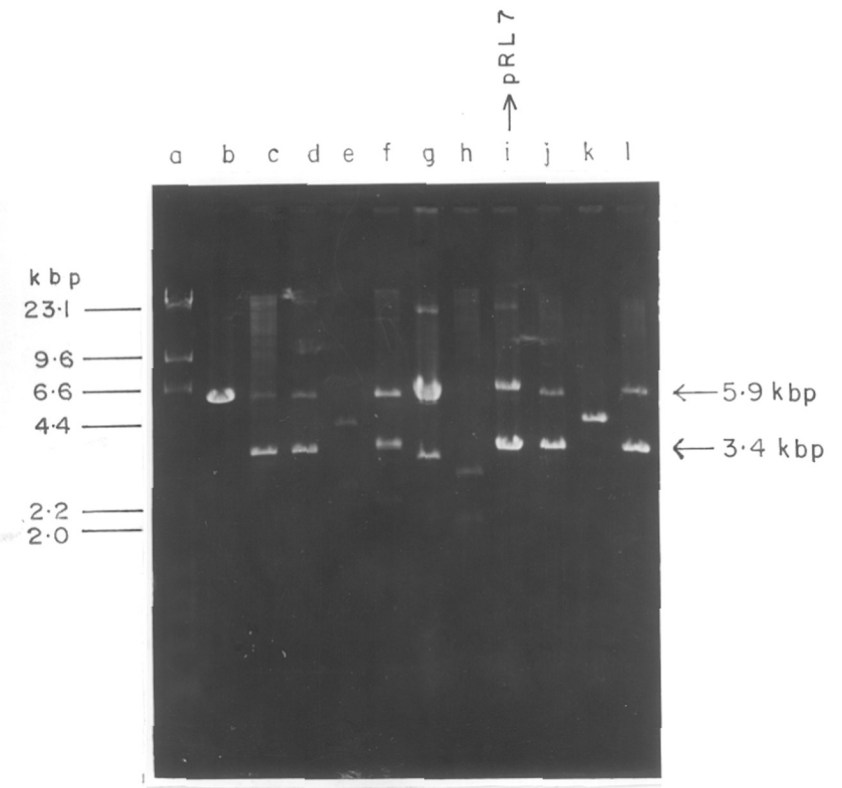


FIG. 4·6



BamHI digestion

FIG. 4·7

FIGURE 4.8 : DIGESTION OF pRL1-pRL10 DNAs WITH

(A) PvuII, (B) HindIII, (C) AvaI, (D) PstI

λ HindIII digest (lane a)
ØX174HaeIII digest
pBR325 (lane b)
pRL1-pRL10 (lanes c-l)

Electrophoresis was carried out on 0.7% neutral agarose slab gels in TAE buffer (pH 8.1) at a constant current of 30 mA.

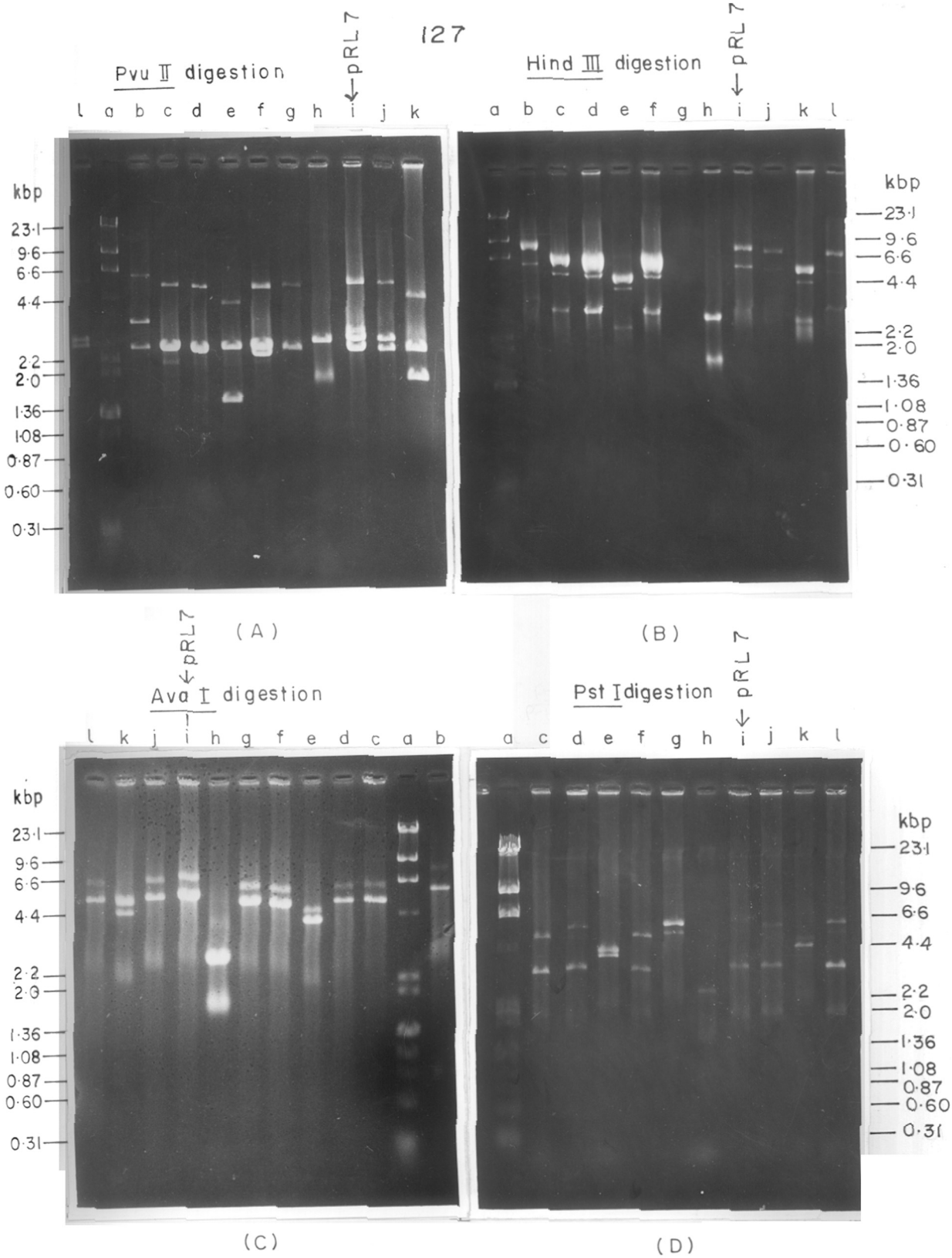

Pvu II digestion
← pRL 7
l a b c d e f g h i j k
kbp
23·1
9·6
6·6
4·4
2·2
2·0
1·36
1·08
0·87
0·60
0·31
(A)
Hind III digestion
← pRL 7
a b c d e f g h i j k l
kbp
23·1
9·6
6·6
4·4
2·2
2·0
1·36
1·08
0·87
0·60
0·31
(B)
Ava I digestion
← pRL 7
l k j i h g f e d c a b
kbp
23·1
9·6
6·6
4·4
2·2
2·0
1·36
1·08
0·87
0·60
0·31
(C)
Pst I digestion
← pRL 7
a c d e f g h i j k l
kbp
23·1
9·6
6·6
4·4
2·2
2·0
1·36
1·08
0·87
0·60
0·31
(D)

FIG. 4·8

present independently inthe genome and is also a part of some sequences which are of higher molecular weights.

4.3.4 Localization of a few restriction enzyme sites in pRL1, pRL7, pRL10

For locating a few restriction enzyme sites in the cloned sequences, three recombinants i.e. pRL1, pRL7 and pRL10 were selected. These plasmids were digested with BamH1 and the corresponding inserts were eluted and used for further characterization using a few enzymes like PvuII and EcoRI in single and double digestions. Since the strategy followed for the various restriction enzyme digestions was the same for all the three plasmids, a detailed analysis only of pRL7 is described. Figure 4.7 (lane i) shows the digestion of total pRL7 plasmid DNA with BamH1. Here, two intense bands of 6.0 kbp and 3.6 kbp are observed. Since the 6.0 kbp band corresponds to the molecular weight of pBR325 (linearised), 3.6 kbp band indicates the size of rice repetitive DNA cloned in this plasmid. On digestion of pRL7 plasmid DNA with PvuII, three bands of 3.6 kbp, 3.4 kbp and 2.6 kbp are observed (lane d) in Fig. 4.10 A. The 2.6 kbp band is due to the occurrence of two PvuII sites at a distance of 2.6 kbp in pBR325. Besides, this the two bands at positions 3.6 kbp reveal clearly a PvuII site in the cloned DNA cutting it into two fragments of molecular weights 1.7 kbp and 1.9 kbp. This is confirmed when the eluted insert DNA is digested with PvuII and two bands of 1.7 kbp and

FIGURE 4.9 : AUTORADIOGRAMS OF pRL7 INSERT DNA HYBRIDIZED TO RICE SHOOT DNA DIGESTED WITH BamH1

Control rice DNA (lane a)
BamH1 digest (lane b)

FIGURE 4.10 : DIGESTION OF PLASMID pRL7 WITH BamH1, PvuII and EcoRI

(A) λ HindIII digest + ϕX174HaeIII digest (lane a)
BamH1 digest (lane b)
BamH1-PvuII digest (lane c)
PvuII digest (lane d)
PvuII-EcoR1 (lane e)

(B) Digestion of pRL7 insert DNA with PvuII

λ HindIII digest (lane a)
PvuII digest (lane b)
Control pRL7 insert DNA (lane c)
ϕX174HaeIII digest (lane d)

Electrophoresis was carried out on 0.7% (A) and 1.4% (B) neutral agarose slab gels in TAE buffer (pH 8.1) at a constant current of 30 mA.

