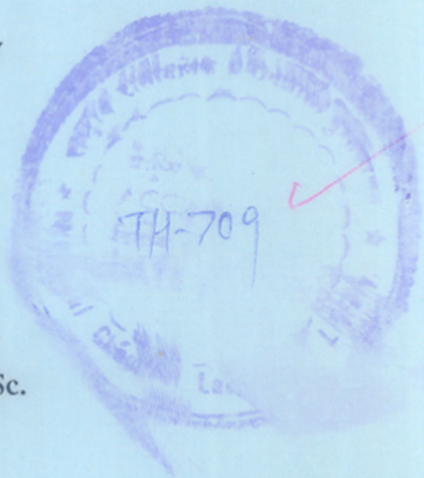
SOME MOLECULAR STUDIES ON PROLAMIN AND GLUTELIN SEED STORAGE PROTEIN GENES IN BASMATI-370 (ORYZA SATIVA L.)

COMPUTERISED

A THESIS
submitted to the
UNIVERSITY OF POONA
for the Degree of
DOCTOR OF PHILOSOPHY
(in Chemistry)



By
VIVEK RAVI VARMA
M.Sc.

Plant Molecular Biology Unit
Division of Biochemical Sciences
National Chemical Laboratory
Pune-411 008 (INDIA)

1993

CONTENTS

COMPUTERISED

CHAPTER 2: CONSTRUCTION OF A RICE GENOMIC LIBRARY : IDENTIFICATION AND ISOLATION OF PUTATIVE PROLAMIN CLONES HOMOLOGOUS TO KAFIRIN

CHAPTER 3 :ANALYSIS OF 5'UPSTREAM REGIONS OF STORAGE PROTEIN GENES OF MONOCOTS AND DICOTS WITH RESPECT TO FREE ENERGY AND DINUCLEOTIDE FREQUENCIES

CHAPTER 4 OPTIMISATION OF PROTOPLAST ISOLATION CONDITIONS FOR THREE INDICA RICE CULTIVARS

ACKNOWLEDGEMENTS

The deep sense of gratitude that I feel towards my guide, Dr. P.K. Ranjekar, cannot be acknowledged to the full by mere words. His quiet confidence and dignified patience in helping me realise my full potential has left a lasting influence on me. With his blessings, I am sure that I will attain the high ambitions that I have set for myself.

The guidance and expert advise of Dr. (Mrs.) V.S. Gupta, Dr. R.R. Hendre and Dr. (Mrs.) N.A. Sahasrabudhe during my entire tenure at N.C.L is appreciated with all my heart. The excellent technical assistance rendered by Sujatha Krishnamurthy deserves a special mention. The enthusiastic typing of Ajit Ranjekar also deserves a special thanks.

The working environment created in the lab by Dr. (Mrs.) M.D. Lagu, Dr.(Mrs) S.A. Tamhankar, Dr. (Mrs.) M.N. Sainani, Dr. (Mrs.) R.H. Bhate, Mr. B.G. Patil, Dr. (Mrs.) L.S. Kumar, Suman, Ramakrishna, Milind, Stella, Savita, Renuka, Sami Subbu, Venkat, Sastry, Manisha, Gagan, Usha, Sarita, Alankar, Jagtap and Karunakaran will never be forgotten.

The encouraging assistance rendered by Dr. Praful Chitnis aand Pushkaraj Apte during emotionally trying times cannot be summed up by a mere thank you.

The morale boosting company of Satish Kulkarni,

Sanjeev Kamte, K.A. Parakh, Sanjeev Khatri, Abhijit Bhave, Anil T, G. Satyanarayan, Ganesh, Aditi, Neelam, and Snehal is gratefully acknowledged.

The enthusiastic company of Snehal, Jawahar, Abhay Sagare, Suhasini, Gulshan, Madhur, Sanjay Palnitkar, Urmila, Hemant, Sangita, Sunita, Mohini, Geeta, Jana and others did help in making the Division an enjoyable place to work in. The encouragement given by Dr. (Mrs.) K.S. Joshi and Dr. (Mrs.) S. Hazra is warmly acknowledged.

I am grateful to the financial assistance from the C.S.I.R which helped me to pursue a higher degree in the field of Plant Molecular Biology. The financial assistance rendered by Rockefeller Foundation is also sincerely acknowledged.

I am thankful to Dr. R.A. Mashelkar, Director for allowing me to work at N.C.L and submit the results in the form of a thesis. The help rendered by Prof. J. Barnabas and Dr. M.C. Srinivasan, in their capacities as the past and present Head of Division of Biochemical Sciences is sincerely acknowledged.

The love and affection showered on me by my uncle and aunt, Mr. and Mrs. M.M.G. Nambudiri, Meenakshi, Sanal and Jayasree during my stay in Pune makes my eyes moist with happiness.

The immense contribution of my parents and brothers, Vinay and Tilak, by their utmost tolerance to my eccentricities is priceless and I would not like to reduce its greatness by attempting to acknowledge it.

Vivek Varma.

DECLARATION

Certified that the work incorporated in the thesis **"Some molecular studies on prolamin and glutelin seed storage protein genes in Basmati-370 (Oryza sativa L.)"**, submitted by **Mr. Vivek R. Varma** was carried out under my supervision. Such material as has been obtained by the candidate from other sources has been duly acknowledged in the thesis.

P K Ranjekar

Dr. P.K. Ranjekar

Research Guide

ABBREVIATIONS

M	-	Molar
mM	-	milli moles
SDS	-	Sodium dodecyl sulphate
°C	-	Degree centrigrade
EDTA	-	Ethylene diamine tetraacetic acid
PFU	-	Plaque forming unit
ug	-	microgram
rpm	-	revolutions per minute
mg	-	milligram
RT	-	Room temperature
mg/l	-	milligram per liter
O/N	-	Overnight
2,4-D	-	2,4-Dichlorophenoxy acetic acid
IAA	-	Indole acetic acid
NAA	-	Naphthalene acetic acid
BAP	-	Benzyl amino purine
CM	-	Coconut milk
MS	-	Murashige and Skoog
A	-	Adenine
T	-	Thymine
G	-	Guanine
C	-	Cytosine
Kbp	-	Kilo base pairs
O.D	-	Optical density
Tris	-	Tris hydroxy amino methane

ATP	-	Adenosine triphosphate
dATP	-	deoxyadenosine triphosphate
dCTP	-	deoxycytosine triphosphate
dTTP	-	deoxythymidine triphosphate
dGTP	-	deoxyguanosine triphosphate
DTT	-	Dithio threitol
cpm	-	counts per minute
BSA	-	Bovine serum albumin
SSC	-	Saline sodium citrate

LIST OF FIGURES

LIST OF TABLES

SYNOPSIS

The indica rice type accounts for 80% of the cultivated rice and is the staple food for more than 2 billion people (1). This makes the indica rice variety an important system for genetic engineering and tissue culture studies.

In my thesis, my basic interest was in studying the expression of rice storage protein genes encoding prolamins. I constructed a genomic library of rice and screened the same for genes encoding prolamin. I then attempted to develop a protoplast system in a few Indica varieties including Basmati-370. After being successful in isolating rice protoplasts, I monitored transient expression of the gus gene, one of the most commonly used marker gene in plant systems. All my data involving gene work and protoplast work can be presented in the following manner.

Cloning of genes encoding prolamin seed storage proteins from an indica rice

Introduction

The proteins found in rice seeds consist of albumin, globulin, prolamin and glutelin. Unlike other cereals, the alcohol soluble prolamins comprise only 5 - 10 % of the total grain protein while glutelin accounts for about 80% of the total endosperm proteins

of rice (2, 3).

The prolamin genes are present in 50 -100 copies per haploid rice genome (4), while glutelin proteins are encoded by a small multigene family each consisting of 5 - 8 copies per haploid genome (8). The low level of expression of prolamin genes despite the presence of a high copy number made these genes interesting to study with regards to their regulation. There are only a limited number of reports in which genomic and cDNA libraries have been used to isolate the prolamin and glutelin genes of rice (4, 5, 6, 7). Hence investigative studies on prolamin genes will help in understanding of the mechanisms involved in their mode of regulation.

Objective

The initial major objective of this study was to construct a genomic library of rice and clone the genes encoding prolamin seed storage proteins.

Results

(i) **Construction of a rice genomic library**

This work involved three major steps :

(a) Isolation of High molecular weight rice DNA and its partial digestion.

(b) Preparation of Charon 4A DNA and its arms.

(c) Development of an efficient *in vitro* packaging system.

The genomic library was constructed by ligating *EcoRI* digested rice DNA with Charon 4A arms and then packaging into the packaging extract obtained from the lysogenic *E. coli* SMR-10 culture. The titer of the library was 1-2 x 10^6 PFU/ml. Taking into consideration the haploid DNA content of rice (1.0 pg), it appeared that the library represented about 7- 8 rice genomes.

(ii) **Screening and identification of clones homologous to Kafirin**

The genomic library of rice was screened for putative prolamin clones, by using a heterologous probe, namely, Kafirin from *S.bicolor* (a gift from Prof. T.C. Hall). The recombinant phages were plated and plaques were lifted onto nitrocellulose membranes for hybridization with labelled heterologous probe. Around 10^5 plaques were screened to obtain three regions which gave signals on autoradiography. Further rounds of plaque lifting and hybridization purified these plaques which were then used for large scale DNA isolation. On further digestion and hybridization, the clones were confirmed to contain sequences that

were homologous to kafirin, the prolamin gene from S. bicolor.

This part of the work has been recently published. (9)

A theoretical analysis of 5' upstream regions of storage protein genes

Along with the experimental studies, I carried out a theoretical analysis of the upstream regions of storage protein genes. 5' upstream regions of storage protein genes were extracted from the Genbank data library and aligned at the translation initiation site. The nucleotide sequences were then classified into monocot and dicot storage protein genes and were analysed using various parameters like DNA homology, free energy distribution; and mono- and di- nucleotide frequencies. Analysis of the 5' upstream regions of monocot and dicot storage protein genes resulted in the following conclusions

There was a distinct difference in the 5' region immediately upstream of monocot and dicot storage protein genes with respect to their free energy distribution.

There was a considerable variation in the consensus sequences of monocots and dicots except for

a few positions which are conserved in both.

The dinucleotides TA and CA contributed to the unstable peak in the region -10 bp upstream of the start site in monocots.

The dinucleotides TA, TT and CA were major contributors to the unstable region -10 bp upstream of the start site in dicots.

Development of a gene transfer and expression system in Indica varieties of rice

Introduction

The indica varieties are widely cultured but unlike japonica rice are generally poor in somatic embryogenesis and plant regeneration (10). Attempts to regenerate plants from indica rice (11) and cell suspension cultures (12) have resulted in limited success due to their poor in vitro cultural response. Moreover, there are genetic variations in terms of in vitro culture within the indica group (13).

Only a small number of japonica (14,15) and an even smaller number of indica varieties (16,17) have been transformed by combining direct DNA transfer and protoplast technologies. Recovering fertile transgenic

rice from protoplasts is not a trivial procedure and this limitation has delayed the application of recombinant DNA methodology to the improvement of elite rice varieties.

Rice in vitro culture has a very strong genotype and culture dependent component. Even though genetic engineering of rice has been reported, only cultivars from japonica varieties could be transformed using protoplast dependent methods. Only recently, transformation and subsequent regeneration of indica rice plants from a specific cultivar was reported utilising protoplasts (18).

Objective

The primary goal of our work was to standardise the protoplast isolation conditions and to develop a gene expression system in an elite Indica cultivar.

Results

Three Indica varieties namely, Basmati-370, VDN-3630, and Indrayani were chosen for this study. The scentedness, fine grain and good cooking quality of the three varieties made them suitable candidates for genetic improvement.

(i) **Callus initiation and maintenance**

I used mature seeds of three indica varieties for callus induction. The successful induction of callus from all the varieties was optimised by inoculation onto media containing various combinations of auxins and cytokinins. The callus thus obtained was maintained by serial subculture.

(ii) **Optimisation of conditions for protoplast isolation**

Protoplasts were isolated from the callus of three varieties namely, Basmati-370, VDN-3630 and Indrayani. The parameters studied for obtaining a high yield of protoplasts included age of culture, enzyme concentration, incubation period, enzyme-mixture volume, pH of the enzyme solution, different osmotica and osmoticum concentration.

(iii) **Transient gene expression**

The protoplasts were used in transformation experiments in order to assess their competence for direct DNA uptake with the help of polyethylene glycol. Transient expression was obtained in all the three indica varieties under investigation.

The protoplasts were plated onto various media like N6 and AA with different osmoticum concentrations in order to obtain microcallus.

Thesis Plan

The thesis is organised in the following fashion :

CHAPTER I

Here I have compiled most of the literature pertaining to molecular and gene expression studies with special reference to seed storage protein genes in rice.

CHAPTER II

In this chapter, I have carried out some molecular studies in an elite Indica rice cultivar namely, Basmati-370. This involves construction of a rice genomic library, screening the library with a heterologous probe and identifying putative clones that are homologous to kafirin, a prolamin gene of S. bicolor.

CHAPTER III

Here, I have carried out a theoretical analysis of the 5' upstream regions of storage protein genes using various approaches like DNA homology, free energy distribution; and mono- and di-

nucleotide frequency distribution.

CHAPTER IV

This chapter consists of the tissue culture investigations carried out in indica varieties and includes :

1) Media optimisation for callus induction and maintenance.
2) Optimisation of protoplast isolation conditions for three indica varieties.
3) Transient gene expression studies in three indica varieties.

Chapter V

This chapter summarises the results of preceding chapters and attempts to correlate them in order to obtain a holistic view.

REFERENCES

1) Swaminathan M.S (1982)

Science, **218** : 967-972.

2) Tecson E.M.S., Esmann B.V., Lontok L.P., and Juliano B.O.,(1971)

Cereal Chem. **48** : 181-186.

3) Villareal R.M., and Juliano B.O., (1978)

Phytochemistry **17** : 177-182.

4) Kim T.W., and Okita T.W., (1988)

Plant Physiol. **88** : 649-655.

5) Masumura T, Hibino T, Kidzu K, Mitsukawa N, Tanaka K, and Fujii S (1990)

Mol. Gen. Genet. **221** : 1-7.

6) Takaiwa F., Ebinuma H., Kikuchi S., and Oono K., (1987)

FEBS Lett. **221** (1) : 43-47.

7) Wang C-S., Shastri K., Wen L., Huang J-K., Sonthayanon B., Muthukrishnan S., and Reeck G.R., (1987)

FEBS Lett. **222** (1) : 135-138.

8) Okita T.W., Hwang Y.S., Hnilo J, Kim T.W., Aryan A.P., Larson R., and Krishnan H.B. (1989)

J. Biol. Chem. **264** (21) :12573-12581.

9) Joshi K.S., Varma V.R., Gupta V.S., and Ranjekar P.K., (1992)

J. of Plant Biochem. and Biotech. **1** : 33-36.

10) Abe T and Sasahara T (1982)
Jpn. J. Breed. **32** : 53-60.

11) Yamada Y, Zhi-Qi Y, Tang D-T (1986)
Plant Cell Rep. **5** :85-88.

12) Ye H.C. (1984)
Acta Bot. Sinica. **26** : 52-54.

13) Abe T and Futsuhara Y (1986)
Theor. Appl. Genet. **72** : 3-10.

14) Tada Y, Sakamoto M, and Fujimura T (1990)
Theor. Appl. Genet. **80** : 475-480.

15) Hayashimoto A, Li Z, and Murai N (1990)
Plant Physiol. **93** : 857-863.

16) Datta S.K., Peterhans A, Datta K, and Potrykus I (1990)
Bio/technology **8** : 33-38.

17) Christou P, Ford T.L, Kofron M (1991)
Bio/technology **9** : 957-962.

18) Peng J, Kononowicz H and Hodges T. K (1992)
Theor. Appl. Genet. **83** : 855-863.

CHAPTER I

STATUS OF RESEARCH ON SEED STORAGE PROTEINS AND GENES IN RICE

1.1 Introduction

The rapid progress in recombinant DNA technology has made it possible to manipulate genes in order to produce plants with altered genotypes and phenotypes. Traits like herbicide tolerance, insect resistance and virus resistance can be incorporated into plants overcoming the limitations of sexual compatibility. The nutritional characteristics of seed proteins are an important aspect of seed quality and may be directly amenable to manipulation through genetic engineering. The economic advantages of increasing the nutritional quality of seed proteins are enormous. The manipulation of seed storage proteins is, however, a multidisciplinary approach involving aspects such as structural characterisation of proteins, cloning and analysis of genes encoding these proteins, site directed mutagenesis of genes followed by an understanding of the overall seed biology as such; and development of a suitable transformation system for monitoring the expression of native and modified seed storage protein genes.

Considerable progress has been made in research involving seed storage proteins and genes in rice. Villareal and Juliano (1) are among the first to study the properties of glutelin seed storage proteins in

rice during development, using SDS-PAGE, showing differences in the amount of glutelin fractions in indica and japonica cultivars. Later, Luthe (2) attempted to monitor the accumulation of storage proteins in developing seeds using ^{35}S sulfate and have also ascertained the relationship of rice glutelin with other seed storage proteins (3). Okita and his group and Krishnan et al have done an indepth analysis involving isolation, biochemical characterisation, and localisation of seed storage proteins of rice (4,5). Their studies have also included isolating and characterising prolamin genes of rice (6). Takaiwa and his co-workers (7,8) have concentrated on isolating glutelin genes and classifying them according to their nucleotide sequence homology. Studies on mechanism of developmental expression of glutelin genes have also been in progress recently. Ray Wu and his group, for example, (9) are interested in identifying regions that contribute to the tissue specific expression of glutelin genes. Tanaka et al (10) are working on the isolation of prolamin clones from cDNA libraries constructed from developing rice seeds and have sequenced some prolamin clones to investigate the homology of these clones with other cereal prolamins. Chen et al (11) have recently cloned and sequenced prolamin clones and have studied the synthesis of

prolamin polypeptides in rice seeds.

Most of the above investigations have been carried out in japonica cultivars like Nipponbare, Mangetsumochi, and Lebonnet. In our laboratory, seed storage protein studies have been undertaken in an indica variety, namely Basmati-370, and these investigations have included biochemical and biophysical characterisation of proteins (12) followed by cloning of prolamin and glutelin genes (13). An attempt will now be made to summarise the overall information that is available on seed storage proteins and genes in rice.

1.2 Rice grain structure and features of rice seed storage proteins

The rice grain a caryopsis, is a dry one-seeded fruit with its pericarp fused with seed coat. The main parts of a seed are husk, pericarp, endosperm and embryo. A peculiarity of the rice embryo is the presence of a well- developed epiblast which is a thin membranous structure situated opposite the scutellum. The outer layer of endosperm cells, which occupies a major part of rice seed constitutes an aleurone layer. Rice seed proteins are mostly contained in the aleurone and subaleurone layer of parenchymatous cells. The endosperm consists of only

parenchymatous cells containing starch grains of the compound type. The rice grain is broadly divided into two types viz., glutinous (waxy) and non-glutinous (non-waxy), depending on the ratio between amylose and amylopectin content of starch (14,15). The glutinous types are characterised by having a low amylose and a high amylopectin content, whereas the non-glutinous types have a high amylose and low amylopectin content.

Unlike other cereals which have prolamins as their predominant protein fraction, rice seed proteins contain large amounts of glutelins. The protein of the starchy rice endosperm typically consists of about 15% albumins and globulins, 5% prolamins and 80% glutelins (16). Another feature of storage proteins in rice is their unique segregation into two immunologically distinct protein bodies, namely PB-I and PB-II (17). PB-I is spherical with a concentric ring structure and contains prolamin, while PB-II is an irregularly shaped structure stained homogeneously by OsO_4 and is rich in glutelin and globulin (17).

Storage protein accumulation in developing rice seeds begins about 6 days post anthesis (DPA) (18). The first protein that is formed is glutelin which comprises 80% of the total seed protein at maturity

(19). Glutelin is sequestered in protein bodies which appear in the rice endosperm at about 7 DPA (20). SDS-PAGE of the protein bodies reveals 2 major subunit groups with apparent molecular weights of 37 - 39 kD and 22 - 23 kD (21). Isoelectric focusing data reveals that the 37 - 39 kD polypeptides are acidic and the 19 - 23 kD polypeptides are basic in nature as in oat (22), soybean (23) and peas (24). It has been shown that glutelin subunits are synthesized as a 57 kD precursor which is subsequently processed in the golgi apparatus (25). The processing mechanism is similar to that as in soybean (26), pea (24) and oat (27). Rice glutelin appears to be preferably synthesized on membrane bound polysomes and is deposited in the vacuoles of sub aleurone cells (25).

Biochemical studies have clearly shown that prolamins accumulate only around 8 - 10 days after flowering (1). The prolamins simply aggregate by dilation of the rough ER and the protein bodies are enclosed by a single lipid bilayer membrane with attached ribosomes (1). Rice prolamins are deposited in the starchy endosperm and have a 13 kDa polypeptide as their major component (28). They appear to be synthesised as a precursor molecule by membrane-bound polysomes attached to PB-I and rough endoplasmic reticulum

and to pass through the membrane into the lumen where they are deposited as aggregates. Yamagata et al (17) have shown that prolamins are present as 10, 13 and 16 kDa polypeptides which also show microheterogeneity in their sequence.

Rice globulins form 8% of the total seed storage proteins and have not been studied very extensively. SDS-PAGE studies reveal 25 and 16kDa bands which are most abundant in the globulin fraction. The 26kDa rice globulin has no cross reactivity with any other seed proteins while the 16kDa displays considerable cross reactivity with similar size polypeptides from other cereals like barley, wheat, and rye, with the exception of pea. Immunogold staining demonstrates that rice globulins are deposited along with rice glutelins in PB-II. Both polypeptides have a 2kDa signal peptide that facilitates their movement across the ER membrane (5). Rice albumins account for about 5 % of the total seed proteins. Mawal et al (12) have identified a 60 kD monomer which is a major fraction of rice albumin and is present maximally between 18 - 20 days post anthesis.

1.3 Structure and organisation of seed storage protein genes

In general, seed storage protein genes are encoded

RR
577.2'.633.18:631.563(043)
VAR

TH-709

by multigene families which exhibit a high degree of conservation and linkage among themselves. The number of genes constituting gene families ranges from 5 - 100 (29) which singly or in clusters; on the same or on different chromosomes.

Okita et al (30) have divided the multigene family of glutelin genes into three subfamilies, Gt1, Gt2 and Gt3 each containing about 5-8 copies. Takaiwa et al, (31) on the other hand, have grouped them into two subfamilies, subfamily A and subfamily B and stated that the Gt3 subfamily could be classified into subfamily A. The glutelin genes subfamilies share a 60-65% homology with each other while there is a 80% homology amongst the subfamily members indicating a clear distinction between the subfamilies (31). An unusual feature of the 5' flanking region of rice glutelin genes is the presence of four very long direct repeat sequences which may be involved in enhancing their expression as shown in soybean β-conglycinin gene (32). Multiple polyadenylation signals, AATAAA or AATAAC, are present in all the glutelin subfamilies which are clustered in Gt1 and Gt2 or separated by 8 bp in Gt3.

Prolamin genes of rice have been first isolated by Kim and Okita (6). They have isolated 3

cDNA clones and have shown that 2 of them have 95% homology with each other while the third has only 63% homology with the other two clones. Genomic Southern blots indicate the presence of 80-100 copies of prolamin genes per haploid genome of rice. Despite the heterogeneity amongst prolamin clones, the isolated polypeptide has an identical molecular weight i.e. 15.2kD. The 5' and 3' untranslated region of prolamin genes are highly heterologous and the coding region does not contain any introns. Putative TATA and CAAT regulatory sequences are present at -271 and -330bp from the translation start site. The proteins formed are immunologically distinct from other cereal prolamins. A genomic clone of prolamin isolated from a lambda Charon 35 library exhibits tandemly arranged repeat fragments which are also observed in wheat gamma - gliadin (33) and maize zein genes (34). The presence of several additional potential AUG translation start sites and multiple polyadenylation signals, AATAAG and AATAAA, in the 3' untranslated segments is also observed. Masumura et al isolated a 10 kDa (10) and 13 kDa cDNA (28) prolamin clones that are present in about 7 - 10 copies/haploid genome. The 13 kDa prolamin gene possesses the AACATGAA sequence which resembles the plant consensus initiation sequence of AACATGGC (35). cDNA sequences of rice globulins have been isolated

but have not yet been characterised (5). Till date, reports on the genetic aspects of rice albumin have not been documented.

1.4 Regulatory features of rice seed storage protein genes

The expression of genes is usually controlled by sequences that are present upstream of the coding regions which are involved in the binding of certain factors that regulate DNA transcription. Such sequences are called as cis-acting elements and the factors are referred to as trans-acting factors. Cis-acting elements and trans-acting factors have been found to be responsible for tissue and developmental specific regulation of seed storage proteins in plants.

All members of the glutelin gene family show the presence of a sequence AACAAACTCTAT between -78 and -65bp upstream of the transcription start site (36). Takaiwa and Oono (37) have detected a legumin box (38) 60bp upstream of the transcription start site of glutelin genes and this box may be important in specifying tissue and developmental expression of glutelin. A consensus sequence TG(T/A/C)AA(A/G)A)(G/T) is present at -694, -486 and -253bp in the promoter regions of both the glutelin gene subfamilies. The dyad

sequence CATGCATG is a second motif prevalent in many seed tissue genes but not in genes from other plant tissues. A SV40 enhancer core-like sequence TGAAAAAT which occurs between -209 and -202 has been demonstrated to be correlated to the expression level of glutelin gene (37).

There are trans-acting factors that bind to specific sequence in the 5' upstream region of glutelin genes. CTTTCGTGTAC is present -130 to -120bp upstream of the transcription start site which binds a trans-acting factor (36). TGAGTCA and ATGCAAAT are two other protein binding sites. A TGAGTCA binding motif to which transcription factors jun and GCN4 are known to bind (39) is located -100bp and -164bp upstream of the transcription start site. Thus this motif might play an important role in tissue specific expression of glutelin gene. The other octameric binding site ATGCAAAT is separated from the TGAGTCA motif by 8-10bp of non-conserved sequence. These are also found to be a part of -300 element of several seed storage protein genes (40,41).

1.5 Homology and evolution of rice seed storage proteins

A similarity in the functional expression of seed storage protein genes has stimulated an interest in

understanding the phylogenetic relationship amongst cereals. Wen and Luthe (3) have shown that the α- and β subunits of rice glutelin share several similarities with other seed globulins. For example, the molecular weights of acidic α and basic β subunits are similar to those of other globulins. The α subunit composition is similar to that of some related monocot species like oat and coconut; while the β subunit composition is similar to that in many species like oat, coconut, pumpkin, soybean, broad bean and hemp. This suggests that except for solubility, rice glutelin is very much like globulin storage protein class of other seeds. But rice glutelin has been found to be immunologically distinct from other cereal proteins (42). Comparison of amino acid sequence of rice glutelin precursor with those of pea legumin A, soybean glycinin and rape cruciferin shows a 37 % homology to each of the 11S globulins (43). The amino acids of the basic subunit of rice glutelin show 74 % homology with a corresponding region of oat 11 S globulin. As the acidic subunit seems to be less conserved than the basic subunit, the α subunit may be the best target for improving the amino acid composition of rice glutelin. The cysteine residues in positions 112 and 313 and the post translational cleavage site between positions 306 (Asparagine) and 307 (Glycine) are strictly conserved

in all 11 S globulins and rice glutelin (44). Despite the divergence in primary structure, secondary structures of rice glutelin, pea legumin and soybean glycinin around the processing site are similar. Hydropathy values of the amino acid sequence of all the three proteins show similar profiles. These evidences may intensely suggest that rice glutelin, pea legumin and soybean glycinin may have similar molecular architectures and that rice glutelin and leguminous 11S globulins have evolved from a common ancestral gene (43).

A similarity of rice glutelin to legume 11 S proteins and oat globulins is also reflected at the genetic level. The 11 S legume proteins are encoded by a small multigene family; with six copies in soybean glycinin (45); eight copies in pea legumin (45) and 7 - 10 copies in oat globulins (45). Rice glutelin gene subfamilies share this feature in possessing 5 - 8 gene copies. The relative constancy of the number of gene copies per glutelin subfamily would be consistent with the hypothesis that genes encoding glutelin proteins already existed as a multigene family which were then duplicated intact. Divergence by point and segmental mutations would account for the diversity of the glutelin subfamilies (46).

Alcohol soluble prolamins of rice share a few common features with prolamins of other cereals. The prolamin genes of maize (47), wheat (48), and rice are encoded by a large gene family consisting of about 80 - 100 copies based on Southern blot analysis. Kim and Okita (1988) have isolated prolamin genes that lack repetitive sequences, a common structural feature of other major cereal prolamins (49). In addition, no significant homology is detected between rice prolamins and prolamins of other cereals (42). Based on these findings, it is suggested that rice prolamin genes have evolved from an independent ancestral gene distinct from that which gave rise to the prolamins of other major cereals. In contrast to these findings, Masumura et al (10,28) have isolated 10 kDa and 13 kDa prolamin cDNA clones from rice, whose polypeptide contained the unique octapeptide, Gln-Gln-Gln-Cys-Cys-Gln-Gln-Leu. This octapeptide has also been reported in maize 15 kDa zein (50) and 27 kDa zein (51); barley B1-hordein (52) and gamma - hordein (53). A comparison of 13 kDa cDNA sequence with cDNA sequences isolated by Kim and Okita reveals a homology of more than 70% indicating that the prolamin genes isolated by Okita et al and Masumura et al belong to the same family. The 30 % dissimilarity may account for the absence of repetitive sequences observed by Masumura et al (10) and not by Kim et al

(49). This also goes on to prove the heterogenous nature of prolamin genes. In another report, Okita et al (42) have shown that rice prolamins are immunologically distinct from other cereal prolamins.

Thus the distinctive property of rice prolamin as well as the close sequence homology evident between rice glutelin and legume seed storage proteins may indicate that rice has diverged very early during the evolution of grasses.

1.6 Expression of rice seed storage protein genes in homologous and heterologous systems

Seed storage protein genes provide an ideal system for the study of developmental and tissue-specific pattern of gene expression. These genes have also been introduced into heterologous systems in order to investigate the empirical nature of their expression.

Okita et al (46) have constructed a plasmid containing Gt3 promoter and the CAT gene and have electroporated it into tobacco to monitor the transient activity of the CAT gene. The CAT activity has been demonstrated in tobacco protoplasts indicating the presence of a functional promoter effective in directing transcription of CAT gene in plant cells.

Similar results have been obtained when a Gt2 - CAT construct is electroporated into rice protoplasts. The expression of Gt3 promoter-CAT construct has also been monitored in tobacco (54) and highest levels of CAT activity are obtained 16 days after flowering. In some plants, low levels of CAT activity are detected in certain non-seed tissues indicating that the expression of this monocot gene is not always strictly controlled in a heterologous tobacco host.

A deletion analysis of a glutelin promoter-GUS construct has shown that regions between -441bp and -237bp are required for temporal and endosperm specific expression of the GUS activity in tobacco cells (37). In this work, the gene promoter used might have been different from that of Gt3 since the more sensitive GUS reporter is not detected in any other tissue of the transgenic tobacco plants whereas the Gt3 promoter-CAT construct is detected in certain non-seed tissues as well (54).

Kim and Wu (9) have fused the 5' flanking sequence of a rice glutelin gene with GUS gene and have bombarded rice seeds with this construct to demonstrate transient gene expression in seeds after two days.

Thus seed storage proteins and genes of rice provide a singular system to study basic mechanisms of

gene regulation in plants. But the rice seed has also been utilised in various other investigations like mechanism of starch (55) and sucrose metabolic (56) pathway, molecular mechanisms of aleurone cell development (57), flavonoid gene regulation (58), waxy gene studies (59), characterisation of β-glucanase (60) and phospholipid transfer genes (61) and identification of seed-borne bacterial diseases of rice (62). Hence a multifaceted and coordinated approach for understanding the rice seed will be of great significance. This will be of immense help in improving the yield and quality of the rice grain so that it can provide a balanced diet for the rapidly expanding world population.

1.7 References

1) Villareal R.M. and Juliano B.O. (1978)
Phytochem. **17** : 177 - 182.

2) Luthe D.S. (1983)
Plant Sci. Lett. **32** : 147 - 158.

3) Wen T-n. and Luthe D.S. (1986)
Plant Physiol. **78** : 172 - 177.

4) Krishnan H.B. and Okita T.W. (1986)
Plant Physiol. **81** : 748 - 753.

5) Krishnan H.B., White J.A. and Pueppke S.G. (1992)
Plant Sci. **81** : 1 - 11.

6) Kim W.T. and Okita T.W., (1988)
Plant Physiol. **88** : 649 - 655.

7) Takaiwa F., Kikuchi S. and Oono K. (1987)
Mol. Gen. Genet. **208** : 15 - 22.

8) Takaiwa F. and Oono K. (1991)
Jpn. J. Genet. **66** : 161 - 171.

9) Kim S.Y. and Wu R. (1990)
Nucl. Acid Res. **18** : 6845 - 6852.

10) Masumura T., Shibata D., Hibino T., Kato T., Kawabe K., Takeba G., Tanaka K. and Fujii S. (1989)
Plant Mol. Biol. **12** : 123 - 130.

11) Shyur L-f., Wen T-n., and Chen C-s. (1992)
Plant Mol. Biol. **20** : 323 - 326.

12) Mawal Y.R., Mawal M.R., and Ranjekar P.K. (1987)

Biosci. Rep. (G.B.) **7(1)** 1 - 9.

13) Joshi K.S., Varma V.R., Gupta V.S., and Ranjekar P.K. (1992)
J. Plant Biochem. Biotech. **1** : 33 - 36.

14) Ghosh A.K. (1971)
Ph.D. Thesis.

15) Chakravarty T.K., Dwarikanath R.K. and Prabhakar B.B. (1972)
J. Fd. Sci. Tech. **9(2)** : 140 - 143.

16) Juliana B.O. (1972)
In : 'Rice Chemistry and Technology'
Ed. D.F. Houston pp 16 - 74.

17) Yamagata H. and Tanaka K. (1986)
Plant Cell Physiol. **27** : 135 - 145.

18) Cagampang G.B., Perdon A.A. and Juliano B.O. (1976)
Phytochem. **15** : 1425.

19) Juliano B.O. and Boulter D. (1976)
Phytochem. **15** : 1601.

20) Yamagata H, Tanaka K. and Kasai Z. (1982)
Agric. Biol. Chem. **46** : 321.

21) Tanaka K, Sugimoto T., Ogawa M. and Kasai Z. (1980)
Agric. Biol. Chem. **44** : 1633.

22) Brinegar A.C. and Peterson D.M. (1982)
Arch. Biochem. Biophys. **219** : 71.

23) Tumer N.E., Thann V.H. and Nielsen N.C. (1981)
J. Biol. Chem. **256** : 8756.

24) Croy R.D., Gatehouse J.A., Evan J.M. and Boulter D. (1980)
Planta. **148** : 49.

25) Yamagata H., Sugimoto T., Tanaka K. and Kasai Z. (1982)
Plant Physiol. **70** : 1094.

26) Barton K.A., Thompson J.F., Madison J.T., Rosenthal R., Jarvis N.P. and Beachy R.N. (1982)
J. Biol. Chem. **257** : 6089 -6094.

27) Rossi H. and Luthe D.S. (1983)
Plant Physiol. **72** : 578.

28) Masumura T., Hibino T., Kidzu K., Mitsukawa K., Tanaka K. and Fujii S. (1990)
Mol. Gen. Genet. **221** : 1.

29) Rubenstein I. (1982)
In `Maize for Biological Research' ed. CWP Sheridon : Plant Mol. Biol. Assoc., Charlotteville, pp 189.

30) Okita T.W. Hwang Y.S., Hnilo J, Kim W.T., Aryan A.P., Larson R. and Krishnan H.B. (1989)
J. Biol. Chem. **264** : 12573 - 12581.

31) Takaiwa F., Oono K, Wing P. and Kato A. (1991)
Plant Mol. Biol. **17** : 875 - 885.

32) Chen Z.L., Schuler M.A. and Beachy R.N. (1986)

P.N.A.S. USA. **83** : 8560 - 8564.

33) Rafalski J.A. (1986)
Gene **43** : 221 - 229.

34) Spena A., Viotti A., and Pirrotta T. (1983)
J. Mol. Biol. **169** : 799 - 811.

35) Lutcke H.A., Chow K.C., Mickel F.S., Moss K.A., Kern H.F., Scheele G.A. (1987)
EMBO J **6** : 43 - 48.

36) Takaiwa F. and Oono K. (1990)
Mol. Gen. Genet. **224** : 289 - 293.

37) Takaiwa F., Oono K. and Kato A. (1991)
Plant Mol. Biol. **16** : 49 - 58.

38) Baumlein H., Wobus U., Pustell J. and Kafatos F.C. (1986)
Nucl. Acid Res. **14** : 2707 - 2720.

39) Curran T. and Fraza B.R. Jr. (1988)
Cell **53** : 395 - 397.

40) Forde B.G., Heyworth A., Pywell J. and Kreis M. (1985)
Nucl. Acid Res. **13** : 7327 - 7339.

41) Fulford N.G., Forde J., Anderson O.D., Greene F.C. and Shewry P.R. (1987)
Theor. Appl. Genet. **75** : 117 - 126.

42) Okita T.W., Krishnan H.B. and Kim W.T. (1988)
Plant Sci. **57** : 103 - 111.

43) Higuchi W. and Fukazawa C. (1987)

Gene. **55** : 245 - 253.

44) Takaiwa F. Kikuchi S. and Oono K. (1986)
FEBS Lett. **206** : 33 - 35

45) Shotwell M.A. and Larkins B.A. (1989)
In. Biochemistry of Plants : A Comprehensive Treatise. (Eds. Marcus A. Vol 15 pp 296 - 345. Academic Press Orlando, FL)

46) Okita T.W., Hwang Y.S., Hnilo J., Kim W.T., Aryan A.P., Larson R. and Krishnan H.B. (1989)
J. Biol. Chem. **264** : 12573 - 12581.

47) Shotwell M.A., Afonso C., Davies E., Chesnut R.R. and Larkins B.A. (1988)
Plant Physiol. **87** : 698 - 704.

48) Reeves C.D. and Okita T.W. (1987)
Gene (Amst.) **52** : 257 - 266.

49) Kim W.T. and Okita T.W. (1988)
FEBS Lett. **231** : 308 - 310.

50) Pederson K., Argos P., Naravana S.V.L., Larkins B.A. (1986)
J. Biol. Chem. **261** : 6271 - 6286.

51) Prat S., Cortadas J., Puigdomenech P. and Palau J. (1985)
Nucl. Acid Res. **13** : 1493 - 1504.

52) Forde B.G., Heyworth A., Pywell J., and Kreis M. (1985)
Nucl. Acid Res. **13** : 7327 - 7339.

53) Mills V.C. and Brandt A. (1988)
Plant Mol. Biol. **11** : 449 - 461.

54) Leisy D.J., Hnilo J., Zhao O.Y. and Okita T.W. (1989)
Plant Mol. Biol. **14** : 41 - 50.

55) Okita T.W., Zhang Y, Roberts M., and Morris J
Abstract presented at the 6th Ann. Meeting of the Intl. Program on Rice Biotech. Feb 1 - 5, 1993 Chiang Mai, Thailand, p118.

56) Salerno G.L., Pontis H.G., Valdez J.J and Luis H-e
Abstract presented at the 6th Ann. Meeting of the Intl. Program on Rice Biotech. Feb 1 - 5, 1993 Chiang Mai, Thailand, p 85.

57) Tanaka K
Abstract presented at the 6th Ann. Meeting of the Intl. Program on Rice Biotech. Feb 1 - 5, Chiang Mai, Thailand. p 84.

58) Reddy A.R.
Abstract presented at the 6th Ann. Meeting of the Intl. Program on Rice Biotech. Feb 1 - 5, 1993 Chiang Mai, Thailand, p89-90.

59) Wang Z-y., Snustad P., Li M-g., Gao J-p., Zheng F-g., Zhang J-l., and Hong M-n.,
Abstract presented at the 6th Ann. Meeting of the Intl. Program on Rice Biotech. Feb 1 - 5, 1993 Chiang Mai, Thailand, p88.

60) Romero G.O., Chandler J., and Rodriguez R.L
Poster presented at the 6th Ann. Meeting of the Intl. Program on Rice Biotech. Feb 1 - 5, 1993 Chiang Mai, Thailand, p172.

61) Vignols F., Lund G., Phami S., Tremousaygne D., Kader J-c., Puigdomenech P., and Delseny M.
Abstract presented at the 6th Ann. Meeting of the Intl. Program on Rice Biotech. Feb 1 - 5, 1993 Chiang Mai, Thailand, p158.

62) Alvarez A.M., Benedict A.A., and Gnanamanickam S.S.
Poster presented at the 6th Ann. Meeting of the Intl. Program on Rice Biotech. Feb 1 - 5, 1993 Chiang Mai, Thailand,

CHAPTER II

CONSTRUCTION OF A RICE GENOMIC LIBRARY: IDENTIFICATION AND ISOLATION OF PUTATIVE PROLAMIN CLONES HOMOLOGUS TO KAFIRIN

ABSTRACT

Construction of a genomic library of rice in a lambda substitution vector namely, Charon4A, is described. The high molecular weight rice genomic DNA was partially digested with EcoRI and purified on a sucrose density gradient. The DNA fragments in the range of 10 - 20 kbp were pooled together and used for ligation with lambda Charon4A arms which were also obtained by digestion with EcoRI. The development of an efficient in vitro packaging system enabled the construction of a genomic library that represented 7 - 8 rice genomes. This library was then screened for putative clones that were homologous to kafirin, a prolamin gene from S. bicolor. Three putative regions were identified and purified by further rounds of plaque lifting and screening. They were isolated on a large scale and confirmed to contain inserts that were homologous to kafirin.

2.1. INTRODUCTION

Prolamins are alcohol soluble proteins predominantly found in the seeds of cereals. Rice is a rather exceptional crop where glutelin type of storage proteins are abundant (1). The prolamin content in rice varies with the extraction procedures and is actually higher than the amounts reported. Although a few rice cDNA clones encoding prolamin polypeptides have been isolated and sequenced (2,3), very little is known about the organization and expression of these genes in rice. Moreover, in spite of the high copy number of the prolamin genes in rice (4), it has a very low expression level. Since there is no information available on prolamin genes from indica rice, my initial objective was to clone the prolamin genes from the genomic library of an indica rice. Construction of a genomic library of rice and screening of this library using a heterologous cDNA probe to identify putative prolamin encoding genes from rice was, therefore, a first step in this direction.

2.2. MATERIALS AND METHODS

The construction of a rice genomic library involves the following steps :

1. Isolation of DNA from rice and vector i.e.Charon 4A.
2. Preparation of packaging extract from E.coli SMR-10

3. Annealing and ligation of vector arms and rice DNA.
4. In vitro packaging and plating of packaged phages;
5. Screening of the library using a heterologous probe.

These steps are described in detail subsequently.

2.2.1. Isolation of high molecular weight rice DNA

The aim of any genomic DNA preparation is to isolate DNA of high molecular weight and of sufficient purity. The major problem during plant DNA extractions are the presence of nucleases and shearing of isolated DNA. Nucleases can be tackled by freezing the tissue quickly and thawing it in the presence of extraction buffer containing detergent and a high concentration of EDTA. Some plant tissues have a major problem of contaminating polysaccharides which affect the DNA purity. These carbohydrates are difficult to separate from the DNA itself and inhibit many enzymes commonly used in cloning procedures. The polysaccharides are removed by performing a chloroform extraction of lysates in the presence of 1 % CTAB (cetyltrimethylammoniumbromide) and 0.7 M NaCl (5). CTAB forms a complex with the proteins and carbohydrates facilitating their removal on addition of chloroform. The general protocol is modified according to the tissue that is to be used for the isolation of high molecular weight DNA.

Seeds of rice (Oryza sativa cv. Basmati-370), obtained from National Seeds Corporation, Pune, were thoroughly washed with distilled water and 50% alcohol and soaked overnight in distilled water. They were grown on wet cotton in the dark at RT for 8-10 days. The etiolated seedlings were harvested and frozen in liquid N_2 and stored at -70°C. This tissue was used for the isolation of high molecular rice DNA essentially according to the procedures of Shure et al (6).

(1) The frozen etiolated seedling tissue was ground in a mortar to a fine powder under liquid N_2.

(2) The powder was suspended in urea-phenol extraction buffer (8.0M urea, 0.35M NaCl, 0.05M Tris-HCl, pH 7.5 and 0.02M EDTA) containing fresh 2% Sarkosyl and 5% Tris-saturated phenol.

(3) The suspension was incubated at 65°C for 15 minutes and then allowed to cool down to RT. This step facilitates the lysis of cells and denatures proteins.

(4) An equal volume of phenol : ($CHCl_3$: IAA) (in a ratio of 3:1) was added to the suspension.

(5) The mixture was centrifuged at 10 K rpm for 10 minutes at 4°C in a Sorval RC 5B centrifuge with a SS34 rotor for deproteinisation of the aqueous phase.

(6) The aqueous phase was separated and mixed with an equal volume of $CHCl_3$: IAA (24:1 v/v ratio).

(7) The mixture was centrifuged at 10 K rpm for 10 minutes at 4°C.

(8) The steps 6 and 7 were repeated twice till there was no interphase between the aqueous and the organic phase to ensure complete deproteinisation.

(9) To the aqueous phase, $1/10^{th}$ volume of 3M Na acetate pH 4.5 and 2 volumes of chilled ethanol was added and the solution was incubated at -20°C overnight.

(10) The solution was centrifuged at 10 K rpm for 10 minutes at 4°C.

(11) The supernatant was decanted and the nucleic acid pellet was washed with 70% ethanol.

(12) The pellet was then dried and suspended in a minimum volume of TE (10mM Tris-HCl and 1mM EDTA pH 8.0).

(13) In order to remove RNA, RNaseA was added to a final concentration of 100 μg/ml and the solution incubated at 37°C for 1 hour.

(14) The RNA free DNA was deproteinized by $CHCl_3$: IAA and reprecipitated by addition of 2 volumes of chilled ethanol.

(15) The DNA was pelleted, dried and redissolved in TE buffer as mentioned earlier.

(16) The quality of DNA preparation was checked by agarose gel electrophoresis while its concentration was measured in a Shimadzu double beam spectrophotometer model UV-210A using the extinction coefficient E = 200.

2.2.2. Preparation of 10 - 20 kbp Fragments of Rice DNA

In order to achieve a complete representation of rice DNA in its genomic library, high molecular weight rice DNA (> 50 kbp) has to be digested into smaller fragments that can be inserted into suitable vectors. Restriction endonucleases are used for such digestions to obtain fragments that have ligatable ends. A complete digestion of the genomic DNA might split the DNA region of interest. To reduce this possibility, rice DNA is partially digested to obtain fragments of sizes that can be accommodated in the vector of choice. The overall strategy, therefore, involves the optimisation of partial digestion conditions, checking of digestion patterns and separation of fragments on a sucrose density gradient.

(a) Partial Digestion of Rice DNA

Since the cloning capacity of Charon 4A vector ranges between 10 to 20 kbp fragments, the rice DNA was partially digested with EcoRI to generate

fragments in this range (7). The protocol followed for setting up partial digestion of DNA was as follows :

(1) 10 μg of genomic DNA was mixed with the EcoRI buffer (10 x) and the volume made up to 100 μl.

(2) 20 μl was dispensed in the 1st vial and 10 μl in the remaining vials and the vials were kept on ice.

(3) 20 units of EcoRI was added to the 1st vial and mixed thoroughly.

(4) The tube was microfuged and 10 μl of the sample was transferred to the next vial.

(5) The above serial transfer was carried out till the 8th vial.

(6) The 9th vial was left as control.

(7) All the vials were incubated at 37°C for 1 hour.

(8) The reaction was terminated by the addition of termination buffer (50 % glycerol, 100 mM EDTA, 0.2 % Bromophenol Blue).

(9) The samples were electrophoresed on a 1% agarose gel along with standard molecular weight markers.

(b) Agarose gel Electrophoresis

All the DNA samples were analysed on 0.8% neutral agarose gels in TAE (40mM Tris-HCl, 20mM Na acetate, 2mM EDTA, pH 8.1) at a constant current of 10mA. After electrophoresis, the agarose gels were

stained with ethidium bromide (100 μg/ml) and visualised on a UV transilluminator and photographed with a 35mm SLR camera (Minolta 700) using a red filter and ORWO 125 ASA, black and white film. Samples showing appropriate digestion pattern were then pooled together, dialysed and size fractionated on a sucrose density gradient.

(c) Size fractionation of rice DNA

The partially digested genomic DNA must be size fractionated before ligation with the vector DNA to remove irrelevant small and large fragments (8). If this step is not performed, the small fragments will ligate together and produce recombinants that are difficult to analyse. Large insert DNA fragments will ligate to the vector DNA and will prevent efficient packaging of the recombinant molecule into the phage. Size fractionation can be done by two methods, sucrose density gradient and preparative gel electrophoresis. Sucrose density fractionation is generally faster than preparative gel electrophoresis. But the latter procedure has a better capacity for resolving large amounts of DNA and has significantly better resolution. In spite of its lower resolving power than preparative gel electrophoresis, sucrose density gradient is the most widely used method fro producing inserts for

genomic libraries. The partially digested DNA fragments migrate through a linear sucrose density gradient at a rate that is dependent on their size. Care should be taken so that the gradient is not overloaded. A maximum of 0.5 mg can be loaded on a 38-ml sucrose gradient or 0.2 mg on a 12-ml gradient. The distance that the DNA migrates is approximately proportional to the time of centrifugation and rpm^2. The gradients are relatively stable, but must be handled with care in order to avoid any sudden jerks. They should not be stored for long periods before or after use. Sucrose gradients can also be used for the purification of bacteriophage lambda arms.

The partially digested and pooled rice DNA was size fractionated on a sucrose density gradient to obtain 10-20 kbp fragments.

1. Sucrose solutions of 10% and 40% were carefully layered in a 12.5ml ultracentrifuge tube and placed horizontally on a table for 3 hours to allow the gradient formation.
2. 200 µl (100 µg) of DNA sample was carefully layered onto the gradients. The gradient was centrifuged at 26,000 rpm for 16 hours at 20°C.
3. After centrifugation, 250 µl fractions were collected from the bottom of the tube.
4. 15 ul of every third fraction was mixed with 15 µl

sterile distilled water, ethanol precipitated, dissolved in 10 ul TE and electrophoresed on 0.5% agarose gels along with standard markers.

5. The fractions containing the 10-20 kbp size range were pooled, dialysed against TE at 4°C for 16 hrs with a buffer change after 6 hrs.
6. The DNA was then precipitated with 2 volumes of ethanol after adjusting the Na acetate concentration to 0.3 M, washed with 70 % ethanol and dried and the pellet was resuspended in TE at a concentration of 1 μg/ul.

2.2.3. Isolation of Charon 4A vector DNA and separation of its arms

Since the first demonstration of the feasibility of using bacteriophage lambda as a cloning vehicle (9), a large variety of lambda vectors have been constructed (10). About the middle third of the lambda genome is dispensable for lytic growth. Derivatives of lambda that are used as cloning vectors typically contain restriction sites that flank some or all of these dispensable genes. The major advantage of using lambda derived cloning vectors is that DNA can be inserted and packaged into phage in vitro. Lambda DNA cannot be packaged into the phage head if its genome is less than 78 % or more than 105 % of the length of wild

type lambda DNA. Thus the left and right arms of the phage are separated and ligated to foreign DNA to form concatamers which are then packaged into phage heads by cleavage at the cos sites. Blattner et al. (11) have constructed a series of bacteriophage lambda vectors that can be used for cloning DNA fragments. This series of constructed bacteriophage vectors were labelled as the Charon phages, after the mythological boatman on the river Styx. The initial Charon vectors had only one or two cloning sites, for example, Charon 4A with EcoRI and XbaI sites. These vectors were improved upon subsequently and resulted in the latest Charon vector, namely Charon 40 with 16 restriction enzyme sites (12). The separation of the Charon 40 arms is also simple. These vectors are used as replacement vectors where the DNA between the restriction sites are replaced by the foreign DNA. The other bacteriophage lambda vectors that can also be used as cloning vectors are lambdaEMBL3 (13), lambda gt10, lambda gt11 (14) and lambda ZAP (Stratagene cloning Systems).