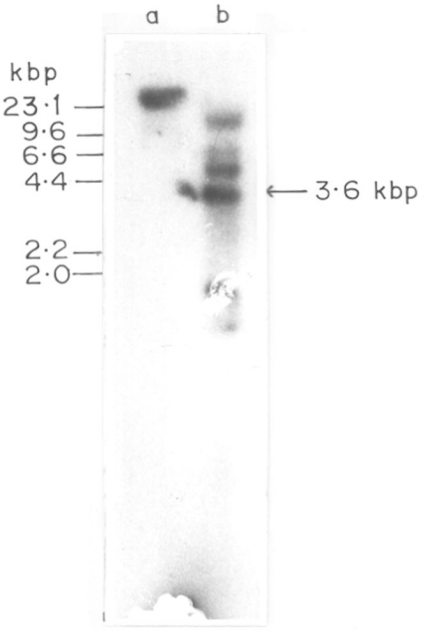


FIG. 4·9

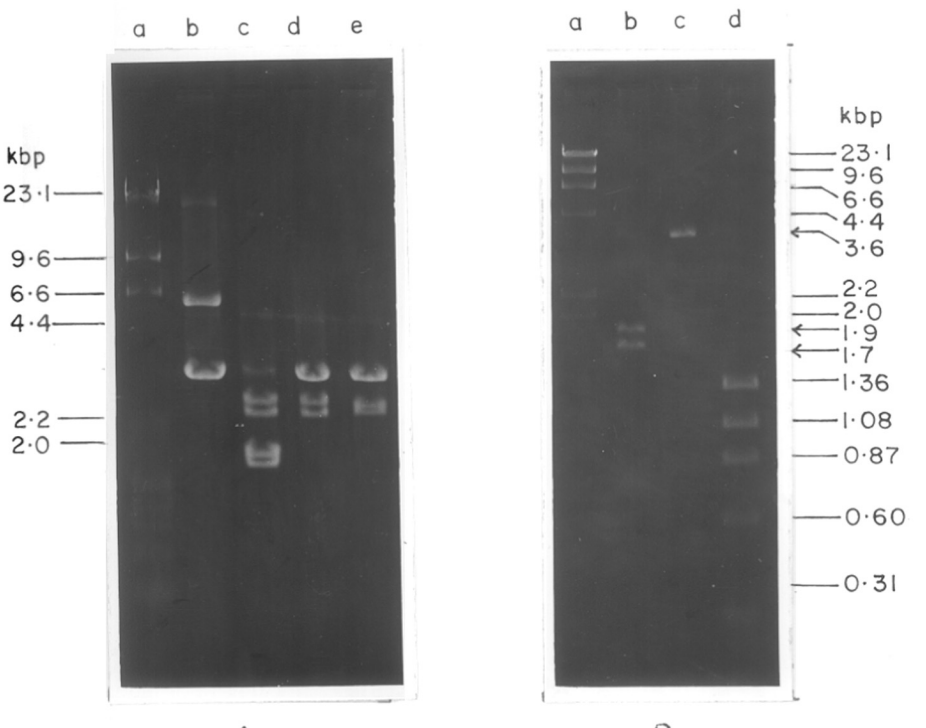


FIG. 4·10

and 1.9 kbp (Fig. 4.10 B) are observed. In order to localise these positions, the pRL7 DNA was double digested with PvuII/EcoRI giving bands of molecular weights 3.6 kbp, 3.3 kbp and 2.6 kbp (Fig. 4.10 A, lane e). Since, the 3.4 kbp band in the PvuII digest of pRL7 (lane d, Fig. 4.10 A) is reduced to 3.3 kbp in PvuII/EcoRI double digest of pRL7 (lane e, Fig. 4.10 A); and there is no change in the molecular weight of the other two bands, it shows the presence of 1.7 kbp fragment in the insert DNA towards the EcoRI site in pBR325 followed by 1.9 kbp PvuII-BamH1 fragment. The difference of 0.1 kbp in the size of the band is due to the difference of 0.1 kbp in PvuII and EcoRI sites in pBR325 (Fig. 4.1 A).

Figure 4.11 A (lane b and c) shows the digestion of pRL7 insert DNA with EcoRI and PstI where single bands of 3.6 kbp and 3.2 kbp respectively are observed. These data reveal an absence of EcoRI site and presence of a single PstI site at a distance of 0.4 kbp from the BamH1 site in the insert DNA. To determine at which end the PstI site is present in pRL7 insert DNA, double digestion of pRL7 with BamH1/PstI was carried out. Fig. 4.11 B shows clearly three bands of about 6.0 kbp, 3.2 kbp and 0.4 kbp. The presence of a band of 6.0 kbp reflects the PstI site at a distance of 3.6 kbp from the position no. 1 in pBR325. Digestion of the cloned DNA in pRL7 with BclI again shows a single band of 3.6 kbp indicating an absence of BclI site (Fig. 4.12).

Finally, enzyme digestion of pRL7 insert DNA with AluI was analyzed on 10% native, polyacrylamide gel. AluI digest shows

FIGURE 4.11 DIGESTION OF pRL7 INSERT DNA WITH EcoRI and PstI

(A) λ HindIII digest (lane a)
EcoRI digest (lane b)
PstI digest (lane c)
ϕX174HaeIII digest (lane d)

DIGESTION OF PLASMID pRL7 DNA WITH BamH1-PstI

(B) λ HindIII digest + ϕX174HaeIII digest (lane a)
BamH1-PstI digest (lane b)

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.

FIGURE 4.12 : DIGESTION OF THE INSERT DNA IN pRL7 WITH BclI

λ HindIII digest (lane a)
BclI digest (lane b)

Electrophoresis was carried out on 1.4% neutral agarose slab gels in TAE buffer (pH 8.1) at a constant current of 30 mA.

FIGURE 4.13 : DIGESTION OF INSERT DNA IN pRL7 WITH AluI

AluI digest (lane a)
ϕX174HaeIII digest (lane b)

Electrophoresis was carried out on 10% polyacrylamide gels in TBE buffer (pH 8.0) at a constant current of 20 mA.

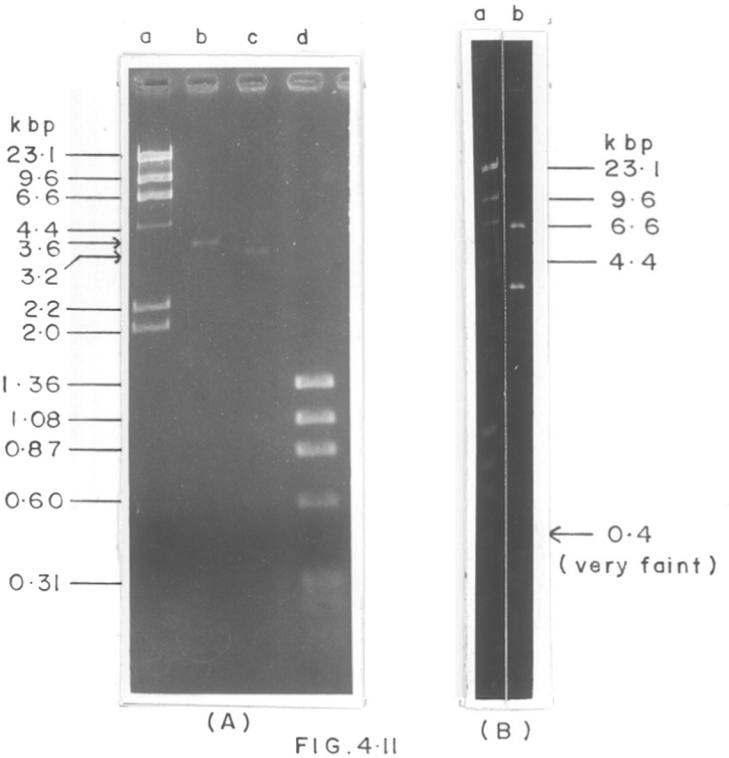


FIG. 4·11

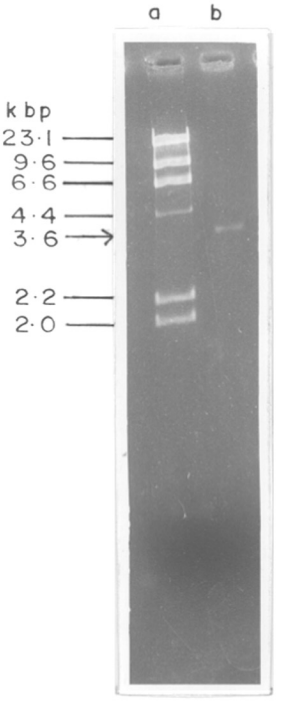


FIG. 4-12

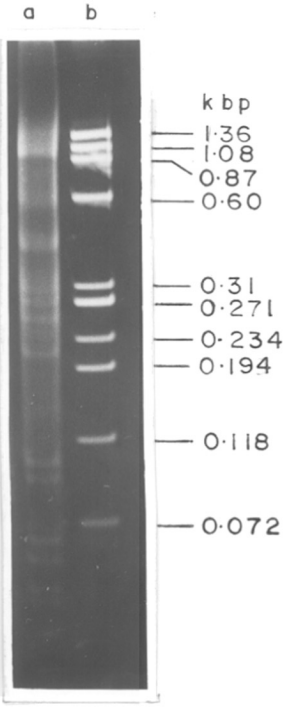


FIG. 4·13

a number of bands in the range of 1.35 kbp to 0.072 kbp indicating a number of AluI sites in the cloned DNA sequence (Fig. 4.13).

Based on these studies, it can be concluded that the rice repetitive DNAs cloned in pRL1, pRL7 and pRL10 contain no sites for EcoRI and BclI; single sites for PvuII and PstI and many sites for AluI (Fig. 4.14).

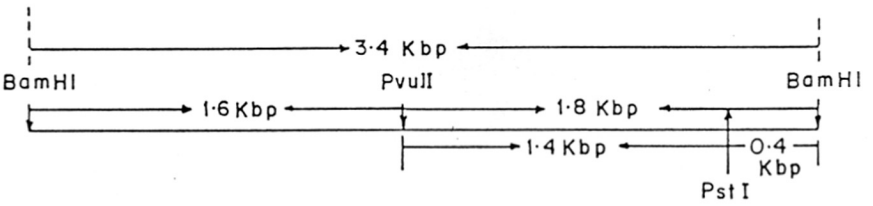
4.3.5 Subcloning of rice long repetitive DNA in M13mp18

After localising a few restriction enzyme sites in the pRL1, pRL7 and pRL10 insert DNAs, the next step was to determine the nucleotide sequence of these repeats. The 0.4 kbp fragment obtained after digestion of pRL7 insert DNA with BamH1 and PstI was selected for this purpose. This fragment was also observed to be present in pRL1 and pRL10 insert DNAs.

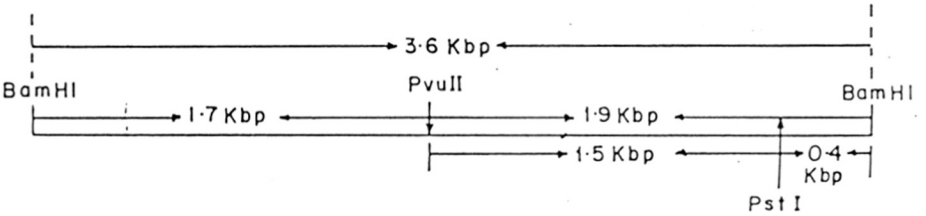
In order to sequence the 0.4 kbp fragment, the pRL7 plasmid DNA was double digested with BamH1 and PstI and was ligated to the vector M13mp18 digested with the same enzymes (Fig. 4.15). The ligated DNA was then used to transform E. coli JM101 and recombinants (white plaques) were selected by direct visualization (Fig. 4.16). A few white plaques which represented the recombinants were used for rapid extraction of DNA and the presence of about 0.4 kbp insert in these plaques was checked by direct gel electrophoresis (Fig. 4.17) (DIGE). The plaques having about 0.4 kbp insert were used for sequencing work.

FIGURE 4.14 : LINE DIAGRAM SHOWING THE RESTRICTION ENZYME SITES IN THE CLONED SEQUENCES IN pRL1, pRL7 and pRL10 DNAs.

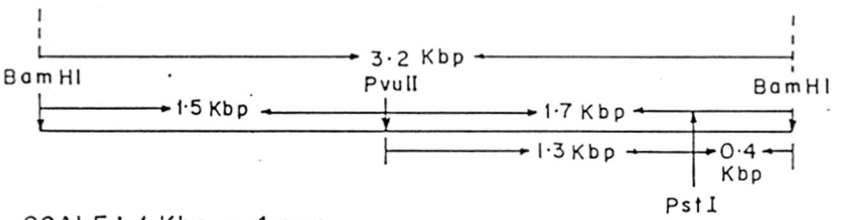
pRL 1



pRL 7



pRL 10



SCALE : 1 Kbp = 4 cms.

FIG. 4·14.

FIGURE 4.15 : DOUBLE DIGESTION OF VECTOR (M13mp18) WITH BamH1/ PstI

λ HindIII digest (lane a)
M13mp18 uncut (lane b)
M13mp18-BamH1 digest (lane c)

Electrophoresis was carried out on 0.7% neutral agarose slab gel in TAE buffer (pH 8.1) at a constant current of 30 mA.

FIGURE 4.16 : SELECTION OF M13 RECOMBINANT PLAQUES ON AGAR PLATES USING X-GAL AND IPTG

YT agar plate showing the negative control (Plate A)
YT agar plates showing the positive control (Plate B)
YT agar plates showing the recombinants (Plate C)

Host culture E. coli DH1 was used as a negative control
M13mp18 DNA (10 ng) was used as a positive control
The circled plaques in Plate C were picked up as recombinants.