The bacteriophage lambda Charon 4A was grown on large scale using E.coli Le392 as a host. The DNA was isolated from the lysate and digested to completion with EcoRI to obtain the left and the right arms. The arms were purified on a sucrose density gradient before

its use for ligation with foreign DNA.

(a) Preparation of Bacteriophage Charon 4A Lysate

E.coli LE392 was used as a host for Charon 4A phages.

1. A fresh bacterial colony was inoculated into 5.0 ml of sterile NZY broth containing 0.2 % maltose and incubated at 37°C overnight at 180 rpm.
2. The cell density of the suspension was determined at $O.D_{600}$ (1.0 O.D = 8 x 10^8 cells/ml).
3. 1.0 ml of the cell suspension was mixed with 10 mM CaMg solution. The lysate of bacteriophage Charon 4A containing 8 X 10^7 PFU/ml was added to the suspension. The divalent cations Ca^{++} and Mg^{++} facilitates the adsorption of the phages onto the cells
4. The mixture was incubated at 37°C for 20 min and transferred to a 500 ml flask containing 100 ml of sterile NZY broth.
5. The flasks were incubated at 37°C on a shaker at 180 rpm till complete lysis.
6. 1 % $CHCl_3$ was added to the flask which was further incubated for 20 min at 37°C to ensure complete lysis.
7. The lysates were centrifuged at 6000 rpm in a Sorvall GSA rotor at 4°C for 20 min to pellet the

cell debris.

8. The supernatant was collected and used for bacteriophage DNA extraction.

(b) Extraction of Charon 4A DNA

1. The lysate obtained was incubated in the presence of DNaseI (1 μg/ml lysate) and RNaseA (5 μg/ml lysate) at 37°C for 1 hour to remove chromosomal DNA and RNA.
2. The phage precipitation was done by the addition of solid NaCl to a final concentration of 1 M and 7.5 % polyethylene glycol 6000 (PEG) and the lysate was incubated at 4°C overnight.
3. The lysate was centrifuged at 5000 rpm at 4°C for 20 min. in a Sorvall GSA rotor, the supernatant was decanted and the phage pellet was suspended in SM buffer.
4. An equal volume of $CHCl_3$ was added to the suspension and centrifuged at 10,000 rpm at 4°C for 10 min. This step was necessary to remove the traces of PEG from the suspension.
5. To the aqueous phase, an equal volume of phenol:$CHCl_3$:IAA (25:24:1) was added and the mixture centrifuged at 10,000 rpm for 10 min. at 4°C.
6. The aqueous phase was separated and an equal volume

of $CHCl_3$: IAA (24:1) was mixed with it. The mixture was centrifuged at 10,000 rpm for 10 min. at 4°C.

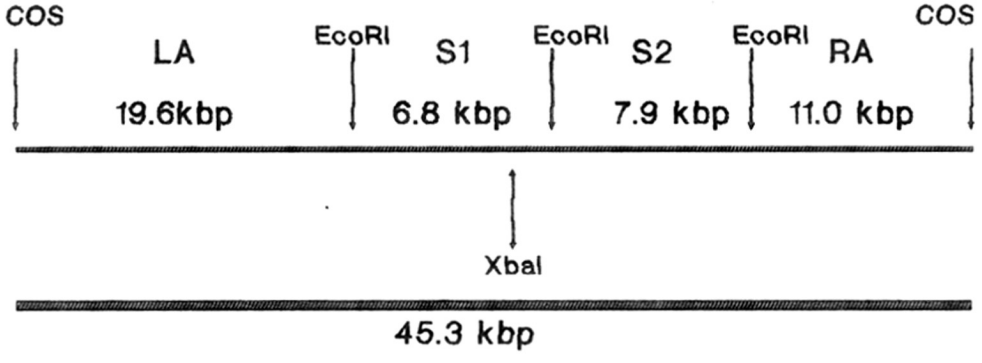
7. An equal volume of 5 M NH_4 acetate and 2 volumes of chilled ethanol were added to the aqueous phase to precipitate the phage DNA.
8. The DNA was pelleted, washed with 70 % ethanol, dried and finally dissolved in $T_{10}E_1$.

(c) Separation of Charon 4A Arms

Bacteriophage Charon 4A was used as a vector for the construction of the rice genomic library. THe linear map of the bacteriophage lambda Charon 4A is depicted in Fig. 2. The total length of the vector DNA is about 45 kbp and can be digested with EcoRI or XbaI. EcoRI digestion of the phage DNA generates two arms of 19.6 kbp and 11.0 kbp and two stuffer fragments of 6.8 and 7.9 kbp. It contains amber mutations in the A and B genes. An EcoRI digestion can accommodate 10 - 20 kbp while the XbaI digestion can take up only 1 - 5 kbp. In order to use this vector DNA, it was necessary to first obtain the adequate amount of the lysate followed by isolation of DNA and separation of its arms.

1. 100 μg of Charon 4A DNA was digested to completion with EcoRI at 37°C for 3 hrs.
2. The digested DNA was heated at 68°C for 10 min for separation of the arms from the stuffer regions

Fig. 2 : Linear map of bacteriophage Charon 4A



LINEAR MAP OF LAMBDA CHARON 4A

LA : LEFT ARM RA : RIGHT ARM S1, S2 : STUFFER REGIONS

Fig 2

3. 10 mM of $MgCl_2$ was added, cooled to 42°C and incubated at the same temperature for 60 min. for the annealing of the cos sites.
4. The annealed DNA was further fractionated by loading it onto a 10-40 % sucrose density gradient as described on page 42.
5. The fractions containing the arms were pooled, dialysed against TE at 4°C overnight, concentrated by ethanol precipitation and finally dissolved in TE buffer.

2.2.4. Preparation of Packaging Extract

Packaging extracts are used to package lambda with high efficiency. It contains all the components necessary to form the complete phage except the concatameric DNA. A mutation in the DNA of lysogenic phage prevents its own DNA from being packaged.

There are two methods by which the components necessary for packaging of phage DNA can be prepared. One is a double culture method where the two cultures have complementary defects in the lambda packaging protein machinery. On addition of the two lysogens, phage DNA is packaged due to the full complementation of packaging proteins (15). The other method is a single culture method, where a single strain coding for all the packaging proteins is used. This extract is

unable to package endogenous phage DNA because of a cos2 mutation in the packaging origin (16).

1. A colony of E. coli SMR-10 was inoculated into sterile 50ml LB broth and incubated at 34°C on a shaker at 180 rpm.
2. Approximately 6.0ml of this overnight broth was added to 115ml of sterile LB broth prewarmed at 34°C in a 2 liter flask to grow the culture in large amounts.
3. The flask was incubated at 34°C on a shaker at 300 rpm till an $O.D._{600}$ of 0.8 was reached.
4. Since the culture is lysogenic, the flask was warmed upto 45°C for 15 min with vigorous shaking in order to induce the culture into lytic cycle.
5. The flask was then incubated at 37°C for 90 min. with vigorous agitation for complete lysis.
6. The culture was quickly chilled by swirling the flask in ice water for 5 min. and then transferred to 2 precooled 250ml centrifuge tubes and centrifuged at 4000g for 6 min. at 4°C in a precooled rotor to collect the pellet. Precooling to 4°C is critical in maintaining the viability of packaging extract.
7. For stabilisation of packaging proteins, 4.5ml of precooled TSP (40 mM Tris-HCl pH 7.9, 10 mM spermidine and 10 mM putrescine) was added to

pellet and centrifuged at 4000g for 6 min. at 4°C.

8. The pellet containing packaging proteins was suspended in 350 µl of precooled TSP. 20 µl was aliquoted into tubes containing 5 µl DMSO + ATP mixture (50 % DMSO + 7.5mM ATP, pH 7.0) and frozen in liquid N_2 and stored at -70°C. This packaging extract was used in the construction of the genomic library.

2.2.5. Ligation, In vitro Packaging and Plating of Packaged Phages.

Partially digested genomic DNA and vector arms have to be ligated in order to generate fragment lengths that can be accommodated in the phage head. This necessitates optimisation of ligation conditions and ratio of insert DNA to vector DNA prior to its addition to the packaging extract. The quality of insert and vector DNA are critical for a successful ligation. Ligated DNA is then introduced into E.coli by packaging the DNA into bacteriophage particles which are then allowed to infect the bacteria. Complete phage particles form well-defined plaques which can be transferred onto filter membranes in order to screen for desired clones or further amplification.

1. EcoRI digested and purified Charon4A arms and rice DNA fragments in the range of 10.20 kbp were mixed

in the ratios of 2:1, 1:1 and 0.5:1 (Charon4A arms: rice DNA) to determine optimum vector : insert DNA ratio.

2. 900 ng of total DNA was mixed with $1/10^{th}$ volume of 10 X ligation buffer (500mM Tris-HCl, pH 8.0, 70mM $MgCl_2$, 10mM DTT, 10mM ATP) and 10 units of T_4 DNA ligase. The mixture was incubated at 14°C for 20 hours for ligation of insert DNA to vector arms.
3. Ligatibility of Charon 4A arms and 10-20 kbp rice DNA was checked by 0.5% agarose gel electrophoresis.
4. The ligated DNA was stored at 4°C till it was used for packaging.
5. The packaging extract was removed from -70°C and placed on ice.
6. 10 ul ligated DNA dissolved in TEK (100mM Tris-HCl, pH 8.1, 5mM KCl) was added to the tube and mixed with a thin glass rod and the tube was incubated at 25°C for 90 minutes to allow for packaging.
7. 500 µl SMCK buffer (SMCK A : 8ml + SMCK B : 2ml + 190ul sterile d/w) containing 50 µg/ml DNaseI was added to the tube and incubated at RT for 5 minutes to remove unpackaged phage DNA.

SMCK A : Na_2HPO_4 1.75g, KH_2PO_4 0.75g, d/w - 10ml.

SMCK B : NaCl 0.5g, NH_4Cl 1.0g, 1M $MgCl_2$, 1M $CaCl_2$.

8. 100 μl of chloroform was added to the tubes, vortexed and the aqueous phase was used for infection.
9. The aqueous phase containing the packaged phage was serial diluted in SM buffer and incubated with overnight grown plating bacteria and 10 mM CaMg.
10. The mixture was incubated at 37°C for 20 minutes for adsorption of the phage onto the bacterial cells.
11. The mixture was mixed with 0.7% molten NZY top agar at 45°C containing IPTG and X-gal. The plates were incubated overnight at 37°C.
12. The number of white plaques were counted to determine the number of recombinant phages.

2.2.6. Amplification of the genomic library

In order to get an overrepresentation of all DNA fragments in the genome, the library of recombinant bacteriophages was amplified by growing a plate stock from the packaged mixture. The principle of amplification is to allow each of the in vitro packaged phage to produce thousands of identical clones. The amplification also allows the library to be stored nearly indefinitely and to be screened as many times as necessary. The amplification procedure was as follows :

1. A colony of E.coli LE 392 was grown overnight at 37°C at 180 rpm in a sterile 5ml NZY broth containing 0.2% Maltose.
2. The packaging mixture was mixed with bacterial cell suspension in presence of 10mM CaMg and incubated at 37°C for 20 minutes for phage adsorption.
3. The infection mixture was mixed with sterile NZY medium containing 0.7% agar kept molten at 45°C and poured onto sterile NZY agar plates. The plates were incubated at 37°C overnight.
4. 30ml of sterile SM buffer was added to the plates and incubated overnight at 4°C on a slow shaker to allow diffusion of phages into the buffer.
5. The bacteriophage suspension was recovered from the plate which was then washed with 1.0ml SM buffer.
6. A few drops of chloroform was added to the suspension and centrifuged.
7. The lysate which contains the amplified library was stored at 4°C with 0.3% $CHCl_3$ without reduction in the titre.

2.2.7. Screening of rice genomic library with a heterologous probe

The usual approach to isolating a recombinant DNA clone encoding a particular gene or mRNA sequence is to screen a recombinant DNA library. The desired clone is identified because it hybridises to a nucleic acid probe or it expresses a segment of protein that can be recognised by an antibody. The basic principle of screening recombinant DNA libraries is that bacteriophage plaques contain large amounts of insert DNA that can be detected either directly by hybridisation or indirectly by the protein that may be expressed from the cloned segment. The first step in the nucleic acid hybridisation procedure is to grow large number of plaques on agar plates. These are transferred onto nitrocellulose filters, where they can be screened. Subsequent treatment of the filter with NaOH destroys the phage particles and denatures the phage DNA which then binds to the nitrocellulose. Neutralisation of filters is required to maintain the integrity of nitrocellulose.The desired clone can be detected by its ability to hybridise to a DNA probe. Hybridisation of these filters to DNA or RNA probe will identify the location of the phage plaque of interest, which can then be recovered from the plate. This is a

rapid and effective screening procedure that allows the identification of a single clone within a population of million other clones. The filters are hybridised with ^{32}P-labelled nucleic acid probe, the excess and incorrectly matched probe is washed off and the filter is autoradiographed. The specificity of probe-DNA interaction can be determined by varying the temperature and salt concentration. A highly homologous probe will remain bound to the immobilised DNA even under stringent conditions.

The rice genomic library represents all DNA sequences that are present in the rice genome. It was screened using a heterologous probe from sorghum to identify corresponding prolamin type of genes in rice. This involved the following steps:

(a) Plaque lifting

For screening the library for recombinant phages, plaques were transferred onto the membranes for hybridization with probes. The plaque blotting was carried out as described by Benton and Davis (17).

1. The plates containing plaques were precooled at 4°C for 1.5 to 2 hours.
2. NEN's (DUPONT) plaque screens were placed on plates for 4-5 minutes for transfer of plaques onto the screens.

3. The membranes were gently lifted and floated first on denaturation solution (0.5M NaOH, 1.5M NaCl) for 3 minutes and on neutralisation solution (1.0M Tris-HCl, pH 7.5, 1.5M NaCl) for 3 minutes.
4. The membranes were air dried and stored till further use.

(b) Labelling of DNA

The probes were labelled with α-P-^{32}-dCTP by a random priming method of Feinberg and Vogelstein (18).

1. 100 ng of probe DNA in a suitable volume of TE was denatured by boiling for 10 minutes and immediately chilled on ice.
2. To the denatured DNA : 5 μl primer, 5 μl labelling buffer, 4 μl of 10 X dCTP, 10 X dGTP 10 X dTTP and 4 μl of 10 X α-P-^{32}-dCTP and 2 μl Klenow fragment were serially added.
3. The final volume was made up to 50 μl with sterile water and the reaction was carried out at 37°C for 60 minutes.
4. The reaction was terminated by adding stop dye (10% SDS, 125mM EDTA, 10mg/ml Bromophenol Blue). Unincorporated labelled nucleotides were removed by spin column chromatography (7).
5. The radioactivity of each fraction was determined separately on a Rack-β liquid scintillation

counter. Fractions containing labelled probe were pooled prior to their use in hybridization experiments. Specific activity of the samples ranged between 1-5 x 10^8 cpm/μg DNA.

(c) Hybridization and autoradiography

1. The plaque screens were prehybridized for 3-4 hours at 62°C in plastic bags in a solution containing 5X SSC, 0.1% SDS, 5X Denhardt's solution (1% Ficoll, 1% Polyvinyl pyrrolidone, 1% BSA, 100 μg/ml salmon sperm DNA and 5mM Na pyrophosphate).
2. The purified probe was added to the hybridization solution, which was the same as prehybridisation solution, boiled for 10 min. and then added to the blots.
3. The hybridizations were carried out O/N with gentle shaking at 62°C.
4. The solution was discarded and blots were washed at 65°C with 2X SSC and 0.5% SDS.
5. The filters were exposed in X-ray films (Kodak) for 12 hours to 4 days at -70°C using Kiran's superfast intensifying screens in X-ray cassettes.

2.2.10. Confirmation of recombinant clones

The putative clones showing hybridisation with the

probe served as the starting material for further work. To confirm the presence of prolamin type genes homologous to kafirin probe, the DNA was digested with EcoRI, electrophoresed and blotted for further hybridisation with labelled probe. The large scale DNA isolation of recombinant clones were performed as on page 47. The DNA was digested to completion with EcoRI using 3-4 µgs of DNA and 20 units of restriction enzyme in a 20 µl reaction volume at 37°C for an appropriate time. Care was taken to adjust the ratio of enzyme to DNA so that non specific (star) activity of the restriction enzyme was avoided. The reaction was stopped by addition of a mixture of 50% glycerol, 100mM EDTA, 0.2% Bromophenol blue. The digested recombinant clones were electrophoresed as described on page 41.

The transfer of DNA from agarose gels to Hybond N/Hybond C extra (Amersham) or Nytran Filter was carried out essentially according to Southern (19) and Maniatis (7).

1. Gels were soaked in 0.25N HCl at RT for 10 minutes for depurination of DNA.
2. The gels were then transferred to a solution containing 1.5M NaCl and 0.5N NaOH and incubated at RT with a gentle shaking for 1 hour with a change of solution after 30 minutes.
3. The gels were neutralised by soaking 1M Tris-HCl,

pH 7.4 and 1.5M NaCl for 60 minutes at RT with gentle shaking and a change of solution after 30 minutes.

4. The gels were placed onto the membrane and transfer was allowed to take place via capillary action for 12-15 hours in 20 X SSC.
5. The membranes were rinsed in 5X SSC, air dried and baked for 2 hours at 80°C under vacuum.
6. The filters were stored at RT till further hybridisation with specific clones as described on page 58.

2.3 RESULTS

(a) Digestion of rice DNA with EcoRI, HindIII and BamHI

For the construction of a rice genomic library, high molecular weight rice DNA was isolated from seedling tissues. About 3 - 4 mg of DNA per 100 g of seedling tissue was obtained. The digestibility of DNA was checked with EcoRI, HindIII and BamHI as shown in Fig. 2.1. The quality of isolated DNA is seen in Fig. 2.1 (lane C) where undigested rice DNA has been loaded. Some amount of undissolved DNA is observed as seen by fluorescence in the well. The digestion pattern with EcoRI, HindIII and BamHI respectively can be seen in lanes D, E and F of Fig.. 2.1. A smear ranging from 20 kbp to 2 kbp, observed in all the three lanes (lanes D,

Fig 2.1 Electrophoresis of high molecular weight rice DNA and digestion with different enzymes

Lane a : undigested lambda DNA
Lane b : Lambda HindIII digest as MW marker
Lane c : Rice (Basmati-370) DNA
Lane d : EcoRI digest
Lane e : HindIII digest
Lane F : BamHI digest

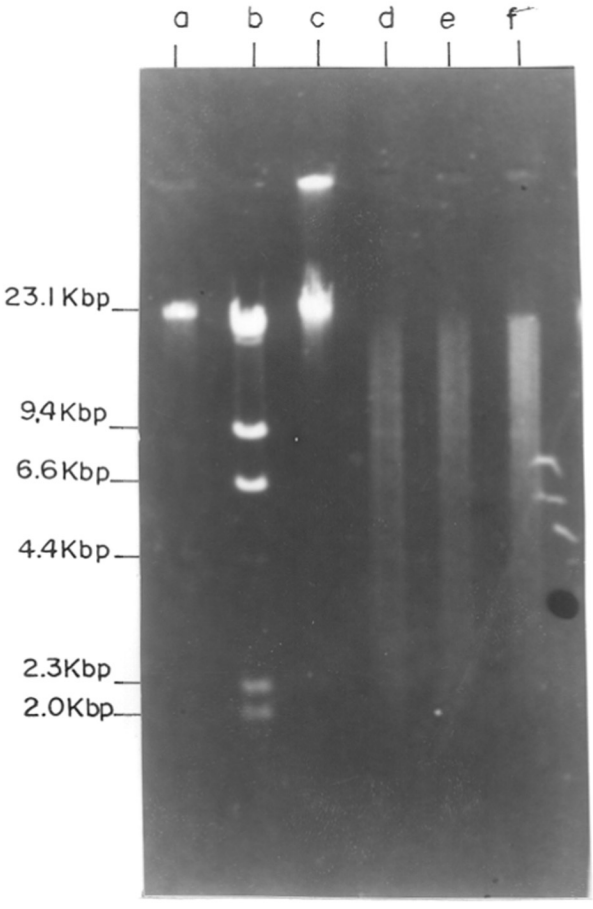


FIG. 2.1

E and F), indicates complete digestion by these enzymes. In fact, all lanes exhibit a band of about 9.0 kbp superimposed on the smear indicative of a repetitive fragment present in rice DNA. The DNA appears to contain a higher number of EcoRI and HindIII sites than BamHI as seen by a smear intensity at a lower range in lanes D and E than in lane F.

(b) Partial digestion of rice DNA with EcoRI

Since Charon 4A DNA has a site only for EcoRI, rice DNA was digested with the same enzyme. If the gene of interest has an internal site for EcoRI, a complete digestion of rice DNA will not prove useful in isolating intact genes. Hence, partial digestion conditions were setup in order to generate fragments of different lengths. Some of the fragments will contain the entire gene providing a better probability of isolating an intact gene with its flanking regions. Since the vector can take upto 10 - 20 kbp fragments, rice DNA was partially digested to generate fragments of such lengths. Figure 2.2 shows an electrophoretic pattern of rice DNA digested with different units of EcoRI. It can be seen that the extent of digestion of the rice DNA decreases with decreasing units of enzyme. Lanes B to F also exhibit the 9.0 kbp repetitive fragment observed earlier in Fig. 2.1 (Lanes D, E and

Fig 2.2 Partial digestion of rice DNA with EcoRI

Lanes a and j : Lambda HindIII digest as MW marker
Lanes b - i : Rice DNA digested with decreasing units of EcoRI

Fig 2.3 Size fractionation of partially digested rice DNA on sucrose density gradient

Lanes a and t : High molecular weight marker DNA
Lanes b - s : Every third aliquot of sucrose density gradient

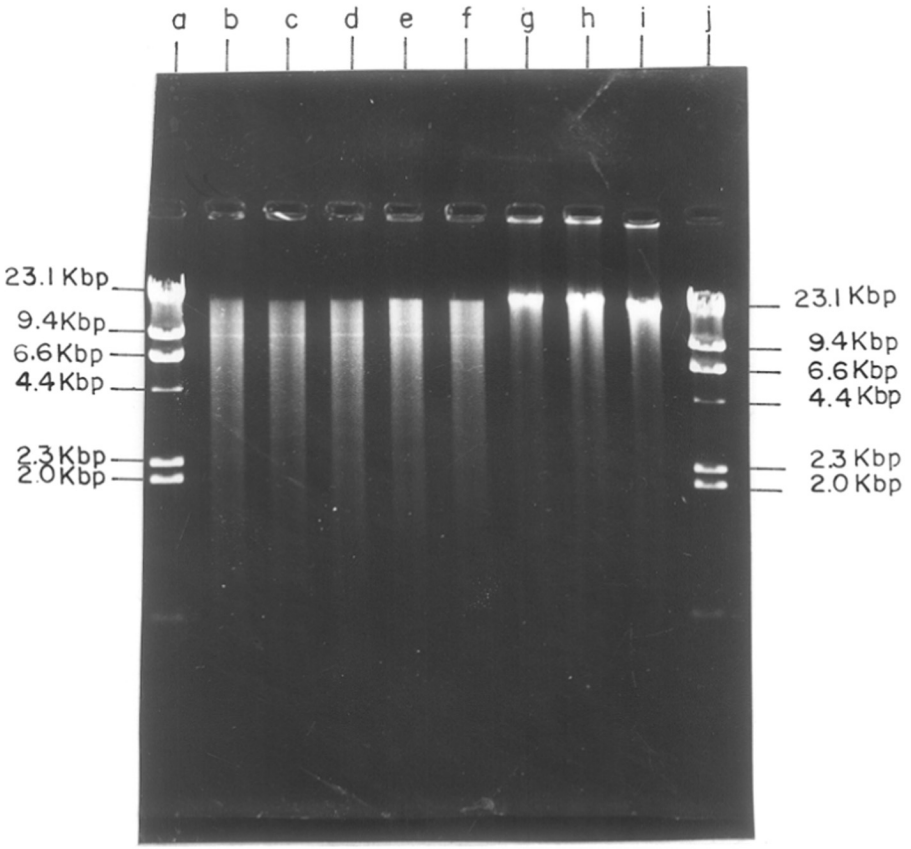

a b c d e f g h i j
23.1 Kbp
9.4Kbp
6.6Kbp
4.4Kbp
2.3Kbp
2.0Kbp
23.1 Kbp
9.4Kbp
6.6Kbp
4.4Kbp
2.3 Kbp
2.0 Kbp

FIG. 2.2

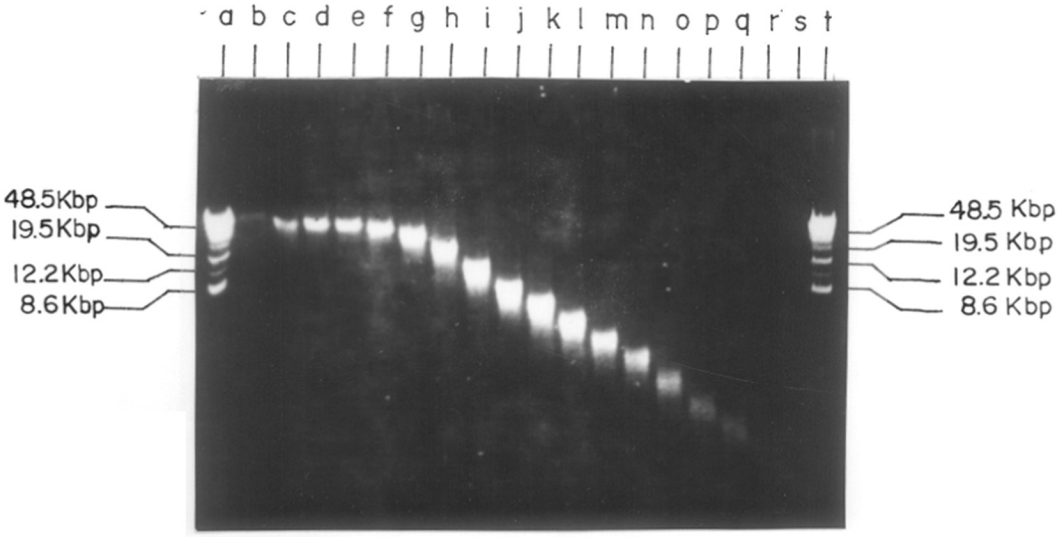

a b c d e f g h i j k l m n o p q r s t
48.5Kbp
19.5Kbp
12.2Kbp
8.6Kbp
48.5 Kbp
19.5 Kbp
12.2 Kbp
8.6 Kbp

FIG. 2.3

F). These lanes also exhibit a smear which steadily reduces from lane B to lane F. It is more or less absent in lanes G, H and I. The digests of lanes B to F were pooled and loaded onto a linear sucrose density gradient for size fractionation of rice DNA.

Figure 2.3 shows the electrophoretic pattern of every third fraction of sucrose density gradient. A gradual decrease in fragment lengths is observed from lane B to lane S. The 10 - 20 kbp fragments are found in fractions 13 - 25 (lanes F to J). Hence these fractions were pooled and checked for ligatibility. Figure 2.4 exhibits 10 - 20 kbp purified rice DNA (lane A), self-ligated rice DNA (lane B) and high molecular weight marker DNA (lane C). Most of rice DNA appears to retain its ability to ligate as observed in Fig. 2.4 (lane B).

(c) Isolation of Charon 4A DNA, ligation and in vitro packaging

Bacteriophage Charon 4A DNA was isolated in a large scale in order to obtain arms for ligating rice DNA. Fig. 2.5 shows the purity of Charon 4A DNA (lane B) present as a single band of about 50 kbp. Complete digestion of purified DNA with EcoRI generates two arms of 19.9 and 11.0 kbp and two stuffer fragments of 7.0 and 6.6 kbp respectively (Lane C). A faint band

Fig 2.4 Ligatibility of cohesive ends of rice DNA

Lane a : 10 - 20 kbp purified size fraction of rice DNA
Lane b : Self ligated purified rice DNA
Lane c : High molecular weight marker DNA

Fig 2.5 Electrophoretic pattern of Charon 4A DNA

Lane a : Undigested lambda DNA
Lane b : Undigested Charon 4A DNA
Lane c : Charon 4A DNA digested with EcoRI
Lane d : Lambda HindIII digest as MW marker

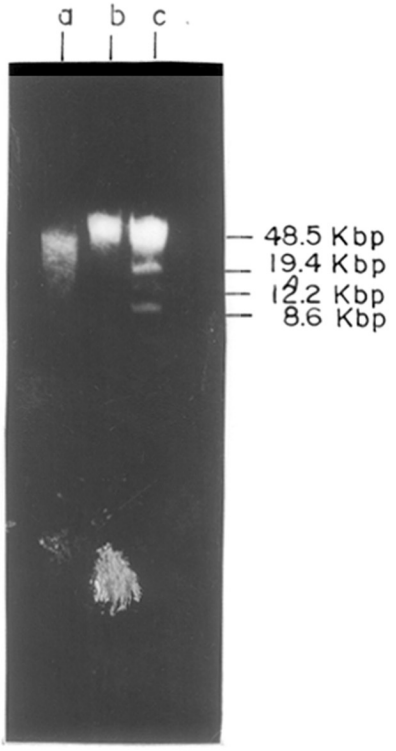

a
b
c
48.5 Kbp
19.4 Kbp
12.2 Kbp
8.6 Kbp

FIG. 2.4

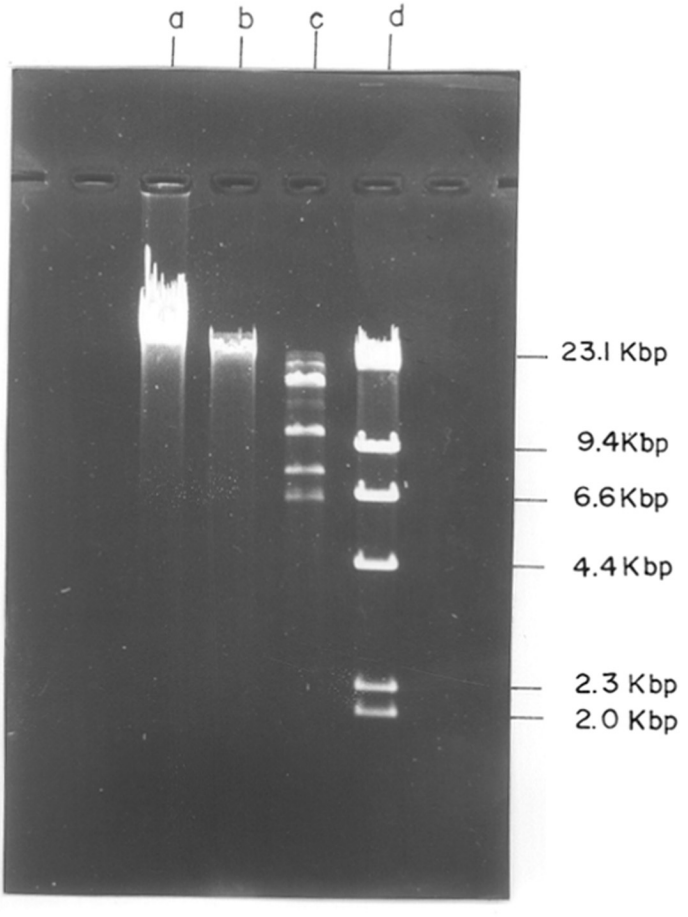

a
b
c
d
23.1 Kbp
9.4Kbp
6.6Kbp
4.4Kbp
2.3 Kbp
2.0 Kbp

FIG. 2.5

observed above the 19.9 kbp band is due to annealing of two arms which can be avoided by heating the sample at 68°C for 10 min. just prior to electrophoresis. Digested Charon 4A DNA was loaded onto a sucrose density gradient in order to separate the arms and stuffer fragments. Fractions were pooled together in order to keep the cos sites intact. Figure 2.6 shows purified arms without the presence of stuffer fragments. The faint upper band observed is again due to annealing of arms.

The vector arms and rice DNA were ligated and packaged into packaging extract. The packaging extract was prepared using a single packaging strain E.coli SMR-10. The efficiency of packaging of preannealed Charon4A DNA was 1 x 10^8 PFU/μg which is similar to that reported for this system (1.0-3.0 x 10^8 PFU/μg of wild lambda DNA). In vitro packaging was carried out using freshly prepared packaging extracts. The packaged phages were plated using E.coli LE392 as a host.

(d) Characterization of rice genomic library

The titre of the rice genomic library was 6.8 x 10^5 PFU/ml. Considering the rice DNA content to be 1 pg, genome representation in the library was calculated as follows

Fig 2.6 Purification of Charon 4A arms

Lane a : Lambda HindIII digest as MW marker
Lane b : Purified Charon 4A arms

Fig 2.7 Autoradiogram of the final round of screening of the genomic library with heterologous probe

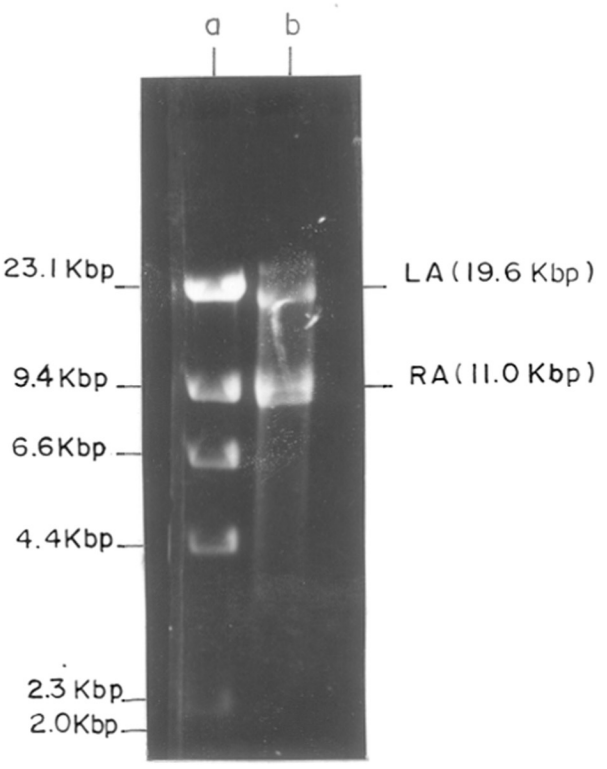


FIG. 2.6

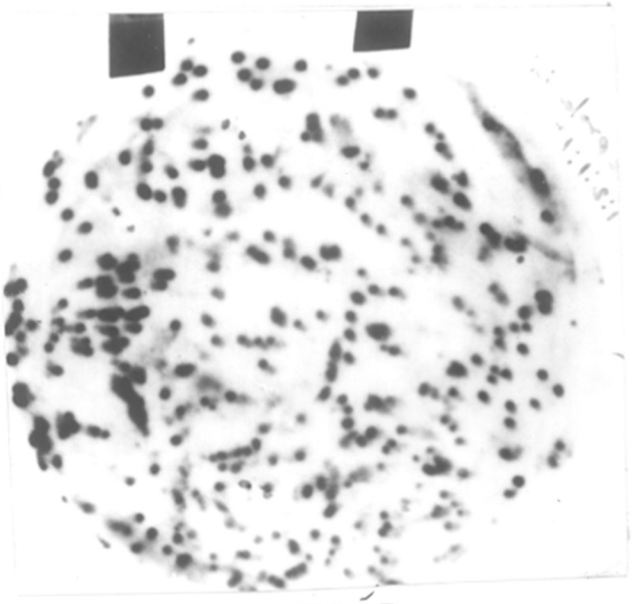


FIG. 2.7

Rice haploid DNA content = 1.0 pg

1.0 pg = 1.096×10^9bp

Average insert size = 15kb (1.5×10^4)

1.5×10^4bp = 1 recombinant phage

1.096×10^9bp = 0.73×10^5 PFU/haploid genome

Observed recombination phage titre = 6.8×10^5 PFU

Observed titre of packaged arms = 1×10^5 PFU

$\therefore$ exact phage titre = 5.8×10^5 PFU

0.73×10^5 = 1 rice genome

5.8×10^5 = 7.95 rice genomes

$\therefore$ the genomic library represents 7.95 rice genomes.

The rice genomic library was amplified in order to get an overrepresentation of low copy clones. Recombinant phages were amplified using E.coli LE392 as host and the titre of amplified library was 1.2×10^9 PFU/ml.

(e) Screening of rice genomic library and isolation of putative prolamin clones

The amplified library was screened using a heterologous probe namely, kafirin, a prolamin gene from S.bicolor. The library was plated onto NZY plates and the plaques were transferred onto nitrocellulose membranes. The membranes were hybridised with a

radiolabelled kafirin probe to identify plaques containing homologous fragments. Plaques showing dark spots on the autoradiogram were separately amplified and subjected to further rounds of plating and hybridisation. Figure 2.7 shows an autoradiogram of the final round of screening of one of the clones. The whole filter can be observed to exhibit intense spots indicating presence of plaques containing fragments homologous to kafirin probe. Three such clones were identified at the end of the final screening which were labelled as 4A1, 4A2 and 4A3.

(f) Characterization of putative prolamin clones

The three putative clones were grown on a large scale for DNA isolation and the DNAs were digested to completion using EcoRI and electrophoresed on a 0.5 % agarose gel in order to obtain a better resolution of the insert from Charon 4A arms. Figure 2.8 exhibits a digestion pattern of the DNA of all three clones. The two arms and an insert of about 15 kbp are observed in all three lanes (lanes A, B and C). The gel was blotted and hybridised to radiolabelled kafirin in order to confirm the homology of insert to the probe. Figure 2.9 shows an autoradiogram which confirms that the clones do contain an insert homologous to kafirin. Superimposing the gel pattern onto the autoradiogram

Fig 2.8 Recombinant clones digested with EcoRI

Lane a : EcoRI digested Charon 4A1 DNA
Lane b : EcoRI digested Charon 4A2 DNA
Lane c : EcoRI digested Charon 4A3 DNA
Lane d : Lambda HindIII digest as MW marker

Fig 2.9 Autoradiogram of digested clones showing hybridisation with the heterologous kafirin probe

Lane a : Charon 4A1
Lane b : Charon 4A2
Lane c : Charon 4A3

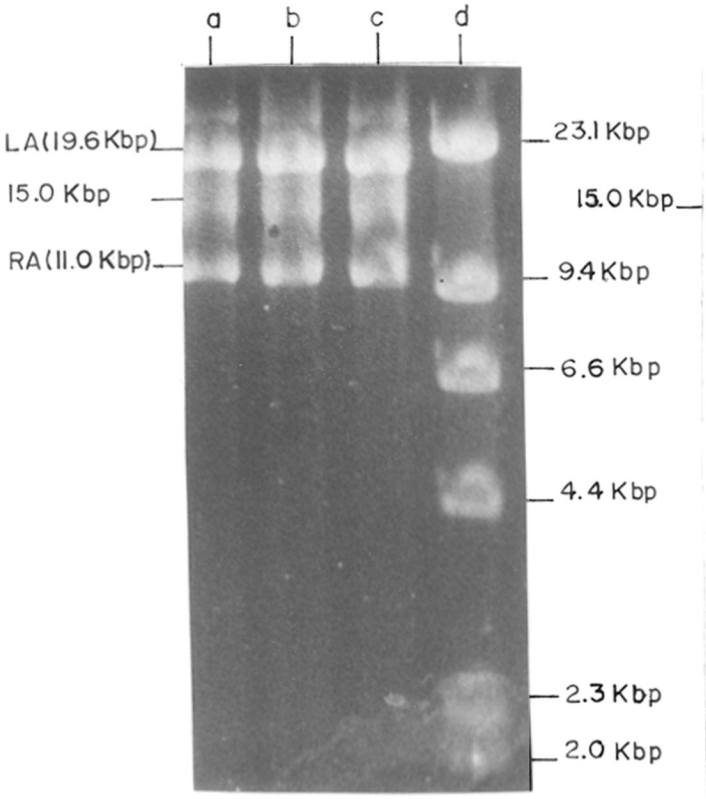

a
b
c
d
LA(19.6Kbp)
15.0 Kbp
RA(11.0 Kbp)
23.1 Kbp
15.0 Kbp
9.4 Kbp
6.6 Kbp
4.4 Kbp
2.3 Kbp
2.0 Kbp

FIG. 2.8


a
b
c

FIG. 2.9

confirmed that the observed bands were of the inserts and not of arms. There also appears to be a variation in intensities of the three bands with 4A3 (lane C) exhibiting the highest band intensity. These clones can now be used as a source material for further subcloning and subsequent sequencing.

DISCUSSION

Construction of genomic libraries is a first step for isolation of genes of specific interest. This involves a number of steps each of which requires some amount of standardisation. I have used one of the earlier vectors of the Charon series, namely Charon 4A for constructing a genomic library of rice. This vector can be used as a replacement vector when digested with EcoRI or as an insertion vector when digested with XbaI. This vector has also been used to generate a maize genomic library from which a 19 kd prolamin clone was isolated (20). During the isolation of Charon 4A DNA, I did encounter a lot of problems. The quality of chemicals like PEG and $CHCl_3$ played an important role in determining the quality of DNA isolated. A lot of degradation was observed due to the low quality of chemicals used. Batchwise variations in the quality of chemicals also added to the difficulties. Despite these difficulties, I did manage to isolate intact vector DNA

which was good in quality and quantity.

Isolation of high molecular weight rice DNA is another important step in this process. Phenolics and carbohydrates present in plant cells affect the final quality of isolated DNA. After trials with various methods of DNA extractions, the method of Shure et al (6) worked well in our laboratory. High molecular weight rice DNA was partially digested with EcoRI in order to generate fragments that can be cloned into the vector. The good quality of isolated rice DNA made it quite simple to optimise partial digestion conditions using EcoRI. Similarly, the quality of Charon 4A DNA was also proved to be excellent by the easy digestion of vector DNA into its arms and stuffer fragments. Thus the DNA obtained from sucrose density gradient runs of partially digested rice DNA and vector arms was found to be good and ligatable.

Another critical step in genomic library construction is the preparation of packaging extract. Though commercial packaging extracts are available (Gigapack Gold), which work extremely well, it was difficult to procure these kits. Moreover, storage of these extracts tends to reduce their packaging capability drastically. Hence I decided to optimise the protocol for preparing a good, efficient packaging

extract so that a fresh extract could be used for genomic library constructions. The single culture method using a lysogenic strain E.coli SMR -10 was optimised which took a period of over 6 - 8 months. During optimisation, I noted that if the culture did not reach the desired O.D within 2.5 - 3 hrs, packaging efficiency was very low. A rapid shaking of 300 rpm at 34°C was necessary to achieve this growth rate. Similarly, temperature changes from 34°C to 42°C to 37°C should be as fast as possible to obtain best results. Addition of DNA had to be done before the frozen packaging extract was thawed. This was contrary to the instructions which stated that the DNA should be added after the extract was thawed. I was able to overcome all these difficulties and achieved a packaging efficiency of about 1 X 10^8 PFU/ug lambda DNA. Using this packaging extract, a genomic library was constructed which represented 7 - 8 genomes of rice.

The library was screened for putative prolamin clones with the help of a heterologous probe namely, kafirin, a prolamin gene from S.bicolor. When I initiated this work, there were few reports on the isolation of prolamin genes of rice and was, therefore, unable to procure a rice prolamin gene that could be

used as a probe. I used the kafirin gene, a kind gift from Prof. T.C. Hall, in order to screen the library for putative prolamin clones. After five rounds of plating, plaque lifting and hybridisation, three clones were identified and isolated. The autoradiogram of digested clones confirmed the presence of 15 kbp inserts in all the three clones.

Kim and Okita (1988) (4) have demonstrated that prolamins are present in 80 - 100 copies per haploid genome of rice. These genes are also known to be heterogeneous in nature. This may probably answer the varying band intensities observed in the autoradiogram confirming the presence of a 15 kbp insert in the three clones homologous to the kafirin gene.

Thus the genomic library of rice constructed in Charon 4A proved to be a marathon effort. Most of the gene work in rice has been confined to japonica cultivars. Genes like actin, waxy gene, glutelins and prolamins have been isolated from japonica varieties. Very little information is available on the genes of indica varieties. Since indica varieties are morphologically distinct from japonica varieties, it is essential to study gene from indica varieties separately. The genomic library constructed here is of an indica cultivar namely, Basmati-370. Hence this

library can be used in screening and isolation of genes from the indica varieties. In addition, this library can also be used in screening for novel rice genes.

Nowadays, better vectors like Charon 40 and lambda ZAP are available which make genomic library constructions less cumbersome and also facilitates ease in screening for genes of interest. An alternative method is the use of PCR technology to identify and isolate genes. Polymerase Chain Reaction or PCR technology is an in vitro method of nucleic acid synthesis by which a particular segment of DNA can be specifically replicated. It involves repeated cycles of annealing of two oligonucleotide primers to complementary regions flanking the DNA fragment to be amplified, extension of the annealed primers with a thermostable DNA polymerase and heat denaturation of the DNA. The ability of PCR (21,22) to generate microgram quantities of a specific DNA fragment can simplify the procedures necessary to clone gene fragments from genomic DNA. This amount of DNA allows direct cloning into plasmid or M13 vector and obviates the need for the construction of phage or cosmid libraries, lengthy screening for recombinant clones, restriction mapping and subcloning (23). But a limitation of PCR cloning compared to conventional genomic cloning is that some sequence information

flanking the desired fragment is required. Though the PCR technology might be an alternative to phage and cosmid library constructions, plasmid libraries will be useful in generating probes for techniques like RFLP where fragments will be of importance as probes. Thus the focus of constructing genomic libraries can be changed from isolation of genes to generation of probes for modern techniques like RFLP.

2.5. REFERENCES

1) Larkins B.A. (1983)
In : A Comprehensive Biochemistry Vol. 6 Eds. Stumpf P.K. and Conn E.E. Academic Press, London pp449

2) Masumura T., Hibino T., Kidzu K., Mitsukawa K., Tanaka K. and Fujii S. (1990)
Mol. Gen. Genet. **221** : 1.

3) Kim W.T. and Okita T.W. (1988)
FEBS Lett. **231** : 308 - 310.

4) Kim W.T. and Okita T.W., (1988)
Plant Physiol. **88** : 649 - 655.

5) Murray M.G. and Thompson W.F. (1980)
Nucl. Acid Res. **8** : 4321.

6) Shure M., Wessler S. and Federoff N. (1983)
Cell **35** : 225.

7) Sambrook J., Fritsch E.F., and Maniatis T. (1989)
In : Molecular cloning : A laboratory Manual 2nd edition, cold Spring Harbor Laboratory Press, NY

8) Ausubel F.M., Brent R., Kingston R.E., Moore D.E., Seidman J.G., Smith J.A., and Struhl K. Eds. (1987)
In : Current Protocols in Molecular Biology. Greene Publishing Associates and Wiley-Interscience, New York.

9) Murray N.E. and Murray K. (1974)
Nature **251** : 476.

10) Williams B.G. and Blattner F.R. (1980)
In : Genetic Engineering Eds. J.K. Setlow and A. Hollaender Vol. 2 Plenum Press, New York. pp 201.

11) Blattner F.R., Williams B.G., Blechl A.E., Denniston K.T., Faber H.E., Furlong L-a, Grunwald D.J., Keifer D.O., Moore D.P., Sheldon E.L. and Smithies O. (1977)
Science **196** : 161 - 169.

12) Dunn and Blattner F (1987)
Nucl. Acid Res. **15** : 2677 - 2698.

13) Frischauf A.M., Garoff H. and Lehrach L (1980)
Nucl. Acid Res. **8** : 5541.

14) Young R.A. and Davis R.W. (1985)
P.N.A.S. USA **80** : 1194 - 1198.

15) Sternberg N.D., Tiemer and Enquist L. (1977)
Gene **1** : 255.

16) Rosenberg S.M., Stahl M.M., Kobayashi I. and Stahl F.W. (1985) Gene **38** :165

17) Benton W.D. and Davis R.W. (1977)
Science **196** :180

18) Feinberg A.P., and Vogelstein B. (1983)
Anal. Biochem **132** : 6.

19) Southern E.M. (1975)
J. Mol. Biol **98** :503.

20) Pederson K., Devereaux J., Wilson D.R. and Larkins B.A. (1982)
Cell **29** : 1015.

21) Saiki R.K., Gelfand D.H., Stoffel S., Scharf S.J., Higuchi R., Horn G.T., Mullis K.B., and Ehrlich H.A. (1988)
Science **239** : 487 - 491.

22) Mullis K.B. and Faloona F.A. (1987)
Met. Enzymol. **155** : 335 - 350

23) Scharf S.J., Long C.M. and Ehrlich H.A. (1988)
Hum. Immunol. **22** : 61 - 69.

CHAPTER III

ANALYSIS OF 5' UPSTREAM REGIONS OF STORAGE PROTEIN GENES OF MONOCOTS AND DICOTS WITH RESPECT OF FREE ENERGY AND DINUCLEOTIDE FREQUENCIES

1) There was a distinct difference in the free energy profiles of 5' regions immediately upstream of monocot and dicot storage protein genes.

2) Dinucleotides TA and CA were predominant contributors to the free energy peak in the region -10 bp upstream of translational initiation site in monocots.

3) Dinucleotides TA, TT and CA were major contributors to the unstable region -10 bp upstream of the start site of dicot storage protein genes.

4) There was a considerable variation in the consensus sequences of monocots and dicot storage protein genes except for a few conserved positions which are common to both.

5) There was a 26 bp nucleotide stretch present in the coding region of rice glutelin gene which was highly homologous to a similar stretch in the noncoding upstream regions of prolamin genes of barley and wheat.

6) Contextual environment of ATG codon was important for its recognition as the authentic translation initiation codon.

3.1. INTRODUCTION

DNA sequences store information necessary for the basic cellular processes of any living system. Some of these sequences carry information that is involved in regulation of expression of genes. Since all genes are not expressed at a given time there must be some fundamental regulatory mechanisms involved in their expression. Cloning and sequencing of genes have resulted in an enormous wealth of information documenting their structure, organization and mode of expression in cells. Identification of TATA and CAAT regions which are important in transcription is a result of analysis of primary sequences of genes. The primary sequences of diverse systems like plants (1), animals (2) and yeasts (3) have been analysed in order to decipher valuable knowledge which would help in understanding gene regulation.

Nussinov (1984) (4) attempted to analyze primary sequences of E.coli promoters and showed the presence of a highly unstable region -10 bp upstream of transcription start site. This region may be assumed to be easily "melted in" by RNA polymerase to facilitate DNA transcription. Margalit et al. (5) attempted to define a relationship between double-helix and activity of E.coli promoters based on their dinucleotide free

energies. As free energy depends on base composition of a DNA sequence, an analysis of 5' upstream regions of genes could probably reveal changes that might take place for transcription. The translational efficiency of mRNA sequence also depends on base composition of 5' untranslated region. Studies on prokaryotic mRNAs have shown that a stable secondary structure does have an inhibitory effect on translation initiation (6). A higher free energy region will tend to have a lesser potential to form such secondary structures. Such studies along with experimental mutational analysis will prove to be valuable in genetic engineering.

Storage protein genes are expressed in a developmental and tissue specific manner in all plants suggesting that there might be certain common features present in their regulatory regions which confer this specificity. In order to identify such consensus regions, the 5' upstream of twenty nine storage protein genes were compiled and analysed separately for monocot and dicot systems for their free energy content, dinucleotide frequencies and direct DNA homology. Significant differences between monocot and dicot storage protein genes were observed in all the three parameters used in this study and their implications in gene regulation are discussed.

3.2 MATERIALS AND METHODS

After a preliminary scanning of the database, GENMON Ver 4.2 EMBL Rel 22.0, obtained from Prof A.S. Kolaskar, Head, Bioinformatics Centre, University of Poona, 14 monocots and 15 dicot storage protein genes were selected which had a minimum length of 250 bp upstream of the translation start site i.e. ATG (Table 3.1 to 3.3). Many of the genes lacked information regarding the position of transcription start site. Hence these genes were aligned at the known translation start site. The 250 bp upstream region should be sufficient to cover the 5' untranslated region and transcription start site. A computer program based on the algorithm of Breslauer et al. (1986) (7) was developed in QuickBasic to calculate the free energy of selected gene sequences. Free energy of dinucleotides, used for calculation of free energies of DNA sequences, were as given by Breslauer et al. (1986) (7). Free energy profiles and dinucleotide frequency distributions of gene sequences of monocots and dicots were calculated using 20 random DNA sequences as control. A separate program was written in order to calculate mononucleotide frequencies of monocot and dicot storage protein genes.