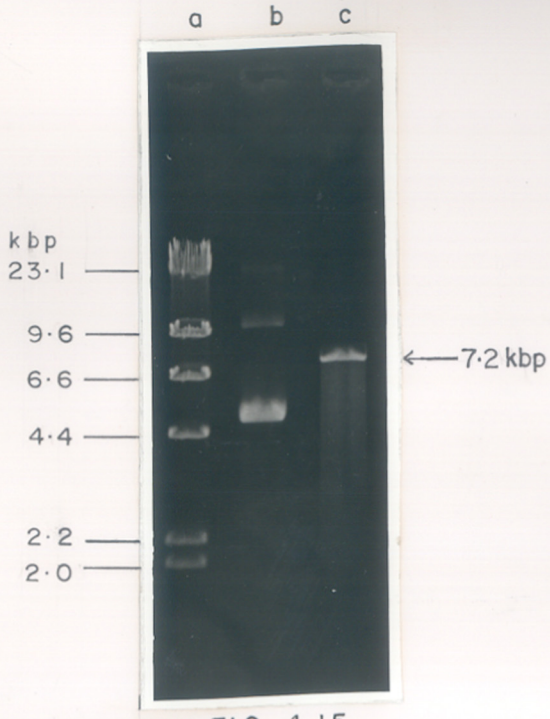
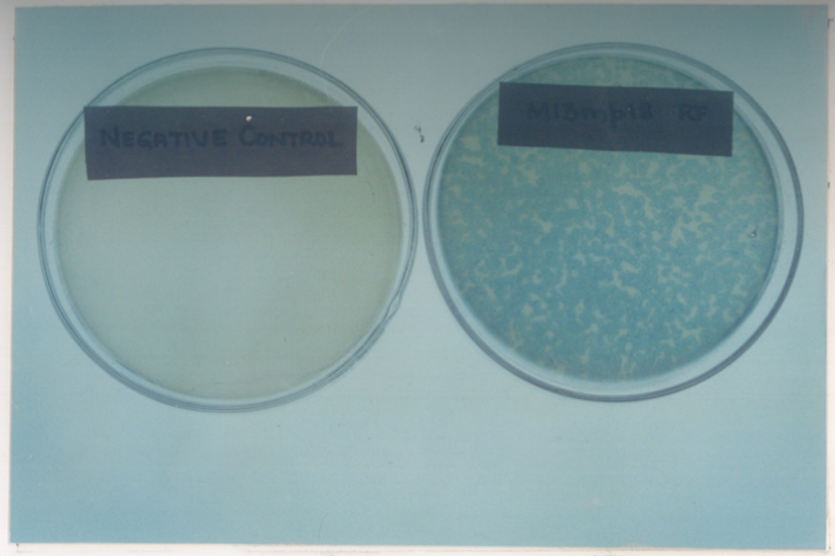
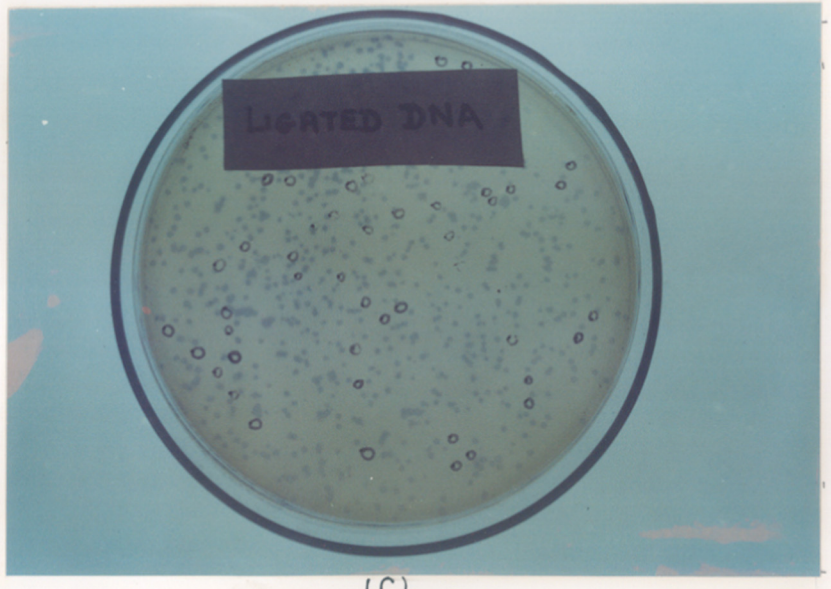


FIG. 4·15



(A) (B)



(C)

FIG. 4·16

4.3.6 Sequence analysis of the 0.4 kbp repeat element (pRL7-0.4 kbp) using dideoxy chain termination method

The 0.4 kbp BamH1-PstI fragment of pRL7 was sequenced by Sanger's dideoxy nucleotide chain termination method as described earlier in this chapter. The ss DNAs were isolated from the recombinants and were checked for their concentrations on 0.7% agarose gel (Fig. 4.18). About 0.5-1.0 ug of DNA was used for the sequencing reaction. The partial sequence obtained from 5'---> 3' end is

5' GCAGGCATCGCTGACGACCGAATCGACACCACAAGCCTGC
GAGGACAGCAATACCATACCGAAGTAAGCTCGAGACTACA
AGTCGGTGCTGAGTCGCCTGTGTTACTAGCGTACACACGT
TCGCTAGTAGAGCAGAGCTAGCTACGTCTCTCATCAGTCA
CATCAGATCTATACGTCTCTGTCAGTGTGGTGCTCAGACT
GCTTAGTCGTTAACTGACACGTACCAGGCGCGATGCGCAC
CCTCTTACATTCGCTCCCATCACATTACATACGATCCATG
CTCTCGAAAGCTACCGATCTCCTCGTATCGCACTCTGCTG
TCTAGCTTCTTCTCTTGTTCGACTGCATGCCTAGTGCGAC
TAGCGCTATCCTCTTCTCTCTCTCTGTTCTCTGTCTCTCG
AGGTACCGACTG 3'

Restriction enzyme analysis, search for open reading frame and dot matrix comparison of the pRL7-0.4 kbp sequence was carried out by computer analysis using program SEQAID II.

The restriction enzyme site analysis shows the presence of single sites for some enzymes like Fnu DII, EcoRII, ScrFI and

FIGURE 4.17 ; DIRECT GEL ELECTROPHORESIS SHOWING THE INSERT SIZES OF THE DNAs FROM THE RECOMBINANT PLAQUES

ss DNAs from recombinant plaques (lanes a-e)
M13mp18 ss DNA (lane f)
λ HindIII digest (lane g)

Electrophoresis was carried out 0.5% neutral agarose slab gel in TAE buffer (pH 8.1) at a constant current of 8-10 mA. Arrows indicate the insert sizes in the recombinants.

FIGURE 4.18 : ELECTROPHORESIS PATTERNS OF THE SINGLE STRANDED TEMPLATE DNAs USED FOR SEQUENCING

λ HindIII digest (lane a)
M13mp18 ss DNA (lane b)
ss DNAs from recombinant plaques (lanes c-g)

Electrophoresis was carried out on 0.7% neutral agarose slab gel in TAE buffer (pH 8.1) at a constant current of 20 mA.

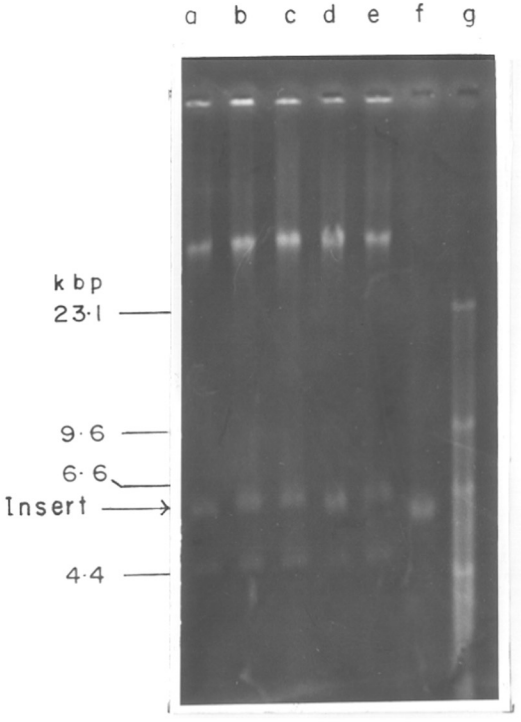


FIG. 4·17

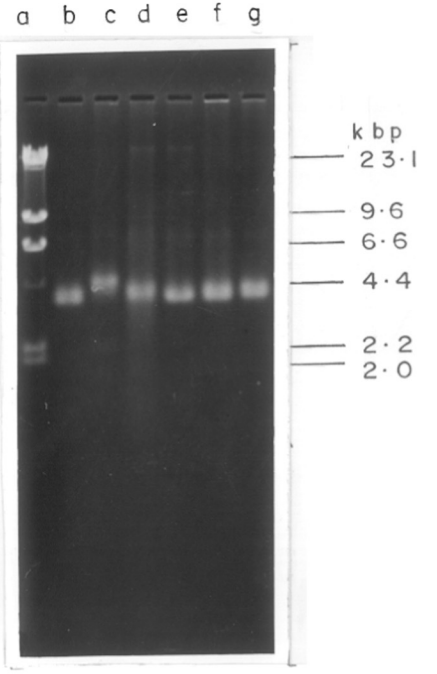


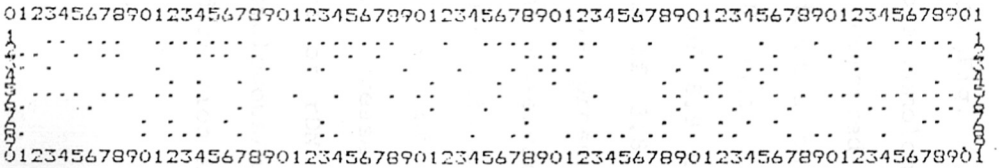
FIG. 4·18

many sites for enzymes like AluI, TaqI, MaeI and MnlI. In case of AluI, it is interesting to note that most of the fragment sizes indicated by computer analysis agree well with the molecular weights of a few bands observed in pRL7 insert DNA digested with AluI (Fig. 4.13 lane a) suggesting thereby that bands of 0.068 kbp, 0.072 kbp, 0.088 kbp and 0.150 kbp in AluI digest (Fig. 4.13) are a part of this BamH1-PstI fragment (pR 7-0.4 kbp). Further analysis shows that this sequence has a G + C content of about 53%. There are clearly three open reading frames at positions 233, 278 and 347 in the sequence.

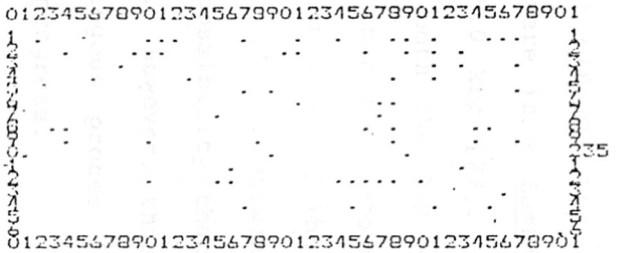
Finally for comparing the pRL7-0.4 kbp sequence with the rice 25S,(A)/17S(C)/rDNA and the spacer fragment between 25S and 17S rDNAs(B)/[24-25], dot matrix was generated using the pRL7-0.4 kbp sequence on Y-axis (Y sequence) and the known sequences on X-axis (X sequence). To reduce the background noise, it was found useful not to compare small fragments but 48 and 26 bases at a time. This eliminated most of the random fortuitous similarity while preserving the features of true homology. In the construction of dot matrix figure (Fig. 4.19), when the segment for the above length from both the sequences is equal to a minimal match of 20 or 13 bases, a dot is placed in the matrix at the coordinates corresponding to the location of a/ nucleotide in the sequence. The horizontal continuous dots show the homology between two different sequences. Continuous lines parallel to a fixed angle, from top to bottom, suggest homology over a long stretch of sequences,

FIGURE 4.19 : DOT MATRIX COMPARISON OF THE RICE REPETITIVE DNA SEQUENCE.

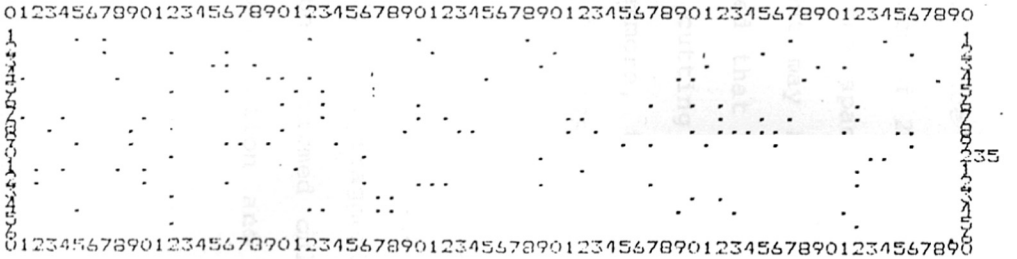
The rice repetitive pRL7-0.4 kbp sequence (Y-axis) is compared to the rice 25S[(A)], 17S[(C)] rDNAs and the spacer fragment between 25S and 17S rDNAs[(B)] (X-axis). The window of comparison was 48 and 26 bases at a time with 50% identity required for a positive result to be recorded.



(A)



(B)



(C)

FIG. 4.19

while short lines indicate homology over a shorter stretch in discrete packets of sequences.

From Fig. 4.19 it is clear that the sequence pRL7-0.4 kbp shows more homology with the nucleotide sequence of 25S rDNA fragment as compared to that of 17S rDNA and the spacer fragment. This indicates that the pRL7-0.4 kbp fragment may be a part of rice rDNA unit. It has already been reported that in rice rDNA fragment of 8.9 kbp, there is a BamH1 site cutting it into two fragments of 3.8 and 5.0 kbp [27]. Furthermore, this 3.8 kbp fragment is present in both the 25S as well as 17S rDNA genes. Thus, the observations that the recombinant pRL7 consists of an insert (repetitive DNA) of size 3.6 kbp which further has a BamH1-PstI fragment of about 0.4 kbp showing some homology with rice 25S rDNA increase the possibility that the 0.4 kbp fragment may be a part of rDNA unit. However, this can be confirmed only by using homologous/heterologous probes for hybridization and work in this direction is in progress.

REFERENCES

1. Jelinek, W.R. (1982)
Annu. Rev. Biochem., 51 : 813-844.

2. Sorenson, J.C. (1984)
Adv. Genet., 22 : 109-144.

3. Sala, C., Biasini, M.G., Morandi, C., Nielson, E., Parisi, B. and Sala, F. (1985)
J. Plant Physiol., 118 : 409-419.

4. Evans, I.J., James, A.M. and Barnes, S.R. (1983)
J. Mol. Biol., 170 : 803-326.

5. Pental, D. and Barnes, S.R. (1985)
Theor. Appl. Genet., 70 : 185-191.

6. Metzlaff, M., Troebner, W., Baldauf, F., Schlegel, R. and Cullum, J. (1986)
Theor. Appl. Genet., 72 : 207-210.

7. Hallden, C., Bryngelsson, T., Sall, T. and Gustafsson, M. (1987)
J. Mol. Evol., 25 : 318-323.

8. Orgel, L.E., Crick, F.H.C. (1980)
Nature, 284 : 604-607.

9. Deumling, B. (1981)
Proc. Natl. Acad. Sci., USA, 78 : 338-342.

10. Kato, A., Yakura, K. and Tanifuji, S. (1984)
Nuc. Acids Res., 12 : 6415-6426.

11. Yannisch-Perron, C., Vieira, J. and Messing, J. (1984)
Gene, 33 : 103.

12. Prentki, P., Karsch, F., Iida, S. and Meyer, J. (1981)
Gene, 14 : 289-299.

13. Norrander, J., Kempe, T. and MEssing, J. (1983)
Gene, 26 : 101-106.

14. Miller, J. (1972)
IN : Experiments in Molecular Genetics, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 431-432.

15. Mandel, M. and Higa, A. (1970)
J. Mol. Biol., 53 : 159-162.

16. Barnes, W.M. (1977)
Science, 195 : 393-394.

17. Birnhoim, H.C. and Doly, J. (1979)
Nuc. Acids Res., 7 : 1513- 1523.

18. Grunstein, M. and Hogness, D.S. (1975)
Proc. Natl. Acad. Sci.(USA), 72 : 3962-3965.

19. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982)
In : Molecular Cloning : A Laboratory Manual. Cold Spring Harbor University, New York.

20. Messing, J. (1983)
IN : Methods in Enzymology, 101 (part C); Recombinant DNA, R. Wu, L. Grossman, K. Moldave (eds.). Academic Press, New York, 20-28.

21. Schreier, P.H. and Cortese, R. (1979)
J. Mol. Biol., 129 : 169-172.

22. Yamamoto, K., Alberts, B., Benzinger, R., Lawhorne, L. and Treiberm G. (1970)
Virology, 40 : 734-740.

23. Sanger, F., Nicklen, S. and Coulson, A.R. (1977)
Proc. Natl. Acad. Sci., USA, 74 : 5463-5467.

24. Takaiwa, F., Oono, K., Iida, Y. and Sugiura, M. (1985)
Gene, 57 : 255-259.

25. Takaiwa, F., Oono, K. and Sugiura, M. (1984)
Nuc. Acids Res., 12(13) : 5441-5448.

26. Takaiwa, F., Oono, K. and Sugiura, M. (1985)
Plant Mol. Biol., 4 : 355-364.

27. Reddy, P.S. and Padayatty, J.D. (1987)
Ind. Journ. Biochem. Biophys., 24 : 293-295.

CHAPTER 5

PREDOMINANCE AND TISSUE SPECIFICITY OF ADENINE METHYLATION IN RICE GENOME

ABSTRACT

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 cloned repeat sequence indicated an abundance of adenine methylation. Although, cytosine methylation in 5'-CCGG-3' sequences suggested more of CpC methylation than CpG; the 'C' methylation in general 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.

5.1 INTRODUCTION

Role of 5-methyl cytosine (5mC) is 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 like rabbit β-globin and chicken ovalbumin [1,2,3]. In plants, methylation patterns of rDNA have been well studied. In general, the ribosomal genes seem to be heavily methylated in a few plant species investigated so far [4-9], except the rice rDNA unit. In rice, there is no change in rDNA methylation from seedlings grown in aerobic and anaerobic conditions [10], though there is a change in the synthesis of rDNA genes. Tissue or stage specific variations are observed in radish rDNA [11] and zein-light chain genes [12]. 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' [13-17].

In the present chapter, I 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 they are variety specific, similar experiments were carried out using the DNAs of two more varieties i.e. R-24 and Sona.

5.2 MATERIALS AND METHODS

5.2.1 Isolation of total DNA and repetitive DNA

Native, high molecular weight DNAs were prepared from embryos and shoots of three different varieties of rice (cvs. Basmati 370, R-24 and Sona) according to the procedures described earlier (Section 2.2.3).

Long (9-20 kbp) and short (0.2-0.3 kbp) repeat DNA sequences reassociating by Cot 0.1 and Cot 50 were isolated as mentioned in Section 3.2.2.

5.2.2 Restriction enzyme analysis and gel electrophoresis

Restriction enzyme digestions of all the DNAs with a few methylation specific enzymes like MboI, Sau3AI, DpnI, MspI and HpaII were carried out according to Maniatis et al [18]. The DNA digests were analysed on 1% and 1.4% agarose gels or 10% polyacrylamide gels, using both the high and low molecular weight DNA markers.

5.2.3 Southern hybridization

The total pRL7 DNA (isolated as described in Chapter 4) was labelled by nick translation [19] and was used as a probe.

Southern hybridization experiments were essentially according to Maniatis et al [18].

5.3 RESULTS

For assesssing the methylation status of rice DNA, the two most commonly used sets of restriction enzymes i.e. (1) MboI/Sau3AI/DpnI and (2) MspI/HpaII were employed. MboI/Sau3AI/DpnI are isoschizomers which recognise the same sequence 5'-GATC-3' but show differential sensitivity towards 'A' and 'C' methylation as per the instructions of the supplier (BRL). MboI digests the DNA when adenine is not methylated while Sau3AI digests it irrespective of methylation. However, DpnI will cleave the DNA only if 'A' is methylated. In addition, Sau3AI does not digest the DNA if the cytosine in sequence 5'-GATC-3' is methylated, whereas MboI cuts the DNA irrespective of 'C' methylation.

MspI/HpaII is another isoschizomeric pair which is specific for the sequence 5'-CCGG-3'. These enzymes, however, differ in their sensitivities towards the external and internal cytosine methylation as per the instructions of the supplier (BRL). HpaII does not cleave the DNA when internal 'C' residue of the sequence

5'-CCGG-3' is methylated, whereas, its digestion is unaffected by external 'C' methylation. On the other hand, MspI cuts the sequence 5'-CCGG-3' only when the internal cytosine is methylated.

5.3.1 Methylation status of rice shoot DNA

Figure 5.1 (A) shows the restriction enzyme digestion patterns of total rice (cv. Basmati 370) shoot DNA with MboI/Sau3AI/DpnI. From this figure, it is clear that MboI digestion of rice DNA is less than that of Sau3AI indicating 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.

Restriction enzyme digestion of total rice shoot DNA with MspI/HpaII is shown in Fig. 5.1 (B). It can be seen from this figure that HpaII shows comparatively more digestion of DNA than that with MspI suggesting the abundance of 5'-^{m}CCGG-3' type of sequences rather than 5'-C^{m}CGG-3' type in rice genome. This also suggests a high frequency of CpC methylation as compared to CpG methylation in 5'-CCGG-3' sequence in rice DNA.

From the above data, it is clear that rice DNA has less of 'C' methylation as compared to 'A' methylation. Recently, McCouch et al [20] have described less of cytosine methylation as compared to other plants like tomato, maize and rye.

FIGURE 5.1 : DIGESTION OF TOTAL RICE (cv. BASMATI 370) SHOOT DNA WITH MboI, Sau3AI, DpnI (A) and MspI, HpaII (B)

(A) λ HindIII digest (lane a)
MboI digest (lane b)
Sau3AI digest (lane c)
DpnI digest (lane d)

(B) λ 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.