In order to identify a possible consensus sequence in monocot and dicot storage protein genes in 5'

upstream region, the 50/75 consensus rule of Cavener et al. (1991) (8) was used. According to this rule, if a nucleotide occurs at a particular position in more than 50 % of the sequences and is greater than twice the frequency of the next most frequent nucleotide, it is assigned as a sole consensus nucleotide. If the sum of frequencies of two most frequent nucleotides is greater than 75 %, (but neither meet the criteria as a sole consensus nucleotide), they are assigned as co-consensus nucleotides. If no single nucleotide or pair of nucleotides meet these consensus criteria, the letter N is assigned to that position to indicate a lack of consensus or a lower case letter is used to indicate the most frequent nucleotide present.

3.3 RESULTS

As explained earlier, a total of 29 storage protein gene sequences were selected and a 250 bp region upstream of the translation start site was examined for three parameters namely (i) Free energy of the 5' upstream regions (ii) mono- and di- nucleotide frequency distributions of the same regions, and (iii) DNA homology.

(1) Free Energy distribution in the 5' upstream regions

Free energy profiles reflect the stability of DNA

sequences. Figure 3.1 depicts free energy profiles of 5' upstream regions of monocot and dicot storage protein genes while Figure 3.2 shows a free energy profile of 20 random DNA sequences. From the graphs, it can be seen that the free energy of random DNA sequences ranges from -17 kcal/mol to -19 kcal/mol (Fig 3.2) while that of storage proteins genes exhibits a broader range between -13 kcal/mol and -17 kcal/mol. It is also observed that there are 2 peaks in the upstream region of monocots regions, which are indicative of highly unstable regions. These peaks are present about -60 bp and -20 bp upstream of the translation initiation site. Dicot genes, however, exhibit a different free energy profile wherein there is only a single peak between -35 bp and the translation initiation site (Fig. 3.1).

(2) Mono - and di- nucleotide frequency

Free energy of 250 bp upstream region of gene sequences has been calculated using the free energies of dinucleotides. In order to identify dinucleotides that could be contributing to the high free energy in these regions, a frequency distribution profile of each dinucleotide was carried out.

Figures 3.3 and 3.4 exhibit frequency profiles of dinucleotide AA and AT respectively. The frequency of AA

Fig. 3.1 : Free energy distribution in monocots and dicots

Fig. 3.2 Free energy distribution of 20 random DNA sequences

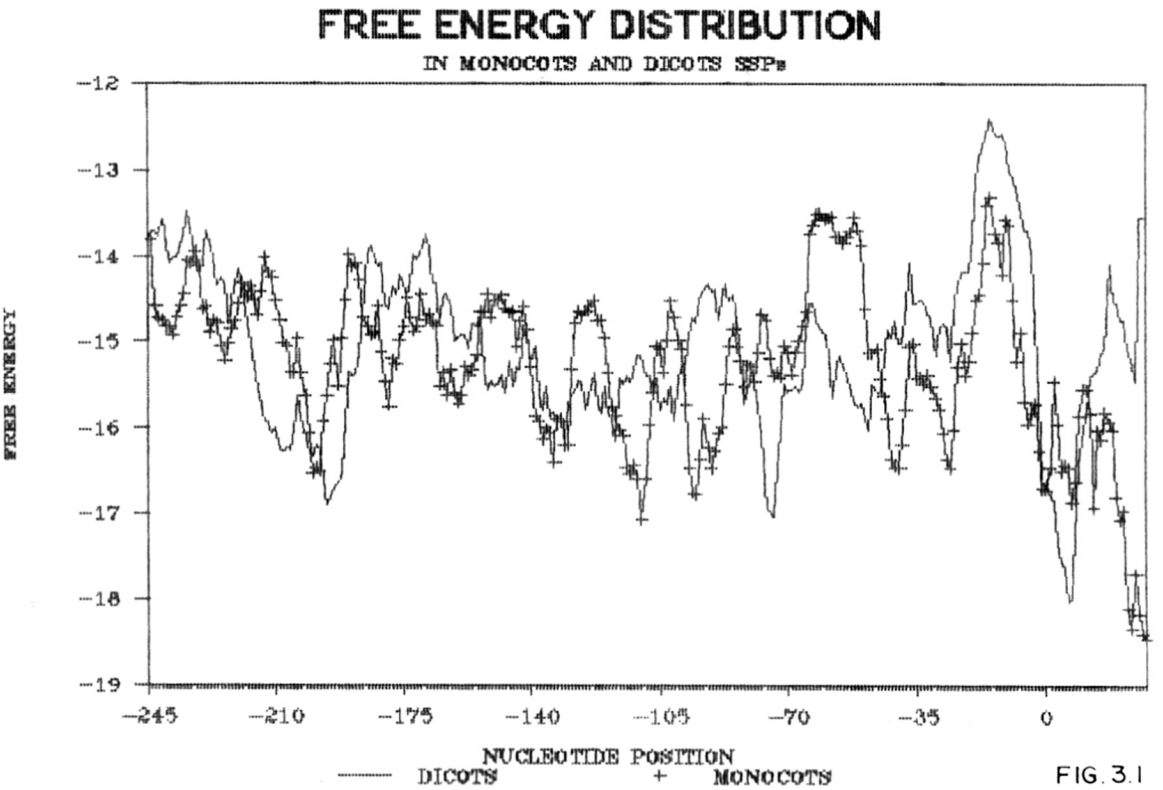

FREE ENERGY DISTRIBUTION
IN MONOCOTS AND DICOTS SSPs
FREE ENERGY
-12
-13
-14
-15
-16
-17
-18
-19
-245
-210
-175
-140
-105
-70
-35
0
NUCLEOTIDE POSITION
DICOTS
MONOCOTS

FIG. 3.1

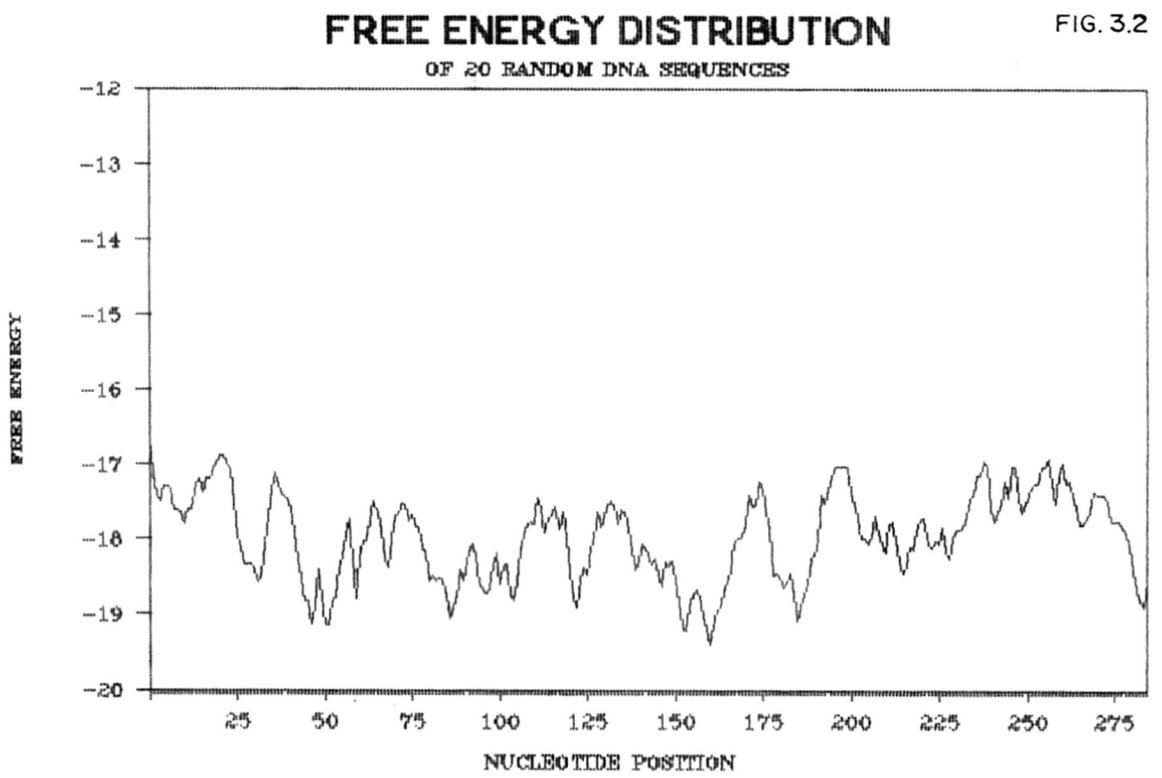

FREE ENERGY DISTRIBUTION
OF 20 RANDOM DNA SEQUENCES
FREE ENERGY
-12
-13
-14
-15
-16
-17
-18
-19
-20
25
50
75
100
125
150
175
200
225
250
275
NUCLEOTIDE POSITION

FIG. 3.2

Fig. 3.3 Frequency of AA in dicots and monocots

Fig. 3.4 Frequency of AT in dicots and monocots

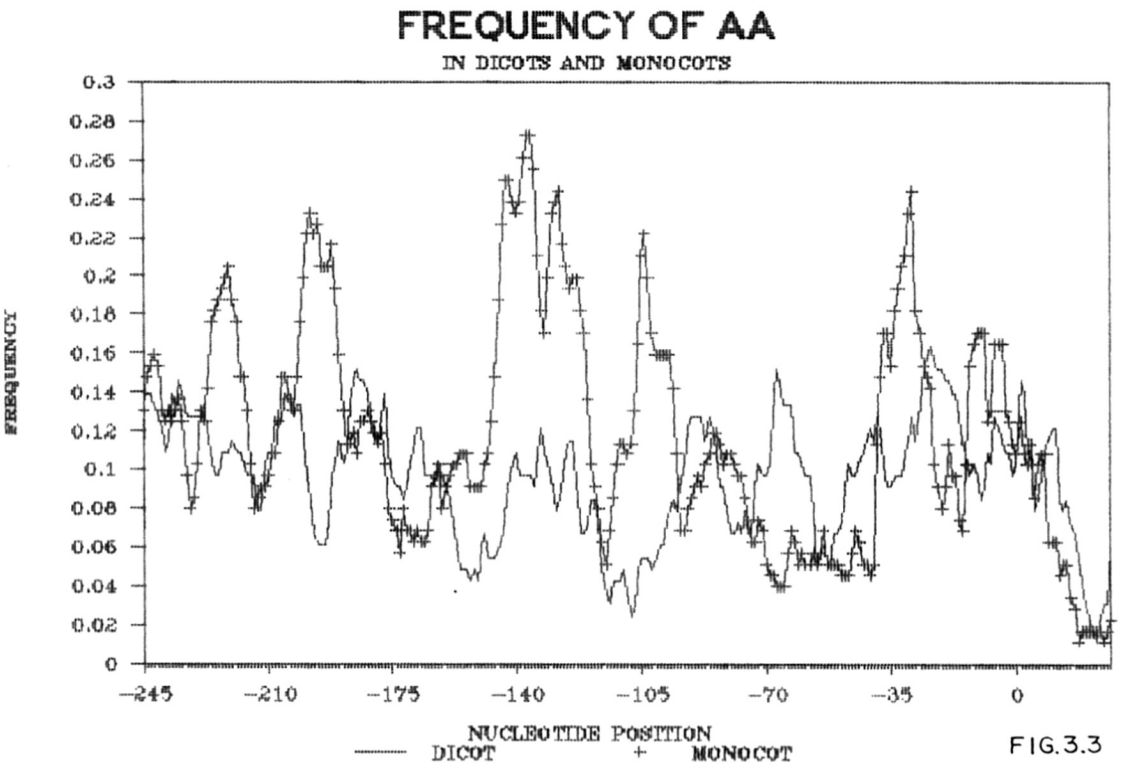

FREQUENCY OF AA
IN DICOTS AND MONOCOTS
FREQUENCY
0.3
0.28
0.26
0.24
0.22
0.2
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
-245
-210
-175
-140
-105
-70
-35
0
NUCLEOTIDE POSITION
DICOT
MONOCOT

FIG. 3.3

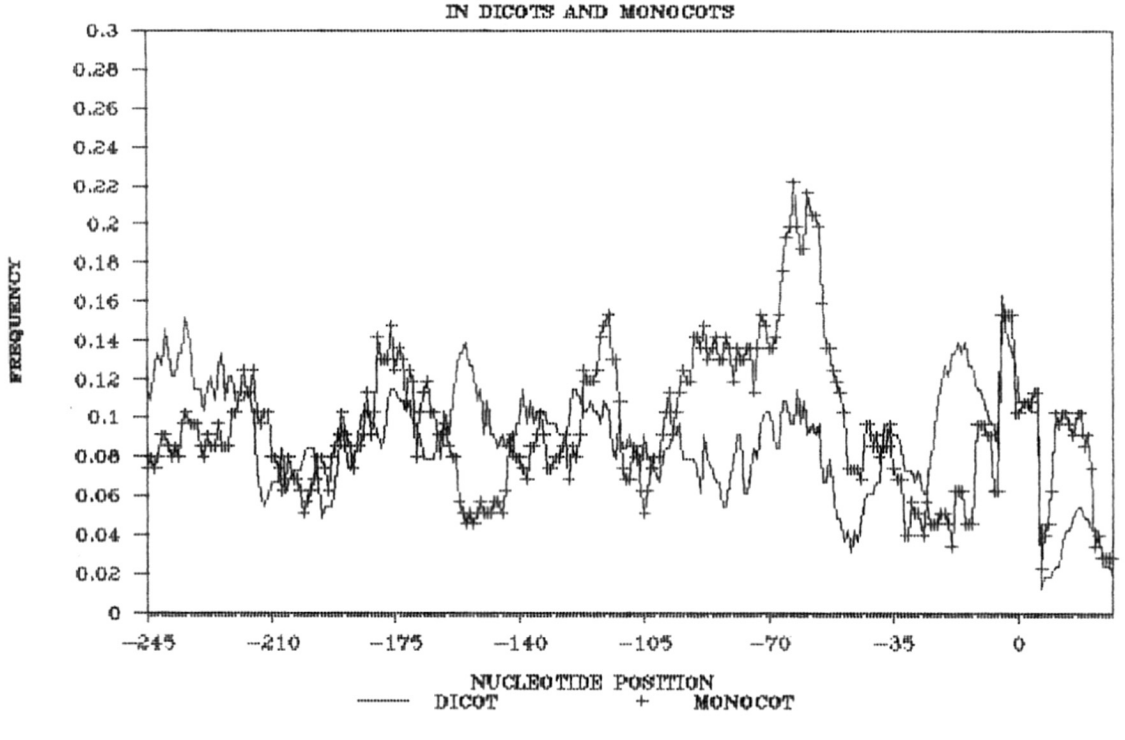

FREQUENCY OF AT
IN DICOTS AND MONOCOTS
FREQUENCY
0.3
0.28
0.26
0.24
0.22
0.2
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
-245
-210
-175
-140
-105
-70
-35
0
NUCLEOTIDE POSITION
DICOT
MONOCOT

FIG. 3.4

Fig. 3.5 Frequency of GC in dicots and monocots

Fig. 3.6 Frequency of TA in dicots and monocots

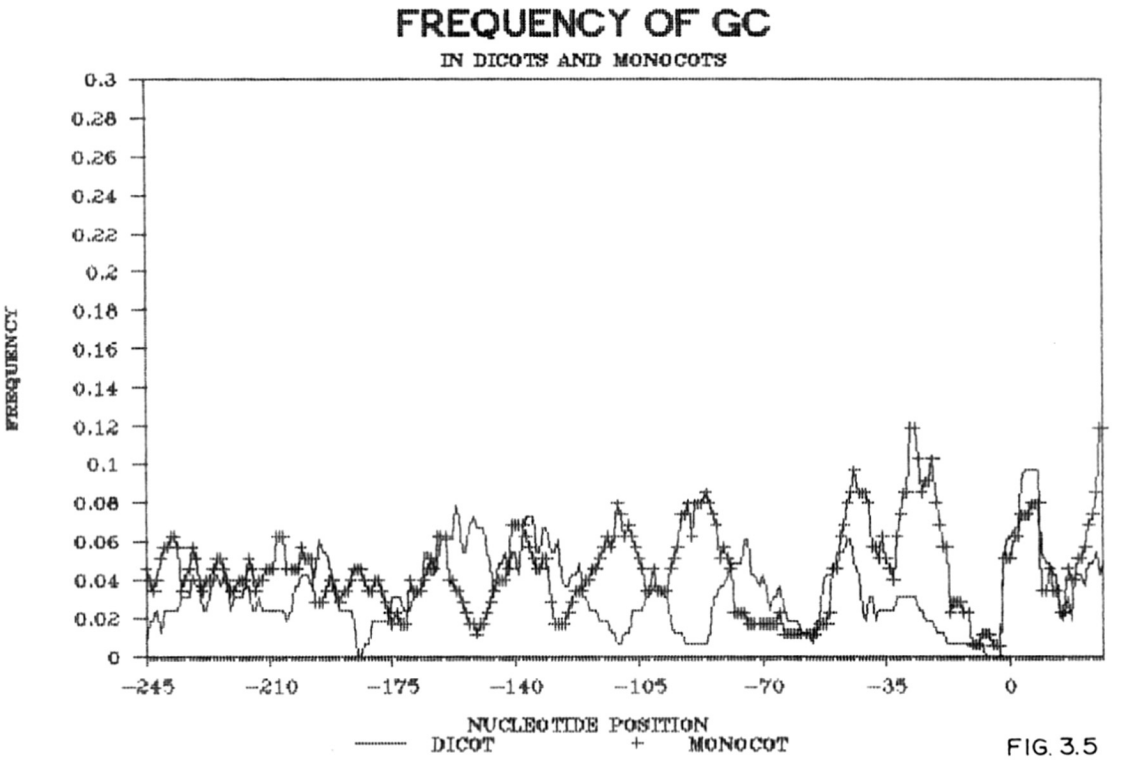

FREQUENCY OF GC
IN DICOTS AND MONOCOTS
FREQUENCY
0.3
0.28
0.26
0.24
0.22
0.2
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
-245
-210
-175
-140
-105
-70
-35
0
NUCLEOTIDE POSITION
DICOT
MONOCOT

FIG. 3.5

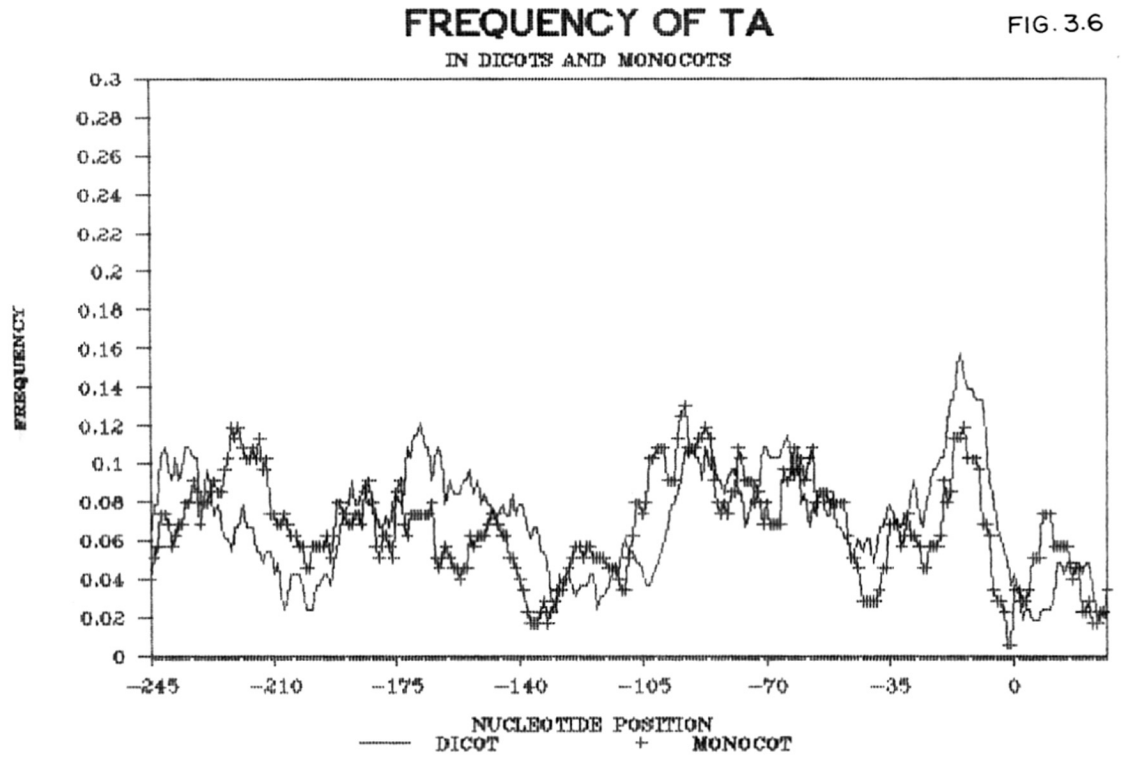

FREQUENCY OF TA
IN DICOTS AND MONOCOTS
FREQUENCY
0.3
0.28
0.26
0.24
0.22
0.2
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
-245
-210
-175
-140
-105
-70
-35
0
NUCLEOTIDE POSITION
DICOT
MONOCOT

FIG. 3.6

Fig. 3.7 Frequency of TT in dicots and monocots

Fig. 3.8 Frequency of CA in dicots and monocots

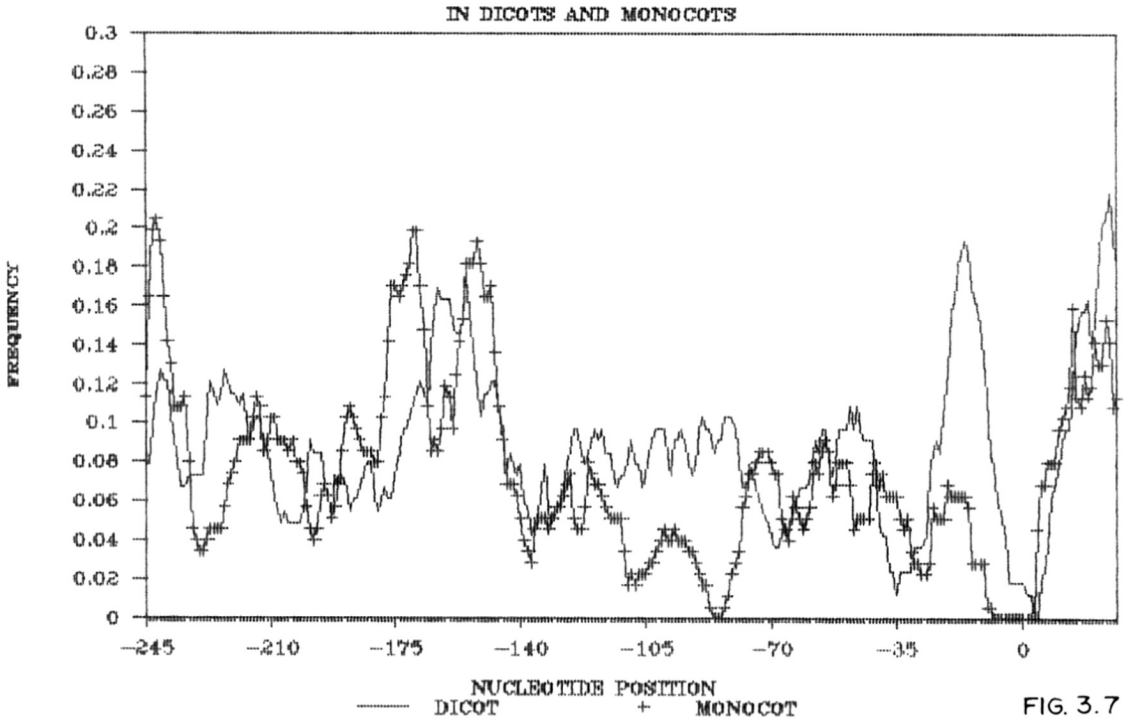

FREQUENCY OF TT
IN DICOTS AND MONOCOTS
FREQUENCY
0.3
0.28
0.26
0.24
0.22
0.2
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
-245
-210
-175
-140
-105
-70
-35
0
NUCLEOTIDE POSITION
DICOT
MONOCOT

FIG. 3.7

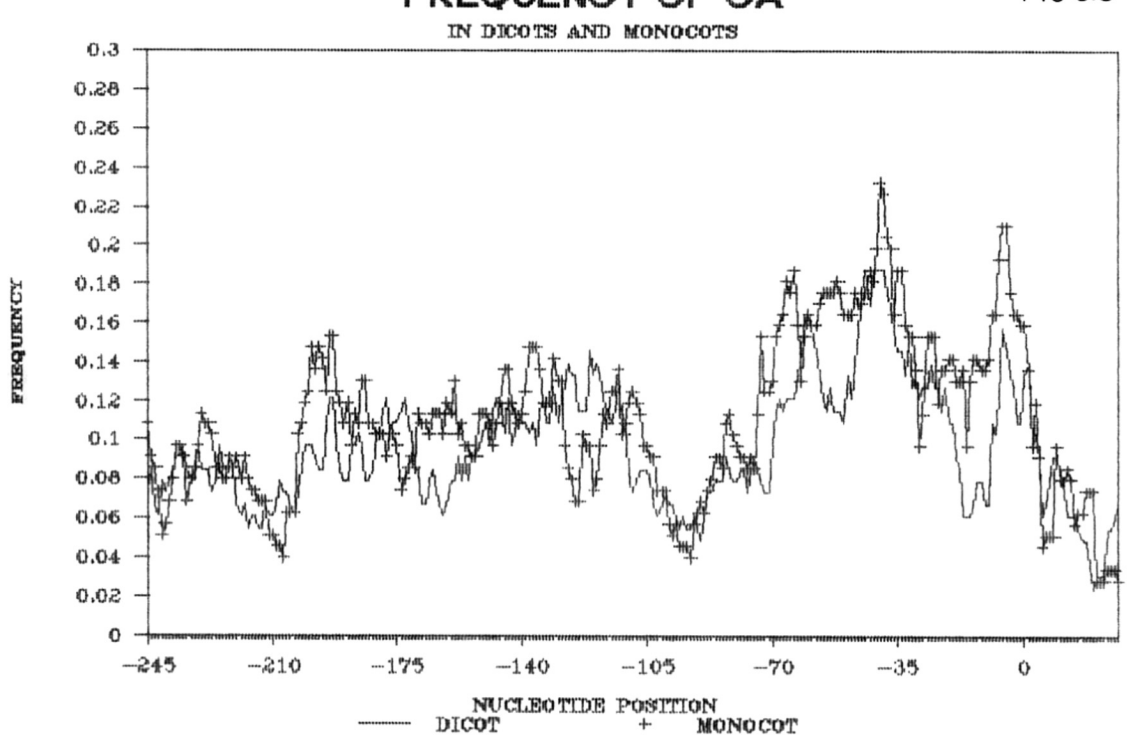

FREQUENCY OF CA
IN DICOTS AND MONOCOTS
FREQUENCY
0.3
0.28
0.26
0.24
0.22
0.2
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
-245
-210
-175
-140
-105
-70
-35
0
NUCLEOTIDE POSITION
DICOT
MONOCOT