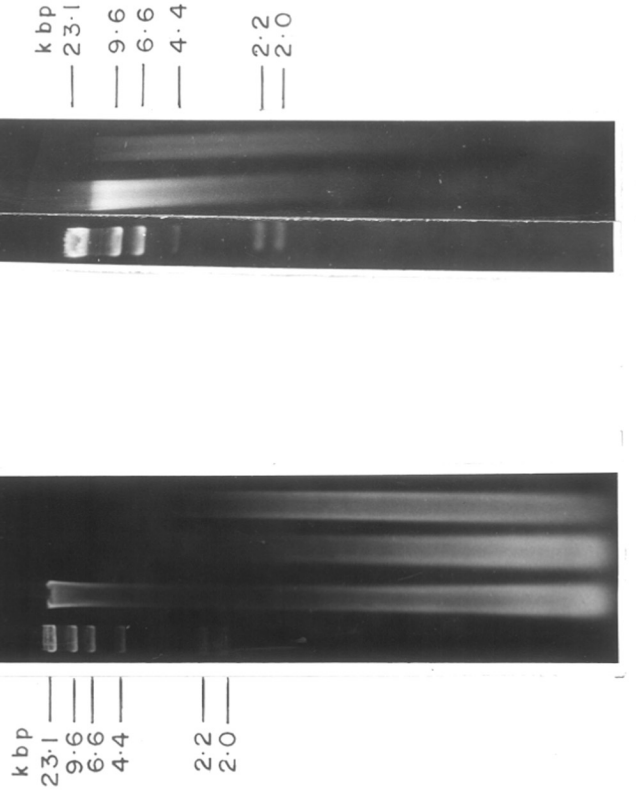


FIG. 5·1

5.3.2 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 sequences. For this purpose, long (9-20 kbp) and short repeats (0.2-0.3 kbp) 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 Figs. 5.2 A,B. From Fig. 5.2 A, it is evident that the digestion of the long repeats of Cot 0.1 and Cot 50 DNAs is less with MboI than that with Sau3AI, thus revealing the presence of methylated adenine in 5'-GATC-3' sequences in rice repetitive DNA fractions. This is further confirmed by the appreciable digestion of the DNAs with DpnI. A comparatively more digestion of the repeat sequences by Sau3AI also suggests the absence of 'C' methylation in the sequence 5'-GATC-3'. Since, DpnI digests Cot 50 DNA more than the Cot 0.1 DNA, it indicates that the adenine methylation of the DNA sequences reassociating between Cot 0.1 and 50 is higher than those reassociating by Cot 0.1.

Fig. 5.2 B shows the digestion patterns of long repeats of Cot 0.1 and 50 with MspI and HpaII. Here, HpaII digests the DNAs more than that with MspI suggesting again the abundance of 5'-^{m}CCGG-3'/5'-CCGG-3' type of sequences in these repeat fractions. However, between Cot 0.1 and 50 repetitive DNA sequences, long repeats of only Cot 0.1 show bands with MspI (4.6 kbp) and HpaII (4.9 kbp). This data indicates the occurrence of a high copy number of 5'-CCGG-3' sequence in highly repetitive DNA as compared to the total repetitive DNA.

FIGURE 5.2 : 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) λ HindIII digest (lane a)

MboI digest (lane b)

Sau3AI digest (lane c)

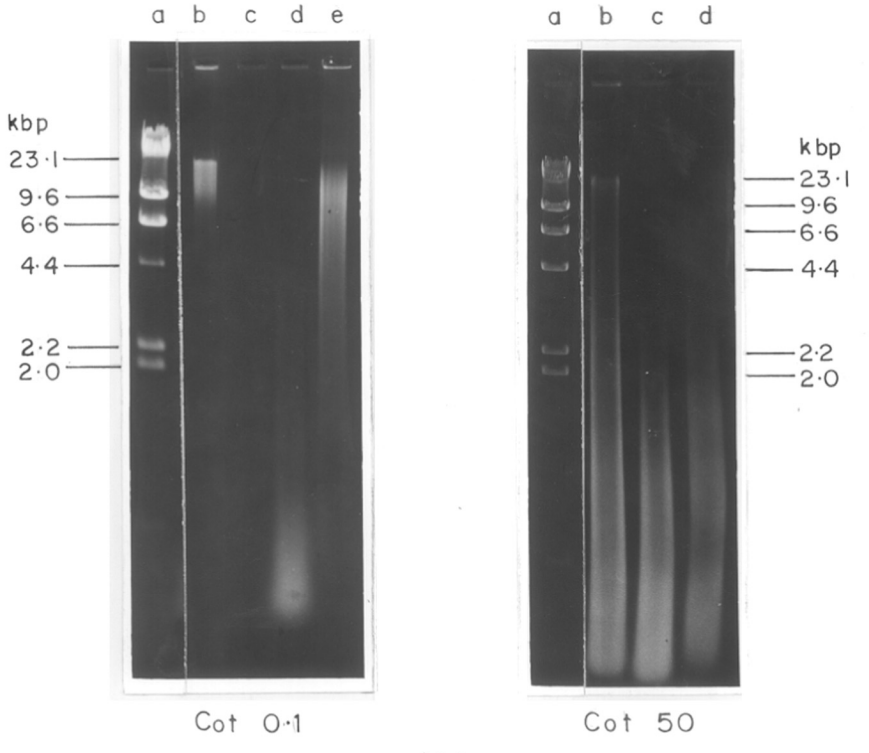
DpnI digest (lane d)

(B) λ HindIII digest (lane a)

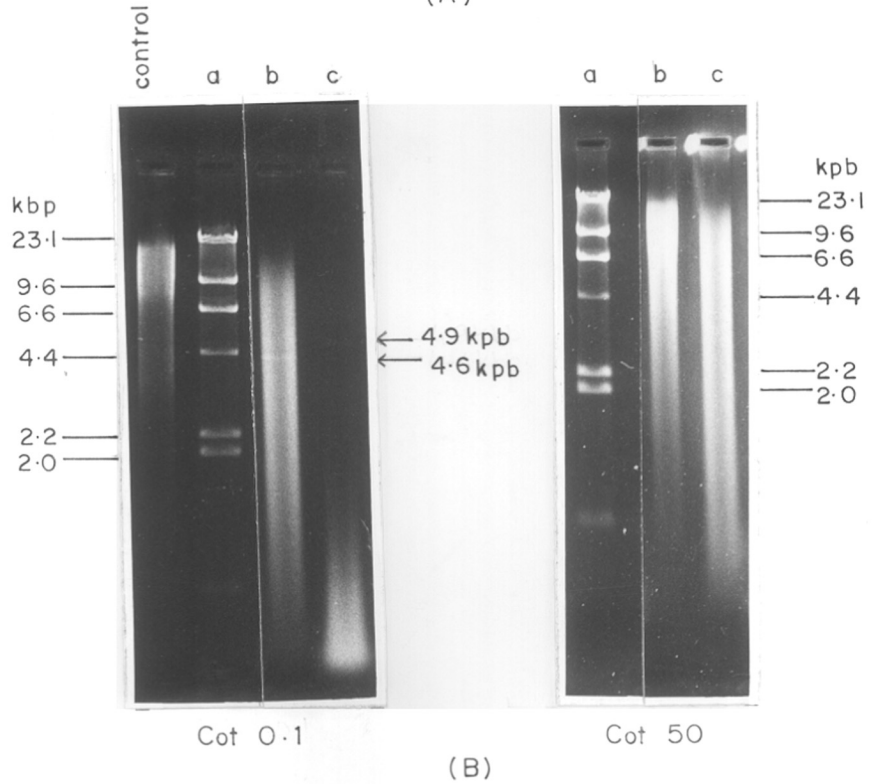
MspI digest (lane b)

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.


a b c d e
kbp
23·1
9·6
6·6
4·4
2·2
2·0
Cot 0·1
a b c d
kbp
23·1
9·6
6·6
4·4
2·2
2·0
Cot 50

(A)


control
a b c
kbp
23·1
9·6
6·6
4·4
2·2
2·0
Cot 0·1
4·9 kpb
4·6 kpb
a b c
kpb
23·1
9·6
6·6
4·4
2·2
2·0
Cot 50

(B)

FIG. 5·2

Figure 5.3 (A,B) depicts polyacrylamide gel electrophoresis of the digests of short repeat sequences of both Cot 0.1 and Cot 50 with the methylation specific enzymes. It is seen from this figure that the short repeats are too small to reveal any distinct differences.

Long repeats of rice DNA were cloned in pBR325 and the methylation status of three cloned repeats (pRL1, pRL7, pRL10) was determined. Figure 5.4 A shows the digestion patterns of pRL7 DNA with MboI and Sau3AI. Here it is seen that MboI does not show any digestion while Sau3AI shows a number of bands in the range 1.36 kbp to 0.072 kbp. This differential digestion with MboI and Sau3AI suggests the methylation of adenine in 5'-GATC-3' sequences in the cloned pRL7 repeat element. Similar digestion patterns were obtained with pRL1 and pRL10. When MspI and HpaII digestions of pRL7 DNA are compared (Fi.g 5.4 B), nearly identical pattern is obtained and many bands in the size range of 1.36 kbp-0.072 kbp are observed in both the digests. This is true for pRL1 and pRl10 also. This indicates the presence of a number of unmethylated 5'-CCGG-3' type of sequences in these cloned repeat units.

FIGURE 5.3 : DIGESTION OF SHORT REPEATS OF COT 0.1 (A) AND COT 50 (B) OF RICE (cv. BASMATI 370) WITH MboI, Sau3AI, DpnI AND MspI, HpaII

(A) and (B) ØX174HaeIII digest (lane a)

MboI digest (lane b)

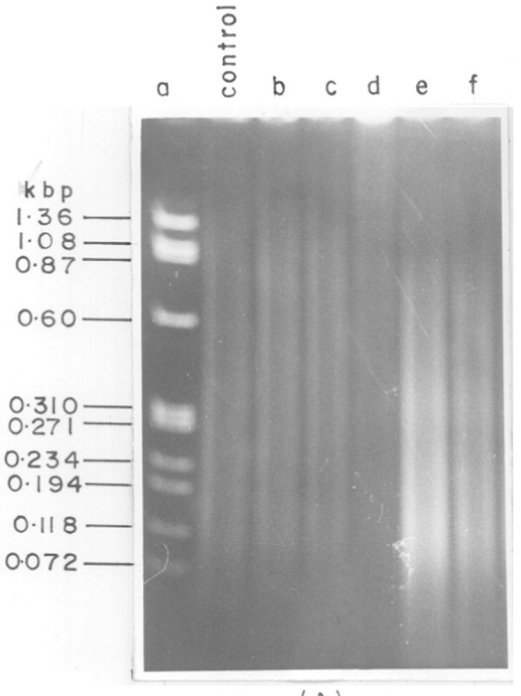
Sau3AI digest (lane c)

DpnI digest (lane d)

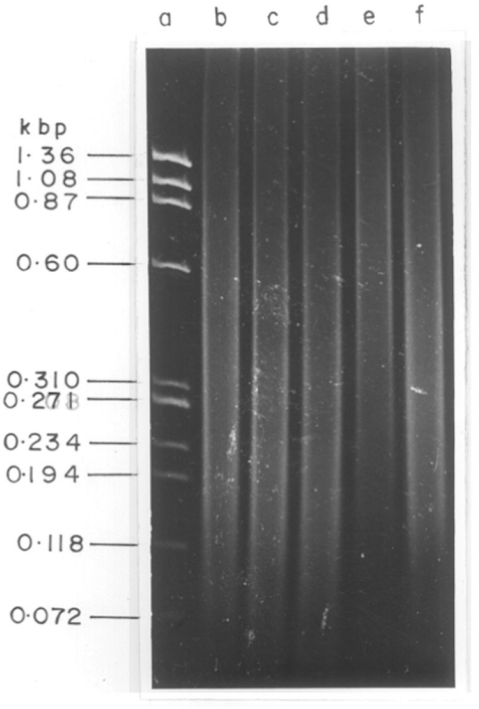
MspI digest (lane e)

HpaII digest (lane f)

Electrophoresis was carried out on 10% (A and B) polyacrylamide gels in TBE buffer (pH 8.0) at a constant current of 20 mA.


a
control
b
c
d
e
f
kbp
1·36
1·08
0·87
0·60
0·310
0·271
0·234
0·194
0·118
0·072

(A)


a
b
c
d
e
f
kbp
1·36
1·08
0·87
0·60
0·310
0·271
0·234
0·194
0·118
0·072

(B)

FIG. 5·3

FIGURE 5.4 : DIGESTION OF pRL7 INSERT DNA WITH MboI, Sau3AI (A) AND MspI, HpaII (B)

(A) ØX174HaeIII digest (lane a)

Sau3AI digest (lane b)

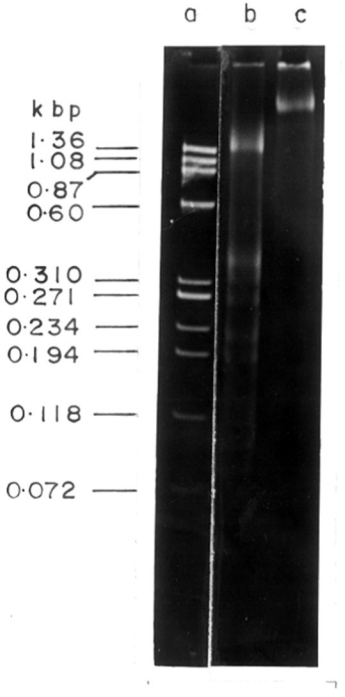
MboI digest (lane c)

(B) ØX174HaeIII digest (lane a)

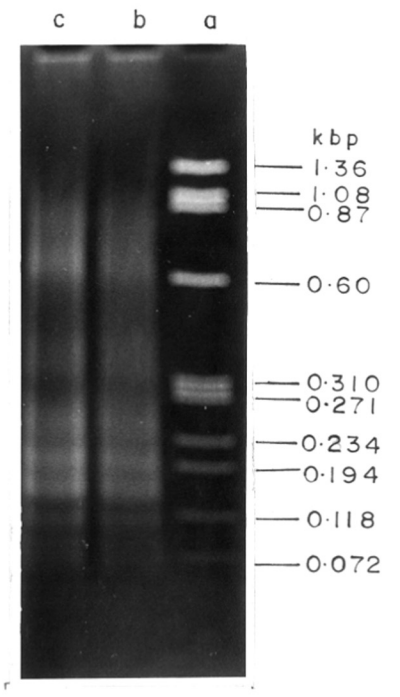
MspI digest (lane b)

HpaII digest (lane c)

Electrophoresis was carried out on 10% polyacrylamide gels in TBE buffer (pH 8.1) at a constant current of 20 mA.


a
b
c
kbp
1·36
1·08
0·87
0·60
0·310
0·271
0·234
0·194
0·118
0·072

(A)


c
b
a
kbp
1·36
1·08
0·87
0·60
0·310
0·271
0·234
0·194
0·118
0·072

(B)

FIG. 5·4

5.3.3 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; and MspI, HpaII. It is clearly seen from Figs. 5.5 A,B that the trend of 'A' and 'C' methylation is reversed in embryo DNA. For example, MboI and Sau3AI digest the embryo DNA extensively suggesting a low frequency of adenine methylation (Fig. 5.5 A). This pattern is again confirmed by very little digestion of embryo DNA with DpnI. Furthermore, Fig. 5.5 B shows MspI digestion to be more than HpaII suggesting a predominance of 5'-$C^{m}CGG$-3' type of sequences. Thus 5mC 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 0.6-4.0 kbp suggesting the presence of these repeat families which are absent in the corresponding shoot DNA digests.

The presence or distribution of a specific cloned pRL7 repeat sequence was next studied in the embryo and shoot system of cv. Basmati 370. 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 (Figs. 5.6 A,B) and were hybridized to ^{32}P-labelled plasmid pRL7 DNA. It can be seen from Figs. 5.7 A,B that a single

FIGURE 5.5 : DIGESTION OF RICE (cv. BASMATI 370) EMBRYO DNA WITH MboI, Sau3AI, DpnI (A) AND MspI, HpaII (B)

(A) λ HindIII digest (lane a)

MboI digest (lane b)

Sau3AI digest (lane c)

DpnI digest (lane d)

(B) λ 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.

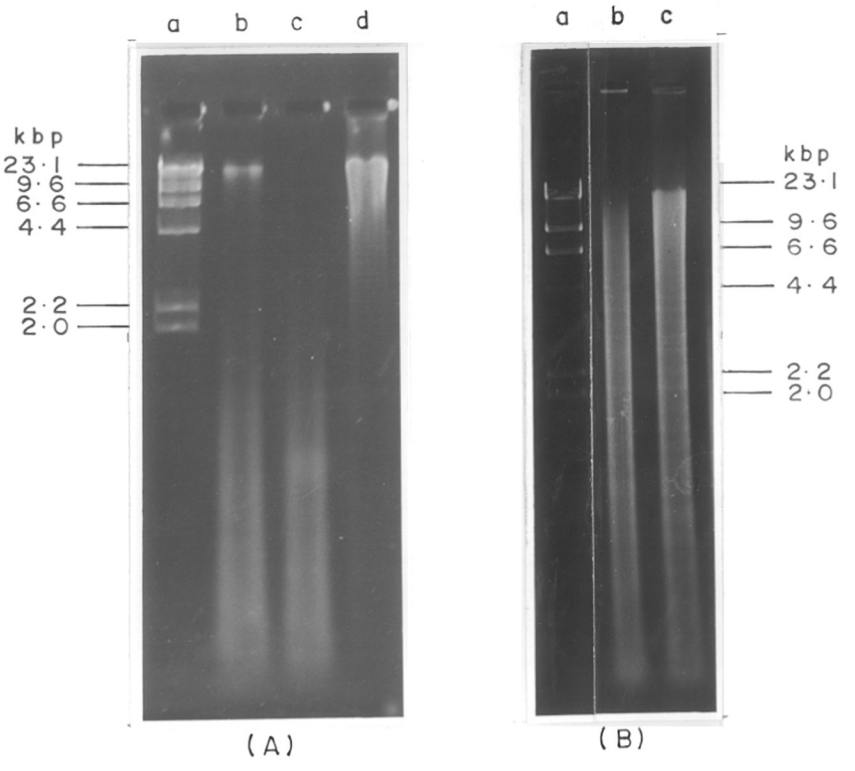


FIG. 5·5

155

strong band is present per lane with each enzyme with the absence of any background smear in case of embryonic DNA. The molecular weight of this band varies in a very small range of 2.3 to 2.5 kbp. In case of shoot DNA, the hybridization pattern is observed to be different with each enzyme (Fig. 5.7 B). With MboI (lane a), for example, two bands of molecular weights of about 15 and 4.5 kbp are prominent. With Sau3AI (lane b), four bandsin the region 0.6 to 1.36 kbp and one band less than 0.3 kbp are observed. In case of AluI and TaqI (lane c and lane d), however, many bands are seen in the range 0.3 to 0.9 kbp and 1.1 to 6.6 kbp.

From these results thus, it is clear that there is some reorganization in this specific cloned repeat unit in embryo and shoot system.

5.3.4 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 5.8 and 5.9 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/DpnI clearly indicate the predominance of 'A' methylation in the shoot

FIGURE 5.6 : DIGESTION OF TOTAL RICE EMBRYO (A) AND SHOOT DNA (B) WITH DIFFERENT RESTRICTION ENZYMES

(A) AND (B) λ HindIII digest (lane a)

MboI digest (lane b)

Sau3AI digest (lane c)

AluI digest (lane d)

TaqI digest (lane e)

ϕX174HaeIII 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.

FIGURE 5.7 : 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)

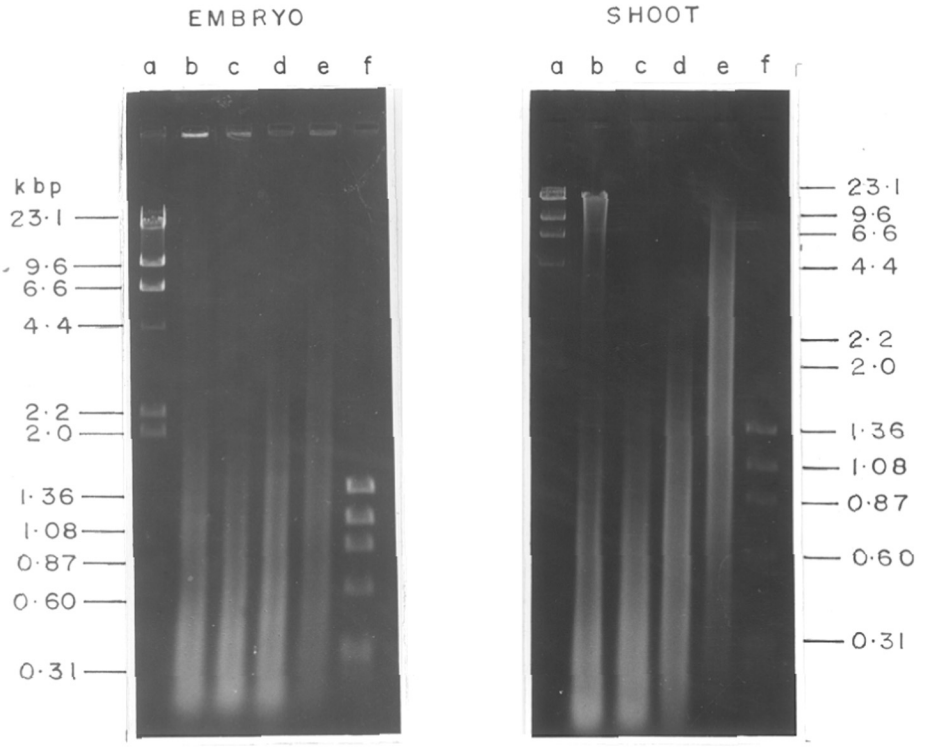

EMBRYO
a b c d e f
kbp
23·1
9·6
6·6
4·4
2·2
2·0
1·36
1·08
0·87
0·60
0·31
SHOOT
a b c d e f
23·1
9·6
6·6
4·4
2·2
2·0
1·36
1·08
0·87
0·60
0·31

FIG. 5·6

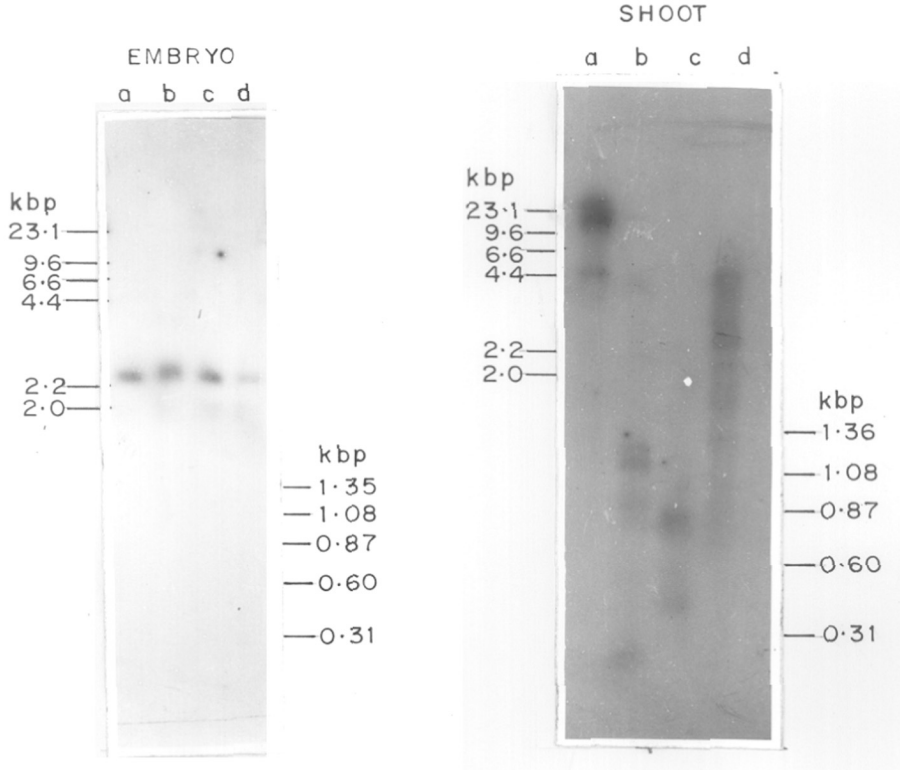

EMBRYO
a b c d
kbp
23·1
9·6
6·6
4·4
2·2
2·0
kbp
1·35
1·08
0·87
0·60
0·31
SHOOT
a b c d
kbp
23·1
9·6
6·6
4·4
2·2
2·0
kbp
1·36
1·08
0·87
0·60
0·31

FIG. 5·7

DNAs as compared to the embryo DNAs in both the varieties. Also, the MspI/HpaII digestion patterns in these DNAs are coherent with that of Basmati 370, wherein a qualitative change in 5mC from CpG to CpC dinucleotides is observed during the transition from embryo to shoot.