FIG.3.8

Fig. 3.9 Frequency of TC in dicots and monocots

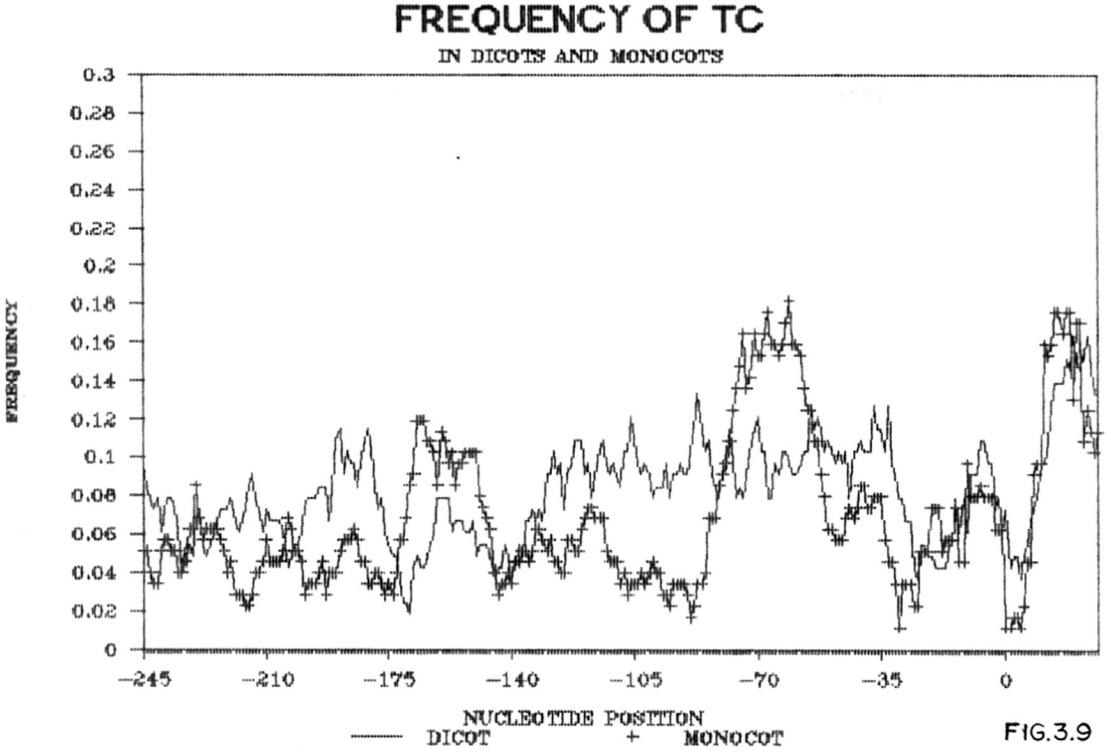

FREQUENCY OF TC
IN DICOTS AND MONOCOTS
FREQUENCY
0.3
0.28
0.26
0.24
0.22
0.2
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
-245
-210
-175
-140
-105
-70
-35
0
NUCLEOTIDE POSITION
DICOT
MONOCOT

FIG.3.9

reaches as high as 0.24 for monocots while it reaches a maximum of only 0.18 for dicots. The frequency of dinucleotides whose first nucleotide is G, namely, GA, GC, GG and GT is rather low in both monocots and dicots. The frequency graph of one of them is shown in Figure 3.5. In Figure 3.6, it can be observed that the distribution pattern of dinucleotide TA is similar in both monocot and dicot storage protein genes with a distinct peak in the regions between -35 bp upstream and the translation start site. Figure 3.7 shows a frequency distribution pattern of dinucleotide TT with a distinct peak in the region between -35 bp and translation initiation site for dicots which is absent in monocots. Figure 3.8 also exhibits a distribution pattern of dinucleotide CA that is similar in monocots and dicots.

Thus, in monocots, dinucleotides TA and CA appear to be major contributors for the high free energy peak between -35 bp and translation start site and in dicots, an additional contributor could be dinucleotide TT. The major contributors to free energy peak around -70 bp region in monocots are dinucleotides AT, TC (Fig. 3.9) and CA; whereas in dicots, there is no peak observed aound this region as stated earlier.

3) DNA Homology studies

The 5' upstream regions of monocot and dicot storage protein genes were analysed for direct DNA homology in order to determine a consensus sequence. Genes were aligned at the translation start site and their positional frequency of mononucleotides was calculated using Cavener's Rule. The consensus sequences for monocots and dicots are as follows :

Monocots : NAANNNAANNAAGNANTNGTTAANACNAANCCACCATG
Dicots : CNANNNNNNACNAANNTNTNTNAATANNANTCANNATG

Considerable variation is observed in the consensus sequences of monocots and dicots, especially as the distance between the nucleotide and the translation start site increases. In spite of the larger number of monocot storage protein genes used, the monocot consensus sequence is similar to that proposed by Joshi (9). But, the dicot storage protein gene consensus sequence NTCANNATG shows more variation than that proposed by Joshi (9) i.e. CTCATAATG. Therefore it appears that there are more conserved sequences in monocots than in dicots. Monocots have a higher number of conserved A residues while dicots have conserved A and T residues. The A residue is conserved in both monocot and dicot consensus sequences at positions -3, -7 and -13 while the C residue is conserved only at

position -4. The T residue is conserved at positions -15 and -19 from the translation start site.

In order to identify homologous regions in the 5' upstream regions of storage protein genes, a comparison was done using the DNASIS software.

OSGLUTG : **CATGGCATCCATAAATCGCCCCATAG**
TAGLIAA : **CATGGCAACTATAAATAGCCCCATAG**
HVBIHORG : **CATGGCTACTATAAATAGGCATGTAG**

During this comparison, a 26 nucleotide stretch in rice glutelin gene which is highly homologous to a similar stretch of B1 hordein gene of barley (10) and alpha-gliadin (11) gene of wheat is detected. This region is present in the coding region of rice glutelin gene. From published data, it is known that these regions, present in barley and wheat prolamin genes, contain the putative TATA box.

3.4 DISCUSSION

Structural analysis of plant genes has been carried out by Joshi (9) in order to determine consensus sequences for transcription and translation sites. Kreis et al. (1985) (12) have compiled the sequences of available seed storage protein gene sequences and attempted to determine their evolutionary relationship. Takaiwa et al. (1991) (13) have compared the DNA sequences of glutelin genes and determined their phylogenetic relationship via a dendrogram. It is for the first

netic relationship via a dendrogram. It is for the first time that in the present work, the storage protein gene sequences are looked into from the point of view of free energy and dinucleotide frequency. The highlight of the present work is a demonstration of a correlation between free energy profiles and dinucleotide frequency of the 250 bp upstream of the translation start site.

The free energy profiles of monocot and dicot storage protein genes exhibit peaks which indicate instability of these regions. A probable explanation to the unstable region adjacent to translation initiation site is that this region could reflect instability in 5' untranslated mRNA leader sequence. From studies carried out on viral mRNAs (14), such regions present in the leader sequences are less amenable to secondary structure formation creating lesser obstacles to the scanning eucaryotic ribosomes (2). Such free energy studies have not been carried out in 5' untranslated regions of genes. DNA sequences of regions exhibiting the peaks do not reveal any significant homology. This indicates that the high free energy associated with these regions is independent of nucleotide sequence. As the transcription start sites of many genes were not available, it is difficult to state whether the peak present -70 bp upstream of translation start site is

present within the untranslated leader sequence.

There appears to be a preference for dinucleotides TT, CA, AA and TA in the unstable regions. The 5' upstream regions seem to have a tendency to avoid the dinucleotides that have a low free energy values like GC (-3.1 kcal/mol), GG (-3.1 kcal/mol), CC (-3.6 kcal/mol) and CG (-3.1 kcal/mol) as observed by the low frequency distribution pattern of these dinucleotides. There appears to be bias towards dinucleotides that have higher free energy values like TA (-0.9 kcal/mol), TT (-1.9 kcal/mol), CA (-1.9 kcal/mol) and AA (-1.9 kcal/mol). This tendency can be directly correlated to the higher free energy profiles of the regions between -250 bp and translation start site. Thus the 5' upstream regions of storage protein genes tend to accumulate dinucleotides that have a lower potential to form secondary structures.

Kozak (2) has compiled the animal sequences and proposed a consensus sequence for the translation start site i.e. CACCATG. Plant DNA sequences have been compiled by Joshi (9) who has proposed AACAATG as the consensus sequence. In the present work, monocot and dicot storage protein gene sequences have been compiled to determine the consensus sequences for each.

The consensus sequences of monocot and dicot

storage protein genes have a predominance of A and T residues. Dinucleotide combinations of A and T have a higher free energy than other combinations tending to make those regions unstable. This can be observed by the highly unstable region between positions -35 bp and translational start site of both monocot and dicot storage protein genes. Kozak (2) has hypothesised that contextual recognition of ATG codon is also critical in determining its functioning as the authentic translation start site. It can be noted that the A residue at -3 position is conserved in the consensus sequences of monocots as well as dicots. The position of A residue at -3 position is universally conserved as observed in the consensus sequences of plants **AACAATG** (1), animals **CACCATG** (2), Drosophila **(C/A)AA(A/C)ATG** (15), and yeasts **(A/T)A(A/C)A(A/C)AATG** (3). Consensus sequences of storage protein genes of both monocots and dicots differ considerably from the plant consensus sequence. The monocot consensus sequence is almost identical to that proposed by Kozak (2) for animal genes. But the dicot consensus sequence differs considerably, except for positions of A at -7 and -3 and C at -4. Hence, it may indicate that these residues could be involved in the contextual recognition of translational start site (9).

Another highlight of this work is the

identification of a homologous nucleotide stretch in rice glutelin, B1 hordein and alpha gliadin. To our knowledge, there is no report which has documented this homologous stretch. In this region, B1 hordein gene contains two ATG codons while α-gliadin and rice glutelin genes contain one each. But only the ATG codon of rice glutelin gene is recognised as the translation initiation codon. These regions are found about -140 bp upstream of the translational start site of B1 hordein and α-gliadin gene. Only the ATG of rice glutelin gene has an A residue present in the -3 position. Absence of A at -3 position of the ATG codons of gliadin and hordein could lead to non-recognition of these codons as the authentic translation start site. This indicates that contextual environment of ATG codon is important due to which ATG codons of gliadin and hordein genes are not recognised as authentic translational initiation codons.

The number of nucleotides conserved is more in storage protein genes of monocots than in dicots. This could probably signify that regulation in monocots is more specific than in dicots. The higher conservation in nucleotides of monocots could probably be related to the fact that monocots are more evolved than dicots. Thus such studies, along with experimental evidences,

will help in understanding differential regulation of storage protein genes in monocots and dicots.

3.5 REFERENCES

1) Heidecker G. and Messing J. (1986)
Ann. Rev. Plant Physiol. : **37** , 439 - 466.

2) Kozak M. (1986)
Cell. **44** : 283 - 292.

3) Hamilton R., Watanabe C. K., and Herman A.B., (1987)
Nucl. Acid Res. : **15**, 3581 - 3593.

4) Nussinov R. (1984)
J. Biol. Chem. **259** : 6798 - 6805.

5) Margalit H., Shapiro B.A., Nussinov R., Owens J. and Jernigan R.L. (1988)
Biochem. **27** : 5147 - 5188.

6) De Smit M.H. and Duin J.V. (1990)
Prog. in Nucleic Acid Res. and Mol. Biol. Vol. 38. pp 1 - 35. Eds. Cohn W.E. and Moldave K. Academic Press

7) Breslauer K.J., Frank R., Blocker H. and Marky L.A. (1986)
P.N.A.S. USA. **83** : 3746 - 3750.

8) Cavener D.C. and Ray S.C. (1991)
Nucl. Acid Res. **9** : 3185 - 3192.

9) Joshi C.P. (1987)
Nucl. Acid Res. **15(16)** : 6643 - 6653.

10) Forde B.G., Heyworth A,. Pywell J. and Kreis M. (1985)

Nucl. Acid Res.. **13** : 7327 - 7339.

11) Andersen O.D., Litts J.C., Gautier M.F. and Greene F.C. (1984)

Nucl. Acid Res. **12** : 8129 - 8144.

12) Kreis M., Shewry P.R., Forde B.G., Forde J. and Miflin B.J. (1985)

In. Oxford Surveys of Plant Molecular and Cell Biology Ed. Miflin B.J. Vol 2 pp 253 - 317, Oxford Univ Press. Oxford.

13) Takaiwa F. and Oono K. (1991)

Jpn. J. Genet. **66** : 161 - 171.

14) Gallie D.R., Sleat D.E., Watts J.W., Turner P.C. and Wilson T.M.A. (1987)

Nucl. Acid Res. **15** : 8693 - 8711.

15) Cavener D. (1987)

Nucl. Acid Res. : **15**, 1353 - 1361.

TABLE 3.1

GENE	DESCRIPTION
DICOTS	
ATCRA1	Arabidopsis CRA1 gene for 12S seed storage protein
ATCRB	Arabidopsis CRB gene for 12S seed storage protein
CGCONA	Canavalia gladiata concanavalin A gene
GHSPA	G.hirsutum (cotton) storage protein (alpha-globulin A) gene,
GHSPH	G.hirsutum (cotton) storage protein (late embryogenesis abundant)
GMGLYCAB	Soybean DNA for glycinin A2B1a subunit
HAG5ALB2	Sunflower HaG5 gene for 2 S albumin storage protein
BNNAPA	B.napus napA gene encoding the 1.7S seed storage protein gene
PSLEGAG	Pea legA gene for legumin
PSLEGB5	Pea legB gene (legumin) 5' flanking region
PSVICIL	Pea vicilin gene
GMBPSP	Glycine max α'-type beta conglycinin storage protein gene
VFLEB4	Vicia faba legumin B gene Leb4
VFLEB1	Vicia faba VfLEB1 pseudogene

TABLE 3.2

GENE	DESCRIPTION
MONOCOTS	
HVB1HORG	Barley gene for B1 hordein
HVGHRDSP	Barley gene for storage protein gamma-hordein
OSGLUTG	Rice glutelin gene
OSPROL	Rice mRNA for prolamine
TAGLIAA	Wheat gene for storage protein gliadin (A-gliadin subgroup)
TAGLIAG1	Wheat gene for alpha/beta-gliadin storage protein (pW1215)
TAGLIAG2	Wheat gene for alpha/beta-gliadin storage protein (pW 8233)
TAGLIAG3	Wheat gene for alpha/beta-gliadin storage protein
TAGLU1AG	Wheat gene for HMW-glutenin subunit located on chromosome 1A
TAGLU1DG	Wheat gene for HMW-glutenin subunit from chromosome 1D
TAGLUT	Wheat gene for low molecular weight glutenin
TAGLUT1	Wheat gene for HMW glutenin subunit
ZMZEI19	Maize gene for Mr 19000 alpha zein and 5'-flanking region
ZMZEP	Maize zein gene (pm11) promoters and partial coding sequence
TAGLGB	Wheat (T.aestivum) gamma gliadin gene, complete cds

TABLE 3.3

GENE	REFERENCE
DICOTS	
GHSPA	Dure L. III unpubl. (1988)
GHSPH	Dure L. III unpubl. (1988)
BNNAPA	Josefsson L.G., Lenman M., Ericson M.L. and Rask L. (1987) J. Biol. Chem. **262** : 12196 - 12201.
GMBSPA	Doyhe J.J., Schuler M.A., Godette W.D., Zenger V., Beachy R.N. and Slightom J.L. (1986) J. Biol. Chem. **261** : 9228 - 9238.
GMGLYCAB	Fukazawa C., Momma T., Higuchi W. and Vdaka K. (1987) Nucl. Acid Res. **15** : 8117 - 8127.
HAG5ALB2	Allen R.P., Cohen E.A., Vander Hear R.A., Adams C.A., Mu D.P., Nessler C.L. and Thomas T.L. (1987) Mol. Gen. Genet. **210** : 211 - 218.
VFLEB4	Baumlein H., Wobus V., Pustell J. and Kafatos F.C. (1986) Nucl. Acid Res. **13** : 6733 - 6743.
PSVICIL	Wattson M.D., Lambert N., Delanner A., Yarweed J.N., Croy R.R.D., Gatehouse J.A., Wright D.J. and Boulter D. (1988) Biochem. J. **251** : 857 - 864.
ATCRAI	Peng P.P., Pruits R.C. and Moyerowitz E.M. (1988) Plant Mol. Biol. **11** : 805 - 820.
CGCONA	Yamauchi D. and Minamikawa T. (1990) FEBS. Lett. **260** : 127 - 130.
VFLEBI	Heim V., Schubroth R., Boeumlein H. and Wobus J. (1989) Plant Mol. Biol. **13** : 653 - 663.
ATCRB	Pany P.P., Pruitt R.E. and Moyerowitz E.M. (1988) Plant Mol. Biol. **11** : 805 - 820.
PSLEGAG	Lycett. G.W., Croy R.R.D., Shirsat A.H., Richards D.N., and Boulter D. (1985)
PSLEGB5	Nucl. Acid Res. **13** : 6733 - 6743.

TABLE 3.4

GENE	REFERENCE
MONOCOTS	
HVB1HORG	Forde B.G., Heyworth A., Pywell J., and Kreis M. (1985) Nucl. Acid Res. **13** : 7327 - 7339.
OSGLUTG	Takaiwa F., Ebinuma H., Kikuchi S. and Oono K (1987) FEBS. Lett. **221** : 43 - 47.
TAGLIAA	Anderson O.D., Litts J.C., Gautier M.F. and Greene F.C. (1984) N.A.R., **12** : 8129-8144
TAGLGB	Rafalski J.A. (1986) Gene **43** : 221 - 229.
TAGLIAG1 TAGLIAG2 TAGLIAG3	Sumner-Smith M., Rafalski J.A., Sugiyama T., Stoll M. and Soell D.G. (1985) Nucl. Acid Res. **13** : 3905 - 3916,
TAGLU1AG	Forde J., Malpica J.M., Halford N.G. and Shewry P,R. (1985) Nucl. Acid Res. **13** : 6817 - 6832.
TAGLU1DG	Thompson R.D., Bartels D. and Harberd N.P. (1985) Nucl. Acid Res. **16** : 6833 - 6846.
TAGLUT	Pitts E.G., Rafalski J.A. and Hedgcoth C. (1988) Nucl. Acid Res. **16** : 11376 - 11376.
TAGLUT1	Sugiyama T., Rafalski J.A., Peterson D. and Soell D.G. (1985) Nucl. Acid Res. **13** : 8729 - 8737.
OSPROL	Kim W.T. and Okita T.W. (1988) FEBS Lett. **231** : 308 - 310.
ZMZEI19	Kriz A.L., Boston R.S. and Larkins B.A. (1987) Mol. Gen. Genet. **207** : 90 - 98.
ZMZEP	Langridge P. and Feix G. (1983) Cell **24** : 1015 - 1022.
HVGHRDSP	Cameron-Mills V. and Brandt A. (1988) Plant Mol. Biol. **11** : 449 - 461.

CHAPTER IV

OPTIMISATION OF PROTOPLAST ISOLATION CONDITIONS FOR THREE INDICA RICE CULTIVARS

ABSTRACT

In the present work, mature rice seeds of three indica cultivars namely, Basmati-370, VDN-3630 and Indrayani were used as a source material for callus induction. A medium containing MS basal medium + 2.0 mg/lit 2,4-D + 0.5 mg/lit kinetin + 100 mg/lit meso-inositol + 2 % sucrose + 0.45 % agar was found to be suitable for callus initiation in all three varieties. Genotypic variation was observed in the calli of the three varieties. VDN-3630 callus had a faster multiplication rate than Basmati-370 or Indrayani calllus; but Indrayani callus was more friable than Basmati-370 or VDN-3630 callus. Conditions for optimum protoplast release differed for all three varieties. VDN-3630 callus gave a high amount of protoplasts as compared to Basmati-370 or Indrayani callus. Protoplasts of VDN-3630 were also larger in size than other two varieties. The purified protoplasts af all three varieties were used to study uptake of foreignDNA by the PEG method. A binary plasmid vector pBI121 containing nptII and gus reported genes was used for transformation. Transient expression of GUS activity was monitored after 48 hrs of incubation. The activity of the GUS protein was spectrophotometrically measured using para-nitrophenol-β-D-glucuronide as substrate. VDN-3630 protoplasts exhibited a higher fold of

expression as compared to transformed Basmati-370 or Indrayani protoplasts.

4.1 Introduction

Introduction and expression of foreign genes into plants represent one of the most significant developments in the systematic advancement in agricultural technology. To acheive this, selection of a suitable source material that is amenable to foreign DNA uptake is a very important factor. Different plant parts such as anthers, roots; and immature and mature embryos have been used to generate callus and suspension cultures which are used for isolating protoplasts for subsequent transformation studies. The ability of a source material to take up foreign DNA is usually monitored by use of marker genes. A number of marker genes such as Neomycin phosphotransferase II (nptII), Chloramphenicol Acetyl Transferase (cat), and β- Glucuronidase (gusA) genes now been identified as potentially useful "reporters" of transformation events. Characteristically, these genes have an ability to differentiate between transformed and nontransformed cells by detection of a gene product for which an assay is available. Thus selection of a suitable cell line and a marker gene is crucial for developing strategies of introducing desirable genes into plants.

Indica and Japonica are two most important subspecies of Oryza sativa. Routine procedures have

been established for regeneration of japonica rice protoplasts to whole plants using electroporation and PEG methods (1, 2, 3, 4). Indica varieties are widely cultivated but, unlike japonica rice, are generally poor in plant regeneration (5). There have been attempts to regenerate plants from protoplasts of Indica rice and successful production of plants has been acheived only recently (6, 7, 8). The particle bombardment technique has been used by Christou et al (1991) (9) to develop a variety independent procedure that could transform suspension cells of both indica and japonica cultivars.

In the present study, of three indica rice cultivars namely, Basmati-370, Vadgaon-3630, and Indrayani were used as a source material with a view of developing a system for transferring genes into these varieties. Indrayani and VDN-3630 are locally cultivated elite varieties with a strong aroma and long, slender grain type similar to that of Basmati. Seeds of Indrayani, VDN-3630 and Basmati-370 were inoculated onto MS medium containing various combinations of hormones in order to select an optimum concentration required for successful initiation of callus which was subsequently used to determine optimum conditions for protoplast isolation. The isolated

protoplasts were used in transformation experiments using PEG and plasmid pBI121 containing gus and nptII reporter genes. Expression of GUS gene in protoplasts was spectrophotometrically assayed using para-nitrophenol-β-D-glucuronide (PNPG) as a substrate.

4.2 MATERIALS AND METHODS

4.2.1. Surface sterilisation of seed material

Rice seeds were used as a source material for callus induction. These seeds need to be pretreated in order to remove extraneous material and microorganisms from their surface. The procedure follows a sequential treatment with chemicals that clean the seeds thoroughly and finally eliminate microbes from the seed surface so as to obtain microbiologically sterile seeds which can be used for callus induction. The precise protocol is elaborated below :

1) Mature rice seeds were dehusked and then washed under running tap water for 5 minutes.
2) This was followed by rinsing with a detergent (Teepol) for 2-3 minutes to remove dirt and other extraneous material from the seed surface.
3) The seeds were washed thoroughly with distilled water in order to remove the detergent traces.
4) They were kept under 70% ethanol for 5 minutes and washed extensively with distilled water and

transferred to a sterile flask containing 20ml of 0.1% $HgCl_2$ for 10 minutes.

5) The seeds were then washed 3-5 times with sterile distilled water in order to remove traces of $HgCl_2$.

4.2.2. Callus induction

Surface sterilised seeds were used for induction of callus which was subsequently used for protoplast isolation. A callus is a mass of undifferentiated cells, each of which has a potential to give rise to whole plants. Seeds were inoculated onto media containing different concentrations of hormones to determine the optimum combination for callus initiation and maintenance. Tubes containing 2-3 seeds and callus initiation medium were incubated at 27°C in the dark for 3-4 weeks for callus formation. The callus was subcultured onto fresh medium every 3-4 weeks and incubated in the dark at 27°C. The composition of media used for induction of callus is described in Table 4 (4.1.1 to 4.1.3).

4.2.3. Isolation of protoplasts from Indica varieties

Isolation of protoplasts is an important step in plant transformation studies and the conditions of protoplast isolation vary from plant to plant. Even

TABLE 4.1.1

MEDIA COMPOSITION

A. MACRONUTRIENTS (expressed in mg/lit)

	MS (10)	N6 (11)	AA (12)	CPW (13)
NH_4NO_3	1650	-	-	-
KNO_3	1900	2830	-	10
$CaCl_2.2H_2O$	440	125	440	148
$MgSO_4.7H_2O$	370	90	370	24.6
KH_2PO_4	170	400	170	2.7
$(NH_4)_2SO_4$	-	463	-	-
KCl	-	-	2940	-

TABLE 4.1.2

MEDIA COMPOSITION

B. MICRONUTRIENTS (expressed in mg/lit)

	MS (10)	N6 (11)	AA (12)	CPW (13)
H_3BO_3	6.20	1.60	6.20	-
$MnSO_4.4H_2O$	22.30	3.30	22.30	-
$ZnSO_4.7H_2O$	8.60	1.50	8.60	-
KI	0.83	0.80	0.83	1.60
$Na_2MoO_4.2H_2O$	0.25	-	0.25	-
$CoCl_2.6H_2O$	0.025	-	0.025	-
$CuSO_4.5H_2O$	0.025	-	0.025	0.25
$FeSO_4.7H_2O$	27.80	27.80	27.85	-
Na_2EDTA	37.30	37.20	37.25	-

TABLE 4.1.3

MEDIA COMPOSITION

C. OTHER CONSTITUENTS (expressed in mg/lit)

	MS (10)	N6 (11)	AA (12)
Thiamine HCl	0.1	2.0	0.1
Pyridoxine HCl	0.5	0.5	0.5
Nicotinic Acid	0.5	0.5	0.5
Glycine	2.0	2.0	-
Inositol	100	200	100
Casein hydrolysate	-	200	-

though a general protocol is available for a particular plant, the exact protocol varies with the variety and hence it is necessary to optimise protoplast isolation conditions for each variety. Optimisation of protoplast isolation conditions for three rice varieties i.e. VDN 3630, Indrayani and Basmati-370 is described here.