Another finding from Figs. 5.8 and 5.9 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 digest shows two bands of 2.6 kbp and 0.9 kbp, while Sau3AI reveals a single band of 2.6 kbp. This indicates the presence of regularly arranged, unmethylated 5'-GATC-3' sequences in the Sona shoot DNA. Also, it shows that there are some 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, the Sau3AI repeat element is absent, whereas the MboI repeat family is present. The R-24 embryo DNA, when digested with MboI, shows many bands with varying molecular weights in the range 0.5-2.5 kbp; while these are absent in the corresponding shoot DNA digest.

When the 2.6 kbp band which is present in the Sau3AI digest of Sona shoot DNA and absent in the embryo DNA is eluted, labelled and hybridised 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. 5.10).

FIGURE 5.8 : DIGESTION OF RICE (cvs. R-24 AND SONA) EMBRYO (A) AND SHOOT (B) DNAs WITH MboI, Sau3AI, DpnI

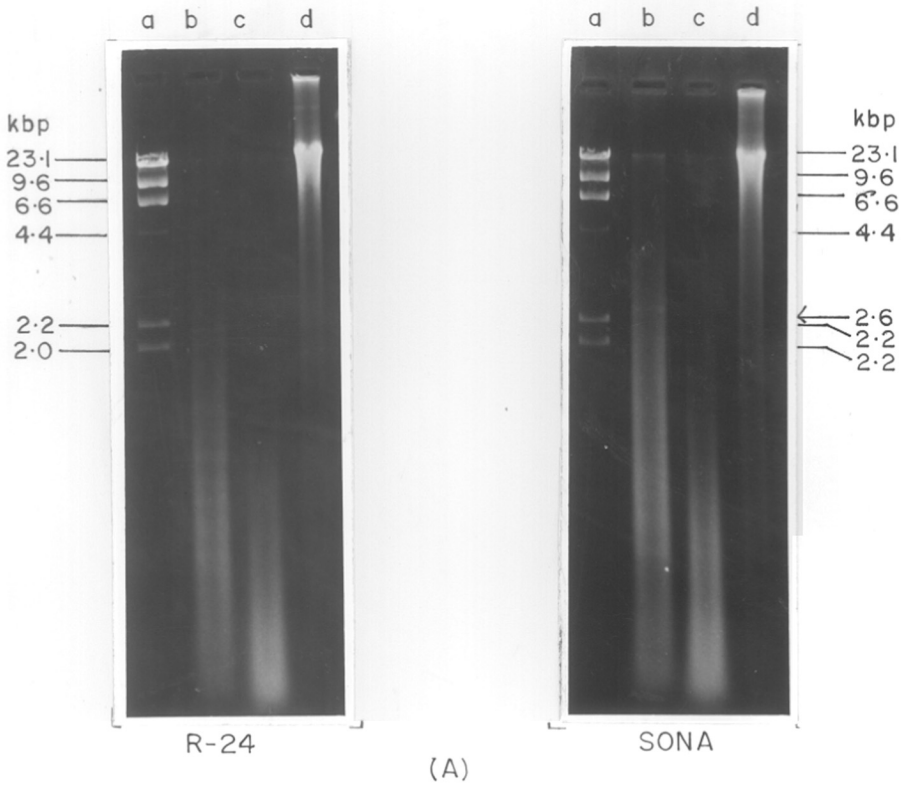
(A AND B) λ HindIII digest (lane a)

MboI digest (lane b)

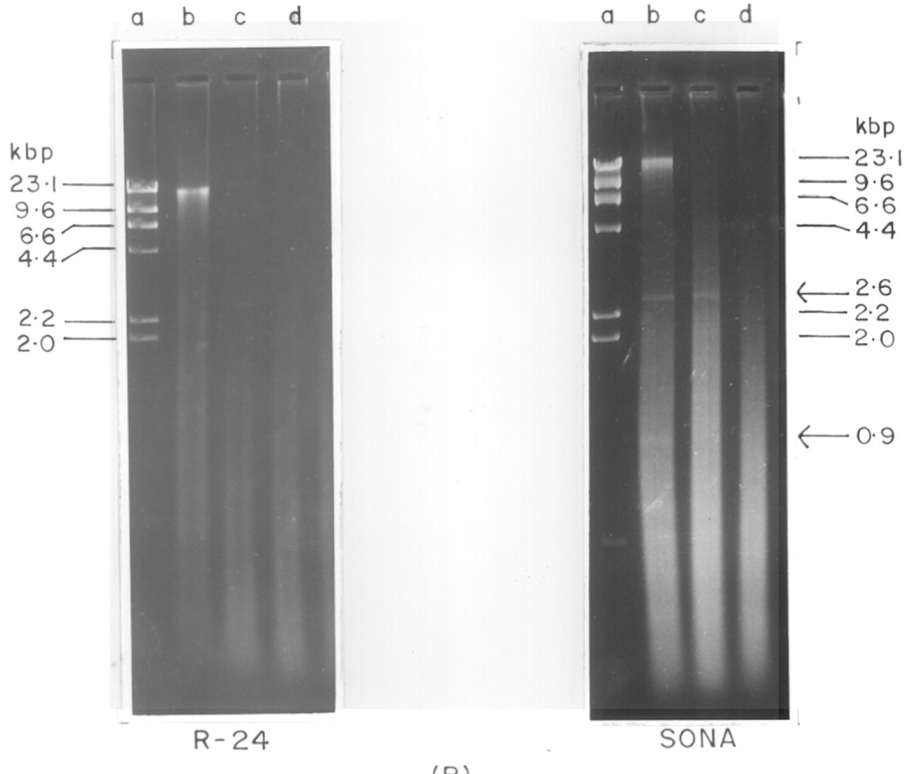
Sau3AI digest (lane c)

DpnI digest (lane d)

Electrophoresis was carried out on 1.4% (A) and (B) neutral agarose slab gels in TAE buffer (pH 8.1) at a constant current of 30 mA.


a b c d
kbp
23·1
9·6
6·6
4·4
2·2
2·0
R-24
a b c d
kbp
23·1
9·6
6·6
4·4
2·6
2·2
2·2
SONA

(A)


a b c d
kbp
23·1
9·6
6·6
4·4
2·2
2·0
R-24
a b c d
kbp
23·1
9·6
6·6
4·4
2·6
2·2
2·0
0·9
SONA

(B)

FIG. 5·8

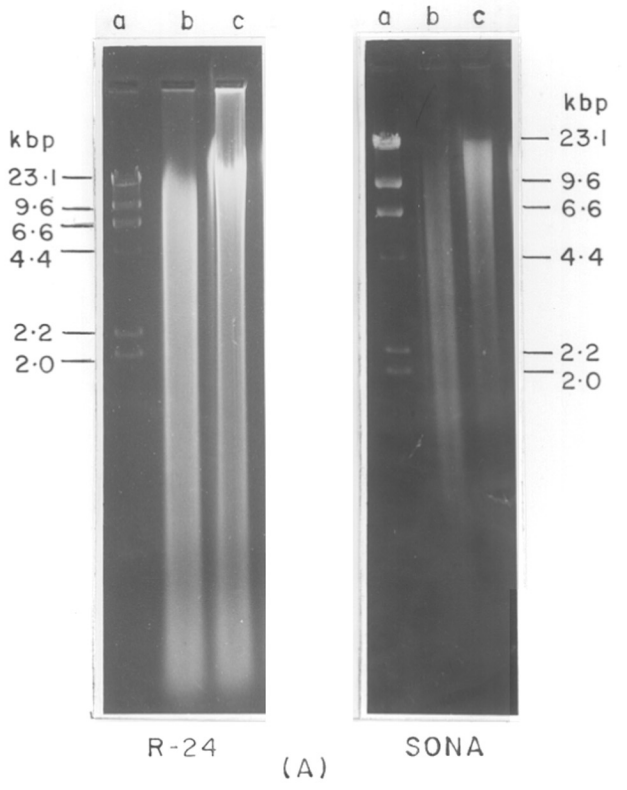
FIGURE 5.9 : DIGESTION OF RICE (cvs. R-24 AND SONA) EMBRYO (A) AND SHOOT (B) DNAs WITH MspI, HpaII

(A AND B) λ HindIII digest (lane a)

MspI digest (lane b)

HpaII digest (lane c)

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


a b c
kbp
23·1
9·6
6·6
4·4
2·2
2·0
R-24
(A)
a b c
kbp
23·1
9·6
6·6
4·4
2·2
2·0
SONA

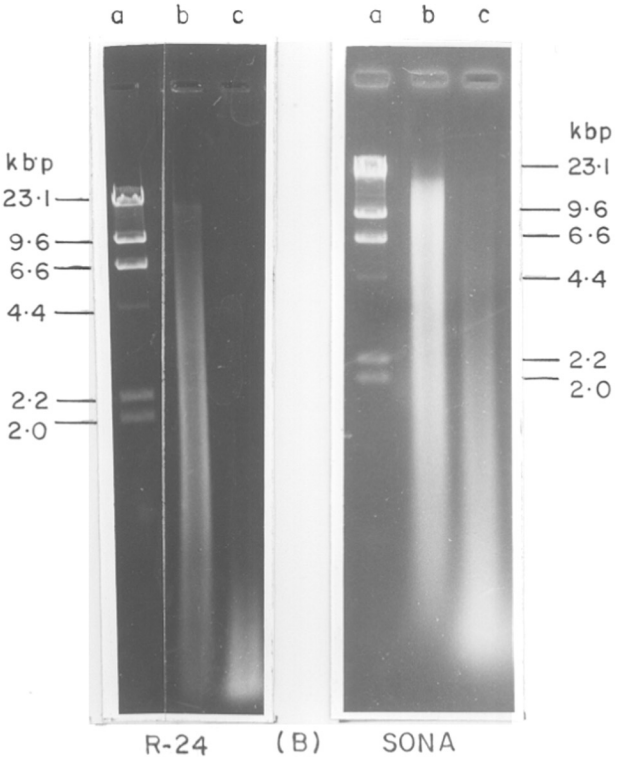

a b c
kbp
23·1
9·6
6·6
4·4
2·2
2·0
R-24
(B)
a b c
kbp
23·1
9·6
6·6
4·4
2·2
2·0
SONA

FIG. 5·9

FIGURE 5.10: DIGESTION PATTERN AND AUTORADIOGRAM OF SONA EMBRYO DNA DIGESTED WITH Sau3AI AND HYBRIDISED TO SONA SHOOT Sau3AI REPEAT FAMILY (2.6 kbp)

λ 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)

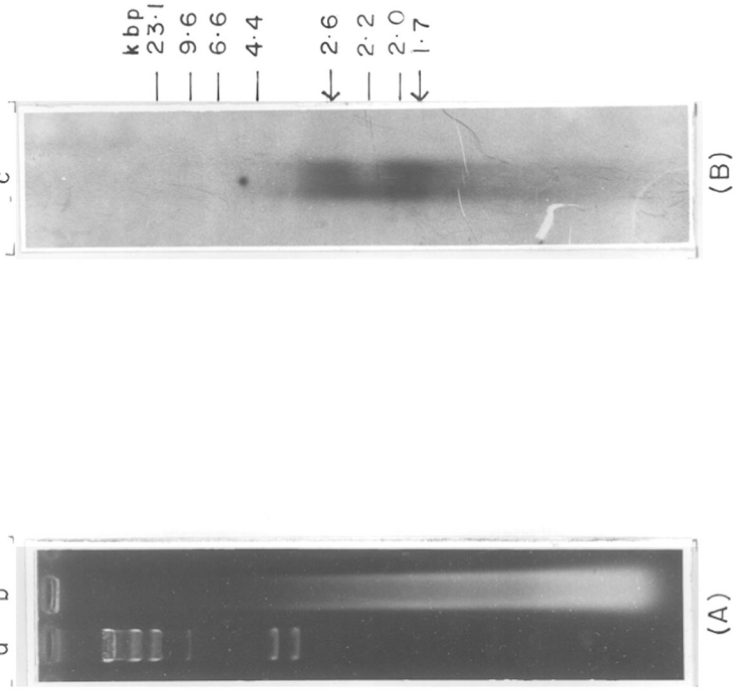


FIG. 5·10

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.

5.4 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 [21]. 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 sequences in this process. The occurrence and role of adenine methylation is still unclear in higher plants. Recently, methylated adenine (mA) has been observed in zein genes, but the experimental evidences suggest that it is unrelated to their regulation of expression [21]. In case of animals and microorganisms, however, mA (methylated adenine) has defined roles [22-28]. For instance, in E. coli, the rate of adenine methylation within 5'-GATC-3' sequences modulates the activity of certain promoters [24,25]. Engel and Von Hippel [26] 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 favoured. Such effects may be significant in protein-DNA interactions [27,28].

Another important feature of the present work is the genomic rearrangement of a specific cloned repeat element during the transition from embryo to shoot stage. When the hybridization patterns of the embryo and shoot DNA digests (using enzymes sensitive to 'A' methylation) with the cloned pRL7 probe are compared, this repeat unit is present in a well organised manner in the embryo DNA. In case of shoot DNA, on the other hand, the hybridization pattern suggests that many copies of this sequence are present in a varied environment with random distribution, while some may be well organised 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 cloned repeat. In the literature, there are a few reports on genomic alterations in tissue culture conditions. In rice, for example, Sala et al [16] have reported the amplification of a specific highly repeated sequence in cultured cells of Oryza sativa cv. Roncarlo during differentiation process. Oono et al [17] have given preliminary results showing variations in the copy number of two rice (Oryza sativa cv. Nipponbare) repeated DNA sequences, one being amplified and the other being reduced in copy number in the dedifferentiation process also under cultured conditions. Ours is

probably the first report which describes the differential methylation status of a specific rice (Oryza sativa cv. Basmati 370) repeat sequence during the transition from embryo to shoot without any external inducing factor.

REFERENCES

1. Waalwijk, C. and Flavell, R.A. (1978)
Nuc. Acids Res., 5 : 4631-4641.

2. Mandel, J.L. and Chambon, P. (1979)
Nucl. Acids Res., 7 : 2081-2103.

3. Compere, S.J. and Palmiter, R.D. (1981)
Cell, 25 : 233-240.

4. Gerlach, W.L. and Bedbrook, J.R. (1979)
Nuc. Acids Res., 7 : 1869 - 1885.

5. GoldsbroughP.B., Ellis, T.H.N. and Cullis, C.A. (1981)
Nuc. Acids Res., 9 : 1551-1558.

6. Uchimiya, H., Kato, H., Ohgawara, T., Harada, H. and Sugiura, M. (1982)
Plant Cell Physiol., 23 : 1129 - 1131.

7. Ellis, T.H.N., GoldsbroughP.B. and Castleton, J.A. (1983)
Nuc. Acids Res., 11 : 3047 - 3064.

8. Segel, A. and Kolacz, K. (1983)
Plant Physiol., 72 : 166 - 171.

9. Olmedilla, A., Delcasso, D. and Delseny, M. (1984)
Plant Sci. Lett., 37 : 123-127.

10. Aspart, L., Got, A., Delseny, M., Mocquot, B. and Pradet, A. (1983)
Plant Physiol., 72 : 115 - 121.

11. Delseny, M., Laroche, M. and Penon, P. (1984)
Plant Physiol., 76 : 627-634.

12. Spena, A., Viotti, A. and Pirrota, V. (1983)
J. Mol. Biol., 169 : 799-811.

13. Durante, M., Cremonini, R., Brunori, A., Avanzi, A. and Innocenti, M. (1977)
Protoplasma, 93 : 289-304.

14. Cremonini, R., Loiero, M. and Durante, M. (1981)
Protoplasma, 108 : 1-7.

15. Bassi, P., Cionini, P.G., Cremonini, R. and Seghizzi, P. (1984)
Protoplasma, 123 : 70-77.

16. Zheng, K.L., Castiglione, S., Biasini, M.G., Biroli, A., Morandi, C. and Sala, F. (1985)
Theor. Appl. Genet., 74 : 65-70.

17. Kikuchi, S., Takaiwa, F. and Oono, K. (1987)
Mol. Gen. Genet., 210 : 373-380.

18. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982)
In : Molecular Cloning : A Laboratory Manual, Cold Spring Harbor University, New York.

19. Rigby, P.W.J., Dieckman, M., Rhodes, C. and Berg, O. (1977)
J. Mol. Biol., 113 : 237-251.

20. 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.

21. Pintor-Toro, J.A. (1987)
Biochem. Biophys. Res. Comm., **147** : 1082-1087.

22. Hare, J.T. and Taylor, J.H. (1985)
Proc. Natl Acad. Sci., USA, **82** : 7350-7354.

23. Pukkila, P.J., Peterson, J., Herman, G., Modrich, P. and Messelson, M. (1983)
Genetics, **104** : 571-582.

24. Hattman, S. (1982)
Proc. Natl. Acad. Sci., USA, **79** : 5518-5521.

25. Plasterk, R.H.A, Vrieling, H. and Van de Putte, P. (1983)
Nature (London), **301** : 344-347.

26. Engel, J.D. and Von Hippel, P.H. (1978)
J. Biol. Chem., **253** : 927-934.

27. Sterglanz, H. and Bugg, C.E. (1973)
Science (Washington D.C.), **182** : 833-834.

28. Soo-Chen Cheng, Gail, H. and Modrich, P. (1975)
J. Biol. Chem., **260** : 191-194.

CURRICULUM VITAE

CURRICULUM VITAE

Name : Mrs. Madhu S. Dhar

Current address : Division of Biochemical Sciences,
National Chemical Laboratory,
Pune 411 008. India.

Permanent address : 'Shreema', 101/24A, Ashok Housing Soc.,
Durganagar. Pune 411 016. India.

Nationality : Indian

Date of birth : September 28, 1961.

Marital status : Married

Academic qualifications :

Degree	University	Year	Class	Subject
B.Sc.	Poona, India	1981	I (77.6%)	Chemistry (Principal) Physics (Subsidiary) Maths (Subsidiary)
M.Sc.	Poona, India	1983	I (67.4%)	Biochemistry
Ph.D.	Poona, India	Thesis submitted		Chemistry

Title of the Ph.D. Thesis : A Detailed Genomic Analysis of a Higher Plant with Special Reference to Rice.

Awards

- Recipient of National Merit Scholarship from 1976-1978.

- Recipient of Junior Research Fellowship of the Council of Scientific and Industrial Research (CSIR), New Delhi, from 1984-1986.

- Recipient of Senior Research Fellowship of the Council of Scientific and Industrial Research (CSIR), New Delhi, from 1986-1989.

Membership of Scientific Bodies

- Member, Society of Biological Chemists, India.

- Member, Indian Society of Cell Biology.

- Member, International Society of Plant Molecular Biology.

Participation at Workshops/Symposia

- Attended the workshop on 'Organization of Genetic Material' held in Zoology Department, University of Poona from 10/3/84 to 3/4/84.

- Attended the workshop on 'Gene cloning and DNA sequencing' held at School of Biological Sciences, Madurai Kamaraj University from 30/12/84 to 21/1/85.

- Was one of the instructing faculty at the workshop on 'Plant Molecular Biology' held in the Division of Biochemical Sciences, National Chemical Laboratory, Poona from 10/6/85 to 30/6/85.

- 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 from 10/9/87 to 30/9/87.

Was one of the instructing faculty at the workshop on 'Analysis of a cloned plant DNA fragment and Agrobacterium mediated gene transfer in Plants', held in the Division of Biochemical Sciences, National Chemical Laboratory, Poona from 28/3/89 to 17/4/89.

List of Publications

1. M.M. Dabak, S.A. Ranade, M.S. Dhar, V.S. Gupta and P.K. Ranjekar (1988)

   'Molecular Characterization of Pigeonpea genome'

   Indian Journal of Biochemistry and Biophysics, 25(3), 230-236.

2. M.S. Dhar, M.M. Dabak, V.S. Gupta and P.K. Ranjekar (1988)

   'Organization and properties of repeated DNA sequences in Rice genome'

   Plant Science, 55(1), 43-52.

3. S.A. Ranade, M.D. Lagu, S.M. Patankar, M.M. Dabak, M.S. Dhar, V.S. Gupta and P.K. Ranjekar (1988)

   'Identification of a dispersed MboI repeat family in five higher plant genomes'

   Bioscience Reports (U.K.), 8(5), 435-442.

4. V.S. Gupta, M.S. Dhar, B.G. Patil and P.K. Ranjekar

   'Molecular cloning, restriction enzyme analysis and sequencing of a specific long repetitive DNA sequence in rice'

   Manuscript under preparation

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

   'Predominance and tissue specificity of adenine methylation in rice'

   Manuscript under preparation

Papers presented at Symposia/Meetings

1. Madhu Malick, Meena S. Karve and N.R. Kale (1983)

   'A fluorometric method for the assay of neutral and acidic carbohydrates'

   Presented at the 52nd Annual Meeting of the Society of Biological Chemists (India) held at Poona during Nov. 26-28, 1983.

2. P.K. Ranjekar, Madhu Malick, Madhavi Atre and S.A. Ranade (1984/85)

   'Plant Vector Development'

   Presented at DAE symposium on 'Newer Approaches to Biol- Applications', held at Baroda during Oct. 19-21, 1984.

3. M.S. Dhar, M.D. Lagu, M.M. Dabak and S.A. Ranade (1986)

   Restriction endonuclease cleavage analysis in four higher plant genomes'

   Presented at the 55th Annual Meeting of the Society of Biological Chemists .(India) held at Trivandrum during Dec. 14-17, 1986.

4. V.S. Gupta, M.S. Dhar, V.V. Pethe and P.K. Ranjekar (1989)

   'Modulation of a cloned rice repetitive DNA sequence during development'

   Presented at the Symposium on 'Current Trends in Biochemical Research in Pune' held at the Department of Chemistry, University of Poona, Feb. 3, 1989.

REFERENCES

1. Dr. P.K. Ranjekar
Scientist E2,
Division of Biochemical Sciences,
National Chemical Laboratory,
Poona 411 008. India.

2. Dr. J. Mascarenhas
University of Albany,
State University of New York,
Department of Biological Sciences,
Albany, New York 12222. USA.

3. Dr. Mrs. Vidya S. Gupta
Scientist C,
Division of Biochemical Sciences,
National Chemical Laboratory,
Poona 411 008. India