1) About 1 g of callus tissue was taken in a flask, weighed accurately and treated with 5 ml of enzyme mixture solution. Initial conditions were arbitrarily chosen to be:

Cellulase R10	- 5%
Macerozyme R10	- 1%
Mannitol	- 12%
CPW (10x)	- 0.5ml
Incubation Period	- 2 hours shaking followed by 2 hours stationary

2) Callus clumps were teased with a sterile rod so that maximum cells were available for enzyme action.

3) The callus was incubated in the enzyme mixture solution and kept on a rotary shaker at 60 rpm for 2 hours. Later, the flask was kept stationary for 2 hours.

4) The enzyme solution was then filtered through a 200u mesh steel wire gauze in order to separate the undigested cell clumps and the filtrate was centrifuged at 800rpm for 5 minutes to pellet protoplasts.

5) The supernatant was decanted and protoplast pellet was suspended in washing solution containing CPW (1x) and 12% Mannitol and centrifuged at 800 rpm for 5 min. This step was repeated twice.

6) The pellet was then suspended in 2.0ml of washing solution and a drop of this solution was used for counting protoplasts on a haemocytometer. An average of three samples were taken to determine the protoplast count.

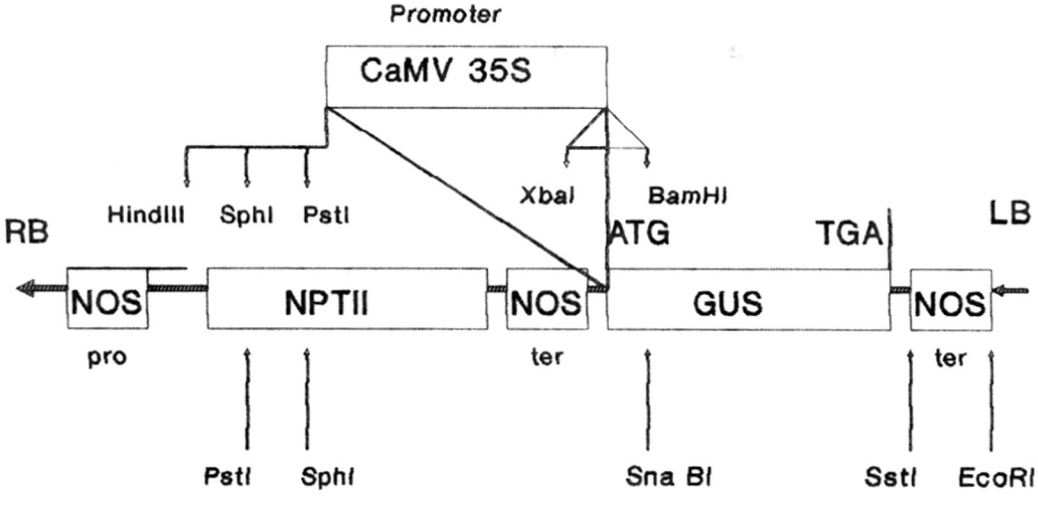
4.2.4. Isolation of plasmid pBI121

Indica protoplasts were transformed by using a binary plasmid vector pBI121 (14). The linear map of the plasmid is depicted in Figure 4. This plasmid is a derivative of pBI101 and is about 12.8 kbp in size. It contains nptII and gus genes which can act upon aminoglycosides like kanamycin and β-glucuronides like para-nitrophenol-β-D-Glucuronide, respectively. The nptII and gus genes are controlled by NOS promoter and constitutive CaMV 35S promoter respectively and have NOS terminator sequences. Multiple cloning sites are present between CaMV 35S promoter and GUS gene. While nptII gene is used to maintain a selection pressure, gus gene is used to monitor transient activity.

Isolation of plasmid DNA was carried out according to the method of Birboim and Doly (1979) (15). Purity

Fig. 4 : Linear map of pBI121

NOS	: Nopaline synthase
NPTII	: Neomycin phosphotransferase
GUS	: β- glucuronidase
pro	: promoter
ter	: terminator
RB	: Right border
LB	: Left border
CaMV 35S	: Cauliflower mosaic virus 35S



LINEAR MAP OF pBI121 (12.8 kbp)

Fig 4

and concentration of the plasmid DNA were determined by agarose gel electrophoresis and spectrophotometric analysis as described in Chapter II. 30 ug of plasmid DNA was used for each transformation experiment.

4.2.5. Transformation of Indica rice protoplasts using PEG

There are different methods by which foreign DNA can be introduced into protoplasts. The commonly used methods are PEG, electroporation and particle bombardment. Each method has its advantages and disadvantages. Here, we have used the PEG method for introducing gus gene into indica rice protoplasts. Various modifications have been introduced for different plant species such as adding PEG after DNA and heat shocking protoplasts at different temperatures followed by chilling on ice. The transformation protocol followed here is essentially as described by Hayashimoto (1990) (16).

1) Isolated protoplasts were suspended in 500 ul MaMg solution containing 15 mM $MgCl_2$ and appropriate concentration of mannitol (7% for Basmati-370 and Indrayani and 12% for VDN-3630) and were distributed into 2 sterile tubes.
2) 100 ul of plasmid DNA solution containing 30 ug DNA dissolved in $T_{10}E_1$ was added to one of the

tubes and incubated at RT for 10 min.

3) 350 ul of 40 % PEG 6000 (dissolved in CPW salts and mannitol) was added to the tubes and mixed gently.

4) The tubes were incubated at RT for 10 min. in order to facilitate a close contact between plasmid and protoplasts.

5) Protoplasts were heat shocked at 37°C for 5 min. and were chilled on ice for 30 min. to allow the plasmid DNA to permeate into them. The tubes were incubated at RT for 30 min.

6) Protoplast plasmid mixture was diluted with MaMg solution so that the PEG concentration reduced to less than 2 %.

7) After proper mixing, the solution was centrifuged at 800 rpm for 5 min. to pellet protoplasts.

8) The supernatant was discarded and 1.0 ml of sterile medium containing Kanamycin (50 ug/ml) was added to maintain selection pressure. The media used for culturing protoplasts were N6 and AA (Table 4.1.1 to 4.1.3) supplemented with hormones.

9) Incubation of tubes at 28°C for 48 hrs. allowed the production of GUS protein.

4.2.6. Extraction and Spectrophotometric Assay of GUS protein

The enzyme β-glucuronidase (GUS) has a monomeric molecular mass of 68 kD and is an exo-hydrolase specific for β-D-glucuronides. It is very stable and tolerates many detergents with widely varying ionic conditions. It is most active in the presence of thiol reducing agents such as β-mercaptoethanol or dithiothreitol (DTT) and can be assayed at any physiological pH with an optimum between 5.0 and 7.8. It is inhibited by glucaric acid-1-4-lactone. It hydrolyses β-linked D-glucuronides to D-glucuronic acid and aglycone.

The introduced gus gene can be monitored in transformed cells by spectrophotometric, fluorometric and histochemical methods (14). The spectrophotometric method is very straightforward, moderately sensitive and can be quantitative without sophisticated instrumentation. It uses para-nitrophenol-β-D-glucuronide as a substrate which is cleaved by the enzyme to liberate the chromophore, para-nitrophenol, the yellow color of which is detected in the visible range of 405 - 420 nm on a spectrophotometer. The protocol followed for the assay is given below :

1) 1.0 ml of Gus extraction buffer (Table 4.1.4) was

TABLE 4.1.4

GUS Extraction Buffer

Components	Final Conc.	Stock Solution	Stock Volume
Sodium phosphate buffer pH 7.0	50 mM	1 M	50 ml
β-mercaptoethanol	10 mM	14.4 M	0.7 ml
Na_2EDTA pH 8.0	10 mM	0.5 M	20 ml
Sarcosyl	0.1 %	10 %	10 ml
Triton X-100	0.1 %	10 %	10 ml
Water	-	-	909 ml
			1000 ml

GUS Assay Buffer

1 mM Para-nitrophenol-β-D-glucuronide (PNPG) dissolved in GUS Extraction Buffer

PEG solution

40 % PEG 6000 in CPW (1 X) and 12 % mannitol for VDN-3630 and 7 % mannitol for Basmati and Indrayani

added to culture medium containing protoplasts and the mixture was sonicated at 160 volts for 15 sec. to lyse protoplasts.

2) It was then centrifuged at 12,000 rpm for 5 min. to remove cell debris.

3) 100 ul of the supernatant was added to 400 ul of pre-warmed substrate (1mM para-nitrophenol-β-D-Glucuronide dissolved in GUS extraction Buffer) and the mixture was incubated at 37°C for 3 hrs to allow GUS enzyme to cleave the substrate and release the product.

4) The reaction was terminated by addition of 0.4 M Na_2CO_3, pH 8.0.

5) Absorbance of the product was measured on a Shimadzu spectrophotometer at 405 nm, the absorption maxima of p-nitrophenol and specific activity of the enzyme was calculated as umoles of p-nitrophenol released/min/10^5 protoplasts.

4.3 RESULTS

4.3.1. Callus induction in Indica varieties

Various combinations of hormones were added to basal MS medium to determine an optimum combination that would induce callus formation in the Indica rices under study (Table 4.2). A concentration of 2.0 mg/l of

2,4-D was found to be optimum for callus initiation while a higher concentration (5.0 mg/l) was inhibitory. Addition of other auxins like NAA (10.0 mg/l) or IAA (5.0 mg/l) did not induce any callus, and led to browning of scutellum. Concentrations of 2.5 mg/l or 0.1 mg/l of kinetin were ineffective in initiating callus in conjunction with 2.0 mg/l 2,4-D. A concentration of 0.5 mg/l kinetin was optimum for callus induction. Use of BAP in concentrations ranging from 0.5 mg/l to 2.5 mg/l failed to elicit any response from the rice seeds. Natural additives like yeast extract and coconut milk were unable to stimulate callus induction in any of the three cultivars. Thus, a combination of MS basal medium + 0.5 mg/l kinetin + 2.0 mg/l 2,4-D+ 100 mg/l inositol was found to be optimum for all three indica rice varieties for callus induction.

Inoculation of dehusked and surface sterilised mature seeds of three Indica varieties on optimised callus medium (Table 4.2) led to the formation of calli after incubation in the dark at 27°C for 4 - 6 weeks. The callus thus formed was regularly subcultured after every 4 weeks. The callus of all three rice varieties showed variable characters. VDN-3630 exhibited a compact, whitish, and embryogenic callus and had a faster multiplication rate (Fig. 4.1). Basmati-370

TABLE 4.2

CONCENTRATION OF GROWTH REGULATORS FOR CALLUS INDUCTION MEDIA (all values are in mg/lit)

Media	I	II	III	IV	V
Kin	0.1	0.5	0.1	0.1	2.5
BAP	0.5	-	0.5	2.5	1.0
CM	10%	-	-	-	-
Inositol	100	100	100	100	100
NAA	10	-	-	-	-
2,4-D	-	2.0	-	5.0	5.0
IAA	-	-	5.0	5.0	5.0
Yeast Extract	-	-	200	-	200
Adenine	-	-	-	1.0	-
Guanine	-	-	-	1.0	-

callus took a longer time to multiply and visual observation showed that the callus was slightly brownish in appearance and compact in nature (Fig. 4.2). Indrayani variety showed a much more friable callus which was whitish in appearance (Fig. 4.3).

4.3.2. Optimization of conditions for protoplast isolation

In order to obtain maximum protoplasts, several parameters had to be optimized for all the three rice varieties. The following parameters were studied in detail :

1) Culture age
2) Cellulase concentration
3) Macerozyme concentration
4) Enzyme mixture volume
5) Incubation conditions
6) pH of the enzyme mixture
7) Different osmotica
8) Osmoticum concentration

The parameter optimization experiments were carried out under non sterile conditions and protoplasts were microscopically observed all the time. Figures 4.4, 4.5 and 4.6 depict isolated protoplasts of VDN-3630, Basmati-370 and Indrayani respectively.

Fig 4.1 25 - day old VDN-3630 callus

FIG.4.1

Fig 4.2 29 - day old Basmati-370 callus

Fig 4.3 21 - day old Indrayani callus

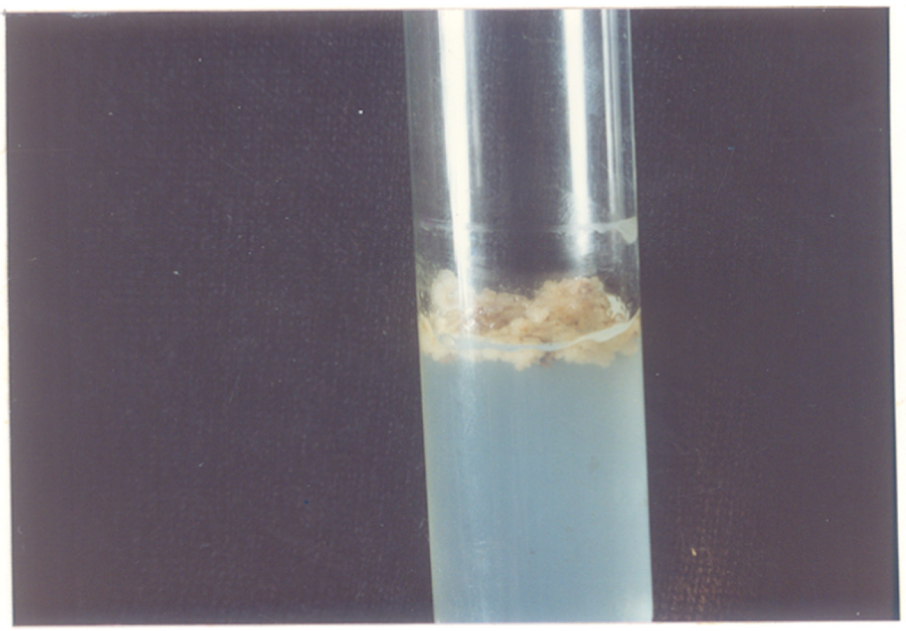


FIG. 4.2

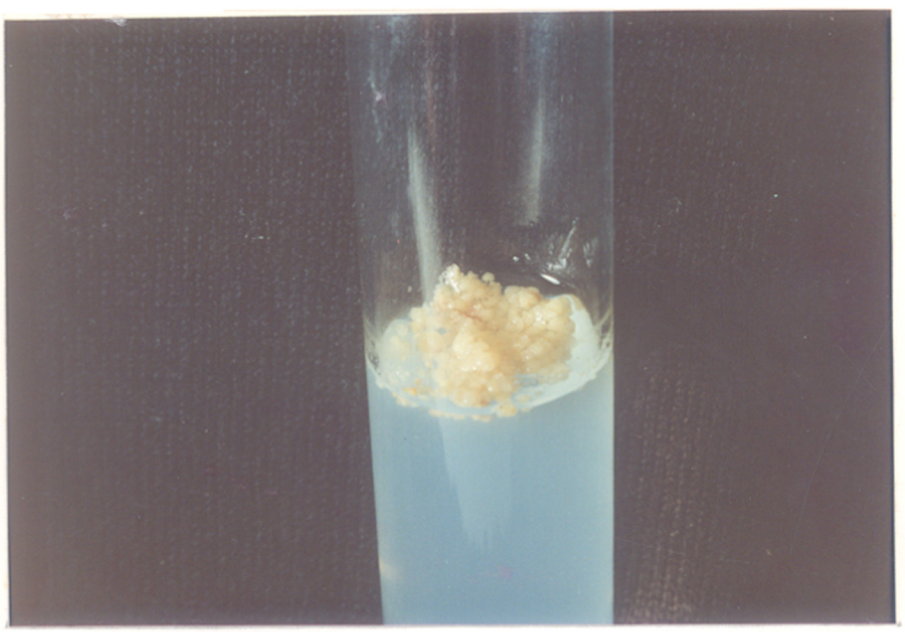


FIG. 4.3

Fig 4.4 VDN-3630 protoplasts

Fig 4.5 Basmati - 370 protoplasts

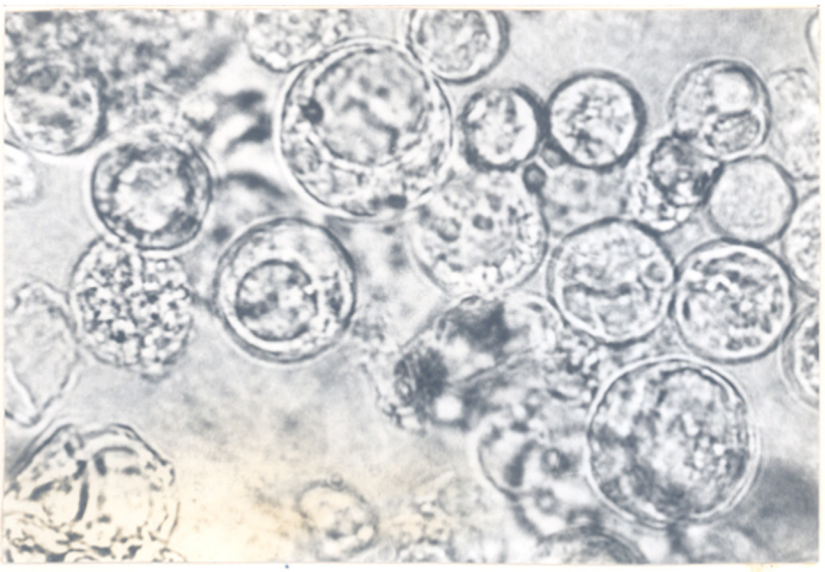


FIG. 4.4

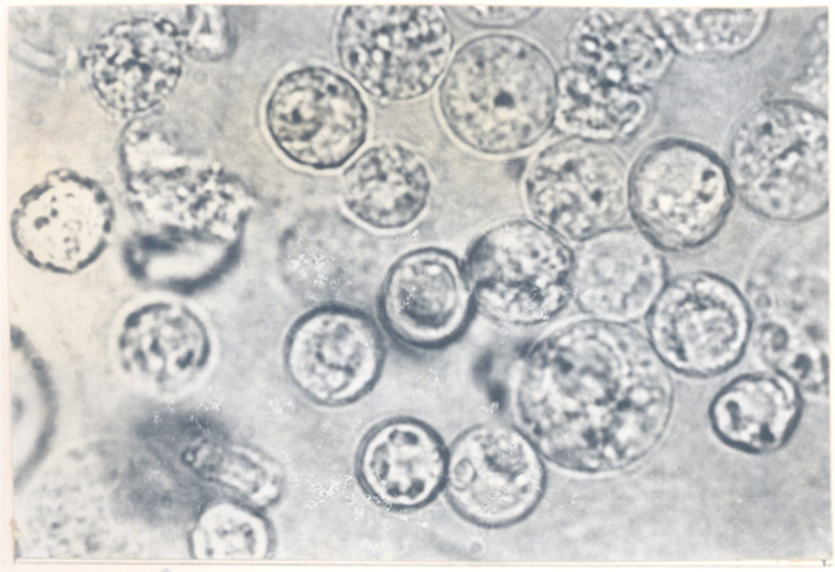


FIG. 4.5

Fig 4.6 Indrayani protoplasts

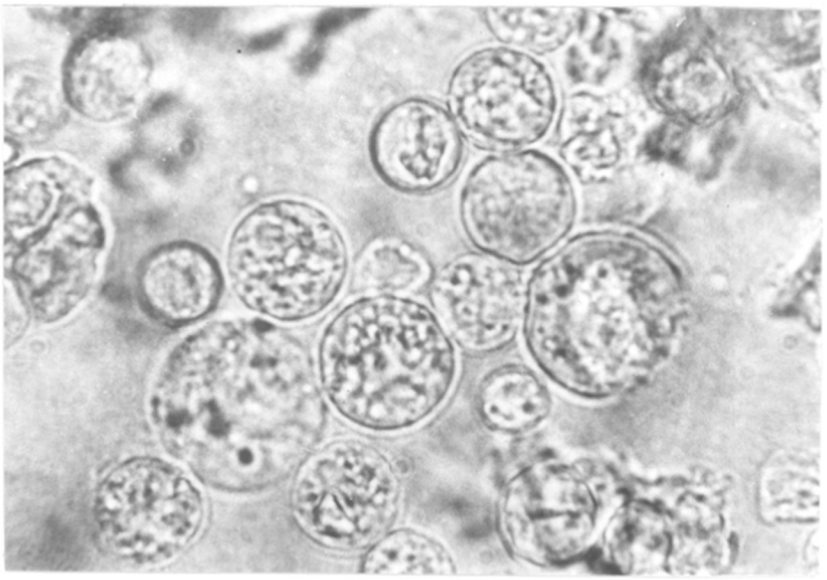


FIG. 4.6

(a) Effect of culture age

The callus of all three rice varieties was subcultured onto fresh medium and protoplasts were isolated from 9th day onwards at 4 day intervals. The results are depicted in Table 4.3.

VDN-3630 callus showed no significant increase in protoplast count (3.97×10^5-4.5×10^5) upto 17 days. However, there was a rapid increase in protoplast formation from 17 days to 25 days and by 29 days, there was a rapid decrease in protoplast count. An optimum yield of 16.15×10^5 protoplast per gram of tissue was obtained from 25 days old callus tissue. Microscopic observations showed the presence of large sized, healthy protoplasts.

In case of protoplast isolation from Basmati-370 callus, a slow but steady increase of protoplasts from 1.19×10^5 to 7.65×10^5 protoplasts/g callus was seen upto 25 days with a rapid doubling (13.4×10^5 protoplasts/g tissue) in count by 29 days. Though protoplast yield increased upto 37 days (18.25×10^5 protoplasts/g tissue), protoplasts were not very healthy.

Protoplasts obtained from Indrayani callus were very small in size. Protoplast count showed little change upto 17 days, increased rapidly by 21 days (7.2

TABLE 4.3

EFFECT OF CULTURE AGE ON PROTOPLAST YIELD

AGE OF CALLUS	:	VARIABLE
CELLULASE ONOZUKA R-10 (W/V)	:	5 %
MACEROZYME R-10 (W/V)	:	1 %
MANNITOL (W/V)	:	12 %
SHAKING PERIOD	:	2 HRS
STATIONARY PERIOD	:	2 HRS
ENZYME MIXTURE VOLUME	:	5.0 mL
WEIGHT OF THE CALLUS	:	1 g

PROTOPLAST YIELD (X 10^5/g CALLUS)

Age	VDN-3630	BASMATI-370	INDRAYANI
9 days	4.0 ± 0.6	1.1 ± 0.0	4.3 ± 0.5
13 days	4.5 ± 1.7	2.6 ± 0.6	4.0 ± 1.7
17 days	3.9 ± 0.5	5.1 ± 0.5	4.2 ± 0.9
21 days	7.2 ± 1.0	7.1 ± 1.0	7.2 ± 1.7
25 days	16.1 ± 1.6	7.6 ± 0.4	4.9 ± 0.3
29 days	5.2 ± 1.2	13.4 ± 1.2	4.0 ± 0.4

x 10^5 protoplasts/g callus) followed by a decline after 25 days.

Thus optimum culture age for isolation of protoplasts from the three rice varieties was 21 days for Indrayani, 25 days for VDN-3630 and 29 days for Basmati-370 callus.

(b) Effect of cellulase concentration

Plant cell wall mainly consists of variable amounts of celluloses, hemicelluloses and pectins which act as a binding material for cells. Thus, in any protoplast isolation protocol, combinations of enzymes viz. cellulases, hemicellulases, and pectinases are used.

In my experiments, cellulase concentration in the enzyme mixture was varied from 1% to 5% and protoplast formation in the three varieties was observed (data presented in Table 4.4). When 25 days old VDN-3630 callus was used to determine the optimum cellulase concentration, there was a rapid increase in protoplast count with increasing concentrations of cellulase. It can be seen from Table 4.4 that maximum protoplasts generated (21.47 X 10^5 /g callus tissue) were with a cellulase concentration of 5%. In Basmati-370, a cellulase concentration of 3 - 5% gave a large number

of protoplasts with 29 days old callus. Microscopic examination showed small size protoplasts with little cell debris with 3% cellulase while 4-5% cellulase concentration showed a larger amount of cell debris. 21 days old Indrayani callus showed a rapid increase in protoplast formation upto 4% cellulase concentration with a subsequent decrease at 5% concentration. Microscopic observation of isolated protoplasts showed that they were small in size. A 4% cellulase concentration generated the maximum amount of protoplasts (10.79×10^5 protoplasts/g callus tissue) with very less amount of cell debris or undigested cells.

In general, cellulase concentration of 3-4% was found to be optimum for the release of protoplasts from Basmati-370 and Indrayani callus while highest yield of protoplasts was obtained with 5% cellulase for VDN-3630 callus.

(c) Effect of macerozyme concentration

Macerozyme R-10, a pectinase, is also used in combination with cellulase. It is isolated from Rhizopus species whose major component is polygalactouronase. Other enzymes that have been used in combination with cellulase are Pectolyase Y-23 and Driselase (17).

TABLE 4.4

EFFECT OF CELLULASE CONCENTRATION ON PROTOPLAST YIELD

CELLULASE ONOZUKA R-10 (W/V)	: VARIABLE
MACEROZYME R-10 (W/V)	: 1 %
MANNITOL (W/V)	: 12 %
SHAKING PERIOD	: 2 HRS
STATIONARY PERIOD	: 2 HRS
ENZYME MIXTURE VOLUME	: 5.0 mL
AGE OF THE CALLUS (IN DAYS)	: VDN-3630 (25), BASMATI-370 (29), INDRAYANI (21)
WEIGHT OF THE CALLUS	: 1 g

PROTOPLAST YIELD (X 10^5/g CALLUS)

CELLULASE CONC.	VDN-3630	BASMATI-370	INDRAYANI
1 %	0.3 ± 0.1	0.1 ± 0.0	0.3 ± 0.0
2 %	1.6 ± 0.3	2.2 ± 0.1	1.3 ± 0.1
3 %	6.3 ± 0.3	14.7 ± 1.8	6.2 ± 0.7
4 %	13.3 ± 1.3	12.6 ± 1.1	10.7 ± 1.9
5 %	21.4 ± 1.8	13.7 ± 1.0	3.2 ± 0.2

To study the effect of macerozyme concentration on protoplast yield of three rice varieties, macerozyme concentration was increased from 0.25% to 1.25% and protoplasts released were counted (Table 4.5).

In VDN-3630 callus, the protoplast yield was maximum at a lower concentration of macerozyme i.e. 0.25 %. Microscopic observations revealed large size protoplasts with small amounts of partially digested cells. At 0.75% to 1.25% macerozyme concentration, 12.80×10^5 protoplasts/g callus tissue were released with a lot of cell debris. In Basmati-370 callus, macerozyme concentration of 0.5% was found to be optimum. In case of Indrayani callus, although 0.25% macerozyme gave the maximum yield, the aamount of cell debris was also more. At 0.5% macerozyme concentration, the protoplast yield was lower but the cell debris was significantly less.

In general, lower macerozyme concentration of 0.25% to 0.5% was found to be capable of releasing the optimum number of protoplasts from the callus of all three rice varieties.

(d) Effect of incubation period

Incubation period of callus and enzyme mixture is known to have a profound effect on protoplast

TABLE 4.5

EFFECT OF MACEROZYME CONCENTRATION ON PROTOPLAST YIELD

CELLULASE ONOZUKA R-10 (W/V)	: VDN-3630 (5%), BASMATI-370 (3%), INDRAYANI (4%)
MACEROZYME R-10 (W/V)	: VARIABLE
MANNITOL (W/V)	: 12 %
SHAKING PERIOD	: 2 HRS
STATIONARY PERIOD	: 2 HRS
ENZYME MIXTURE VOLUME	: 5.0 mL
AGE OF THE CALLUS (IN DAYS)	: VDN-3630 (25), BASMATI-370 (29), INDRAYANI (21)
WEIGHT OF THE CALLUS	: 1 g

PROTOPLAST YIELD (X 10^5/g CALLUS)

MACEROZYME CONC.	VDN-3630	BASMATI-370	INDRAYANI
0.25 %	54.2 ± 3.5	3.8 ± 0.3	17.9 ± 1.1
0.50 %	44.1 ± 6.6	6.9 ± 0.6	15.0 ± 1.9
0.75 %	12.9 ± 0.8	4.3 ± 0.2	15.8 ± 0.8
1.00 %	12.8 ± 3.2	5.6 ± 0.1	8.2 ± 0.3
1.25 %	12.8 ± 0.3	2.3 ± 0.2	13.4 ± 1.7

isolation. In order to determine the optimum incubation period, calli of three varieties were subjected to different periods (0 - 4 hrs) of shaking and stationary phases during protoplast isolation. The total period of incubation including stationary and shaking phases (60 rpm) was 4 hrs. From Table 4.6, it is clear that in case of VDN-3630 callus, a shaking period of 1 hour followed by a stationary period of 3 hours seemed to be optimum for releasing a large number of protoplasts (15.72×10^5/g callus). Increasing the shaking period with a subsequent reduction in stationary period reduced the protoplast count.

A shaking period of 1 and 2 hours followed by a stationary period of 3 and 2 hours, respectively released approximately the same amount of protoplasts (3.7×10^5 per/g callus) in Basmati-370 callus. A stationary period of 4 hours and a shaking period of 3 and 4 hours produced lesser amount of protoplasts. Incubation periods of 2-4 hours shaking gave rise to a large amount of cell debris. In general, It can be seen from Table 4.6 that Basmati-370 callus showed a lesser amount of protoplast release as compared to VDN-3630 callus.

A very high count of protoplasts (6.49×10^5 /g callus tissue) was observed on 4 hours shaking and

TABLE 4.6

EFFECT OF INCUBATION PERIOD ON PROTOPLAST YIELD

CELLULASE ONOZUKA R-10 (W/V)	: VDN-3630 (5%), BASMATI-370 (3%), INDRAYANI (4%)
MACEROZYME R-10 (W/V)	: VDN-3630 (0.25%), BASMATI-370 (0.5%), INDRAYANI (0.25%)
MANNITOL (W/V)	: 12 %
SHAKING PERIOD	: VARIABLE
STATIONARY PERIOD	: VARIABLE
ENZYME MIXTURE VOLUME	: 5.0 mL
AGE OF THE CALLUS (IN DAYS)	: VDN-3630 (25), BASMATI-370 (29), INDRAYANI (21)
WEIGHT OF THE CALLUS	: 1 g

PROTOPLAST YIELD (X 10^5/g CALLUS)

PERIOD (HRS) SHAKING	STATIONARY	VDN-3630	BASMATI-370	INDRAYANI
0	4	9.2 ± 0.9	1.3 ± 0.2	1.2 ± 0.2
1	3	15.7 ± 1.6	3.7 ± 0.3	0.7 ± 0.3
2	2	7.2 ± 1.3	3.7 ± 0.5	1.6 ± 0.3
3	1	7.7 ± 0.7	2.0 ± 0.2	1.5 ± 0.2
4	0	3.4 ± 0.6	2.0 ± 0.2	6.4 ± 1.5

without any stationary incubation periods. But microscopic observation of the protoplasts revealed a large amount of cell debris and partially digested cells at this incubation condition.

In general, incubation periods of one hour shaking, 3 hours stationary; 2 hours shaking, 2 hours stationary and 4 hours shaking without stationary incubation were found optimum for protoplast isolation from VDN-3630, Basmati-370 and Indrayani callus respectively.

(e) Effect of volume of the enzyme mixture on protoplast yield

The objective of this experiment was to study the effect of enzyme mixture volume on protoplast yield. Appropriate concentrations of cellulase and macerozyme were taken and the volume of enzyme mixture was varied from 2.5ml to 12.5 ml keeping the amount of callus tissue at 1g. Protoplasts were enumerated for each volume and the results are tabulated in Table 4.7.

When VDN-3630 callus was incubated in varying volumes of enzyme mixture, the maximum amount of protoplasts (26.46 x 10^5protoplasts/g tissue) was released by 5.0 ml of enzyme mixture. Microscopic observation showed that the amount of cell debris was

TABLE 4.7

EFFECT OF THE VOLUME OF ENZYME MIXTURE ON PROTOPLAST YIELD

CELLULASE ONOZUKA R-10 (W/V)	: VDN-3630 (5%), BASMATI-370 (3%), INDRAYANI (4%)
MACEROZYME R-10 (W/V)	: VDN-3630 (0.25%), BASMATI-370 (0.5%), INDRAYANI (0.25%)
MANNITOL (W/V)	: 12 %
SHAKING PERIOD	: VDN-3630 (1 HR), BASMATI-370 (1 HR), INDRAYANI (2 HRS)
STATIONARY PERIOD	: VDN-3630 (3 HRS), BASMATI-370 (3 HRS), INDRAYANI (2 HRS)
VOLUME OF ENZYME MIXTURE	: VARIABLE
AGE OF CALLUS (IN DAYS)	: VDN-3630 (25), BASMATI-370 (29), INDRAYANI (21)
WEIGHT OF THE CALLUS	: 1 g

PROTOPLAST YIELD (X 10^5/g CALLUS)

Vol (in ml)	VDN-3630	BASMATI-370	INDRAYANI
2.5	15.5 ± 0.2	12.1 ± 1.6	3.3 ± 0.6
5.0	26.4 ± 1.3	2.5 ± 0.1	6.8 ± 0.2
7.5	13.1 ± 1.8	1.4 ± 0.2	2.2 ± 0.2
10.0	9.7 ± 1.4	2.2 ± 0.6	1.3 ± 0.2
12.5	10.5 ± 1.7	1.6 ± 0.3	1.2 ± 0.4

quite high and protoplasts were not healthy. On the other hand, 7.5 ml of enzyme mixture gave half the number of protoplasts but they were healthy in appearance and less cell debris was observed. So 7.5 ml of enzyme mixture for 1g callus tissue was considered as optimum.

2.5 ml of enzyme mixture solution gave a high protoplast count (12.17 x 10^5 protoplasts/g tissue) with Basmati-370 callus. Microscopic observations revealed a large amount of cell debris and a large number of undigested and partially digested cells. Although the protoplasts were high in number, they were not very healthy. There was a drastic reduction in the protoplast count on increasing the volume of enzyme mixture to 5.0 ml. The protoplasts were also healthy and contained lesser amount of cell debris and undigested cells. So for further experiments, 5.0 ml of enzyme mixture for 1g of Basmati-370 callus was considered.

5.0 ml of enzyme mixture gave highest protoplast yield (6.8 x 10^5/g tissue) for Indrayani callus and the protoplasts were healthy in nature.

Depending on surface area of the flask, enzyme mixture volume should be sufficient to submerge the tissue completely. An excess volume seems to have a

deleterious effect on protoplast yield as volumes above 7.5 ml releases a lower number of protoplasts from the tissues.

In general, an enzyme mixture volume of 5.0ml to 7.5ml appeared to be optimum for protoplast isolation from all the three varieties.

(f) Effect of pH of enzyme mixture

The activity of any enzyme depends on an appropriate pH value of the solution. To determine optimum pH value for maximum protoplast yield, pH of enzyme solution used for protoplast isolation was varied from 4.0 to 6.0.

It can be seen from Table 4.8 that pH values of 4.0 to 5.0 generated a high amount of protoplasts and that there was a drastic drop in protoplast formation beyond these pH values of 5.5 and 6.0. Microscopic observation of VDN-3630 protoplasts showed medium sized protoplasts with lots of cell debris.

Basmati-370 callus protoplasts appeared to tolerate a pH range of 5.0 to 5.5. The high count of protoplasts at pH 4.0 was accompanied by a large amount of cell debris and partially digested cells. At pH values of 5.0 and 5.5, there was less cell debris as well as a lower amount of undigested or partially

TABLE 4.8

EFFECT OF pH ON PROTOPLAST YIELD

CELLULASE ONOZUKA R-10 (W/V)	: VDN-3630 (5%), BASMATI-370 (3%), INDRAYANI (4%)
MACEROZYME R-10 (W/V)	: VDN-3630 (0.25%), BASMATI-370 (0.5%), INDRAYANI (0.25%)
MANNITOL (W/V)	: 12 %
SHAKING PERIOD	: VDN-3630 (1 HR), BASMATI-370 (1 HR), INDRAYANI (2 HRS)
STATIONARY PERIOD	: VDN-3630 (3 HRS), BASMATI-370 (3 HRS), INDRAYANI (2 HRS)
ENZYME MIXTURE VOLUME	: VDN-3630 (7.5 mL), BASMATI-370 (5.0 mL), INDRAYANI (5.0 mL)
AGE OF THE CALLUS (IN DAYS)	: VDN-3630 (25), BASMATI-370 (29), INDRAYANI (21)
WEIGHT OF THE CALLUS	: 1 g

PROTOPLAST YIELD (X 10^5/g CALLUS)

pH OF ENZ. MIX	VDN-3630	BASMATI-370	INDRAYANI
4.0	46.5 ± 4.6	16.1 ± 1.5	28.3 ± 4.9
4.5	33.8 ± 2.7	5.8 ± 0.2	37.1 ± 2.1
5.0	49.3 ± 2.6	10.1 ± 0.1	17.8 ± 0.4
5.5	0.4 ± 0.0	16.5 ± 0.1	11.7 ± 0.8
6.0	1.6 ± 0.1	7.0 ± 0.8	7.5 ± 0.4

digested cells.

Indrayani callus also released a large number of protoplasts at a lower pH range of 4.0 to 5.0 (17.85×10^5- 37.13×10^5). At higher pH values, protoplasts revealed lower amounts of undigested or partially digested cells and the cell size also seemed to be varying between small to medium.

The high protoplast count at pH 4.0 for Basmati-370 and VDN-3630 could be due to the large amount of cellular debris appearing as spherical protoplasts giving rise to aberrant counts. Among the three rice varieties, VDN-3630 protoplasts seem to be unable to withstand higher pH values of 5.5 and 6.0 as seen from the extremely low counts.

Overall, lower pH values of 4.0 to 5.0 were preferable for protoplast isolation from callus cultures of VDN-3630 and Indrayani; whereas for Basmati-370 the optimum yield of protoplasts was obtained at a pH between 5.0 to 5.5.

(g) Effect of different osmotica

Three different osmotica, namely, Glucose, Sucrose and Mannitol were tried and protoplasts were isolated at a 12% concentration of each osmotica.

TABLE 4.9

EFFECT OF DIFFERENT OSMOTICA ON PROTOPLAST YIELD

CELLULASE ONOZUKA R-10 (W/V)	: VDN-3630 (5%), BASMATI-370 (3%), INDRAYANI (4%)
MACEROZYME R-10 (W/V)	: VDN-3630 (0.25%), BASMATI-370 (0.5%), INDRAYANI (0.25%)
OSMOTICUM (W/V)	: 12 %
SHAKING PERIOD	: VDN-3630 (1 HR), BASMATI-370 (1 HR), INDRAYANI (2 HRS)
STATIONARY PERIOD	: VDN-3630 (3 HRS), BASMATI-370 (3 HRS), INDRAYANI (2 HRS)
ENZYME MIXTURE VOLUME	: VDN-3630 (7.5 mL), BASMATI-370 (5.0 mL), INDRAYANI (5.0 mL)
AGE OF THE CALLUS (IN DAYS)	: VDN-3630 (25), BASMATI-370 (29), INDRAYANI (21)
WEIGHT OF THE CALLUS	: 1 g

PROTOPLAST YIELD (X 10^5/g CALLUS)

Osmoticum	VDN-3630	BASMATI-370	INDRAYANI
Glucose	3.1 ± 0.5	3.2 ± 0.3	8.8 ± 0.4
Sucrose	1.9 ± 0.1	1.8 ± 0.4	8.9 ± 1.2
Mannitol	5.8 ± 0.6	5.1 ± 0.5	16.0 ± 1.4

It can be seen from Table 4.9 that in the presence of glucose and sucrose, the amount of protoplasts released was less as compared to those in mannitol. In the presence of mannitol, Indrayani callus gave the highest amount of protoplasts (16.0×10^5/g callus). The protoplasts isolated from all the three calli with mannitol as osmoticum were healthy and dense as compared to those isolated with glucose and sucrose as osmotica.

It can be concluded from Table 4.9 that mannitol is the best osmoticum for protoplast isolation for all three varieties.

(h) Effect of mannitol concentration

In this experiment, the objective was to determine optimum mannitol concentration to be used in the enzyme solution for optimum release of protoplasts from three varieties. Mannitol concentration was varied between 7% to 15% i.e. from 0.4M to 0.8M and its effect on protoplast yield was observed. Results of this experiment are depicted in Table 4.10.

VDN-3630 callus gave highest yield of protoplasts with a mannitol concentration of about 11% i.e. 0.6M. The amount of cell debris and undigested cell number was less and the protoplasts released were very

healthy in nature. There was a significant reduction in protoplast yield at higher mannitol concentrations.

Lower concentrations of mannitol i.e. 0.4 -0.5 M was observed to be very appropriate for protoplasts isolated from Basmati-370 callus. There was a gradual reduction in protoplast yield with increasing mannitol concentration. A very high protoplast yield was observed at 0.8 M mannitol concentration; however the protoplasts were not healthy in appearance.

Indrayani callus gave a higher protoplast yield at 0.4 M mannitol concentration. The protoplast yield was approximately same (6×10^5/g callus) from 0.5 - 0.7 M mannitol concentration. Protoplasts were smaller in size and variable amounts of cell debris were observed. At 0.7 M mannitol concentration, cells were very healthy in appearance with a lesser amount of cell debris.

Thus it can be seen from Table 4.10 that VDN-3630 callus and Indrayani callus gave high protoplast yields at 0.6 M (11%) and 0.7 M (12.5%) mannitol concentration, respectively; while Basmati-370 callus gave a higher protoplast yield at 0.4-0.5 M (7 - 9%) mannitol.

TABLE 4.10

EFFECT OF MANNITOL CONCENTRATION ON PROTOPLAST YIELD

CELLULASE ONOZUKA R-10 (W/V)	: VDN-3630 (5%), BASMATI-370 (3%), INDRAYANI (4%)
MACEROZYME R-10 (W/V)	: VDN-3630 (0.25%), BASMATI-370 (0.5%), INDRAYANI (0.25%)
MANNITOL (W/V)	: VARIABLE
SHAKING PERIOD	: VDN-3630 (1 HR), BASMATI-370 (1 HR), INDRAYANI (2 HRS)
STATIONARY PERIOD	: VDN-3630 (3 HRS), BASMATI-370 (3 HRS), INDRAYANI (2 HRS)
ENZYME MIXTURE VOLUME	: VDN-3630 (7.5 mL), BASMATI-370 (5.0 mL), INDRAYANI (5.0 mL)
AGE OF THE CALLUS (IN DAYS)	: VDN-3630 (25), BASMATI-370 (29), INDRAYANI (21)
WEIGHT OF THE CALLUS	: 1 g

PROTOPLAST YIELD (X 10^5/g CALLUS)

Molarity	VDN-3630	BASMATI-370	INDRAYANI
0.4 M	4.4 ± 0.5	6.7 ± 0.8	9.5 ± 0.2
0.5 M	5.9 ± 0.9	6.6 ± 0.5	6.5 ± 0.2
0.6 M	7.7 ± 0.8	3.0 ± 0.5	6.0 ± 0.6
0.7 M	1.7 ± 0.3	2.1 ± 0.1	6.9 ± 0.9
0.8 M	0.9 ± 0.4	8.2 ± 1.7	4.8 ± 1.0

4.3.3. Transient gene expression in indica rice protoplasts

Transient gene expression is now widely used as a tool to optimise DNA delivery conditions as well as to examine functional importance of particular DNA sequences for gene expression. After a few days or even hours the gene transferred to plant cells can be monitored by assaying for reporter enzyme expression. This short time interval between the start of experiment and actual enzyme assay makes transient gene expression advantageous in comparison to regeneration of stably transformed plants.

In the present work, a plasmid pBI121 containing gus and nptII genes was introduced into indica rice protoplasts via PEG method. Expression of GUS gene was monitored spectrophotometrically by assaying the chromogenic product para-nitrophenol at 405 nm. The GUS activity was measured in terms of umoles of PNPG/min/10^5 protoplasts. GUS expression was not obtained when protoplasts were subjected to a heat shock of 42°C for 1 min; whereas detectable GUS activity was observed when the protoplasts were incubated at 37°C for 5 min. From the results in Table 4.11, it can be seen that Basmati-370 protoplasts

TABLE 4.11

ACTIVITY OF GUS IN INDICA PROTOPLASTS

umoles of PNP/min/10^5 protoplasts

DNA Conc (in ug)	BASMATI-370	VADGAON-3630	INDRAYANI
- DNA (0 ug)	118.60 ± 8.07	14.94 ± 1.43	21.99 ± 1.72
+ DNA (30 ug)	169.60 ± 8.79	61.55 ± 0.36	28.37 ± 1.81
Fold increase	1.41	4.32	1.29

Legends :

- DNA : Control

+ DNA : Transformed

$$\text{GUS activity (G)} = \frac{\text{O.D.}}{\text{H x V x E}} \text{ umoles/min}$$

$$\text{GUS activity umoles/min/}10^5 = \text{G / PC}$$

where :

G = GUS activity in umoles/min

O.D. = Optical Density

H = Reaction Time (in min.) = 180 min.

V = Sample Volume = 0.1 ml

E = Extinction Coefficient = 14,000

PC = Protoplast Count

showed a higher GUS expression than VDN-3630 or Indrayani protoplasts in terms of absolute values. The background values of untransformed Basmati-370 protoplasts were also higher than those of untransformed VDN-3630 and Indrayani protoplasts. But VDN-3630 showed a 4.3 fold expression of GUS activity than Basmati-370 or Indrayani.

In general, the protoplasts of three indica varieties showed a varied pattern of expression of gus gene it being highest (4.3 fold) in VDN-3630.

4.4 DISCUSSION

One of the primary steps in development of a transformation system in rice is selection of an appropriate source material that can give rise to a high protoplast yield. In the present work, mature seeds have been used for callus induction though other materials like anthers (18, 6), roots (19) and immature embryos (20) have also been used.

The choice of basal medium for initiation of callus is an important factor. The most commonly used basal medium for rice is Murashige and Skoog's Medium (21, 22). Other media like Linsmaier and Skoog's medium (23); N6 medium (7), modified White's medium (20) and

AA2 medium (24) have also been used. Zapata et al. (25) have used Gamborg's B5 medium for induction of callus from anthers of Taipei 309. In the present investigation, MS medium containing 2.0 mg/lit 2,4 - D, 0.5 mg/lit kinetin, and 100 mg/lit meso-inositol has been used for induction and maintenance of callus and this has given us satisfactory results. The rate of multiplication of callus on the medium varies with cultivar. The VDN-3630 callus has a faster multiplication rate than Basmati-370 or Indrayani callus. The Indrayani variety has a much more friable callus than VDN-3630 and Basmati callus.

The successful isolation of protoplasts depends upon various factors like i) source material ii) protoplast isolation conditions, and iii) media composition. Though Lee et al. (26) have stated that continuous subculturing of suspension culture leads to loss of regenerative potential of the cell line, most of the available reports have used suspension culture for isolation of protoplasts. Our attempts to obtain a rapidly dividing suspension culture have not been fruitful and hence protoplasts were isolated from callus material.

Protoplasts can be isolated from various parts of the plants such as stem, leaf, and root. Conditions for obtaining maximum number of protoplasts differ

from species to species. Table 4.12 summarises the optimum conditions for a maximum yield of protoplasts from three Indica varieties. It can be seen from Table 4.12 that each variety requires different enzyme concentrations of cellulase and macerozyme pointing towards the differences in their cell wall compositions.

The period of contact of enzymes with callus tissue is also critical in obtaining a high yield of protoplasts. A higher enzyme concentration will allow a shorter contact time between tissue and enzyme solution. Baset and Cocking (1991) (23) have used a longer incubation period (18 hrs) with a lower concentration of cellulase (0.25 %) and Pectolyase (0.025%). On the other hand, Jenes and Pauk (1989) (27) have used a higher enzyme concentration of cellulase (6.0 %), Pectinase (2.0 %) and Driselase (2.0 %) and incubated the tissue only for a period of 3 - 3.5 hrs on the shaker. VDN-3630, Basmati-370 and Indrayani require increasing periods of shaking of 1 hr, 2 hrs and 4 hrs respectively. These periods of shaking have to be followed by a stationary phase of 3 hrs and 2 hrs, for VDN-3630 and Basmati-370, respectively. A stationary phase facilitates the release of protoplasts without damage. Our data indicates that VDN-3630 cells

TABLE 4.12

OPTIMISED PROTOPLAST ISOLATION CONDITIONS FOR THREE INDICA VARIETIES

PROTOPLAST ISOLATION CONDITIONS	BASMATI -370	VADGAON -3630	INDRA-YANI
Culture Age	29 Days	25 Days	21 Days
Cellulase concentration	3 %	5 %	4 %
Macerozyme concentration	0.5%	0.25%	0.25%
Incubation period shaking stationary	2 HRS 2 HRS	1 HRS 3 HRS	4 HRS -
Enzyme mixture volume for 1g callus	5.0 ml	7.5 ml	5.0 ml
pH	5.5	5.0	4.5
Osmotica	Mannitol	Mannitol	Mannitol
Osmoticum concentration	0.4 M (7 %)	0.7 M (12 %)	0.4 M (7 %)

may be more sensitive to the enzymes than Basmati-370 or Indrayani cells.

The nature of osmoticum and its concentration play a very important role in isolation of protoplasts from any plant. Initially, inorganic salts such as NaCl, $NaNO_3$ and $CaCl_2$ were used as osmoticum during protoplast isolation (28). These were ionic in nature and found to be detrimental to protoplast viability. Later, researchers used nonionic sugars such as mannitol, sucrose and sorbitol as osmoticum. Mannitol is a nonionic and non-nutritive carbohydrate that seems to be the best osmoticum for isolation of protoplasts from rice. Mannitol may be creating an appropriate osmotic potential in a much better fashion than glucose or sucrose. In one report by Yamada et al. (1986) (29) 0.3 M Glucose has also been used. Various reports have shown that different rice cultivars have different levels of osmotolerance. Taipei 309 can tolerate 0.7 M mannitol (30) while Nipponbare, Iwaimochi, Norin 14 and Fujisaka 5 can tolerate only upto 0.4 M mannitol (3). Our investigations have also shown that mannitol is the best osmoticum for rice cells. The three indica varieties exhibit differences in their ability to tolerate mannitol concentrations. VDN-3630 tolerates upto 0.7 M mannitol concentration while optimum protoplast yields from Basmati-370 and

Indrayani callus is around 0.4 M.

Various protocols have been attempted in order to obtain transient gene expression in indica protoplasts. There was no expression in any variety when the protoplasts were suspended in CPW salts and used for transformation studies. On suspending the protoplasts in MaMg solution prior to the addition of DNA and PEG, GUS activity was observed in all three varieties. This result agrees with the observation of Negritiu (1987) (31) that Mg^{++} ion plays a significant role in uptake of plasmid DNA by protoplasts. The detection of transient gene expression in spite of the use of a slightly lesser sensitive spectrophotometric assay clearly indicates that transformation has indeed occurred. Results presented in Table 4.11 show that VDN-3630 protoplasts are in a more competent state than Basmati-370 or Indrayani protoplasts. This rather agrees with the faster multiplication rate of VDN-3630 callus. The Table also indicates that Basmati-370 protoplasts have a higher background activity as compared to VDN-3630 or Indrayani protoplasts. This means that Basmati-370 protoplasts contain a higher level of endogenous GUS-like activity which can interfere with the activity originating from introduced GUS gene. Kosugi et al. (1990) (32) have demonstrated

that the Oryza sativa cv. Musashikogane callus has significant levels of endogenous GUS.

Transient gene expression was detected only when protoplasts were cultured under a selection pressure of Kanamycin at 50ug/ml concentration in the culture medium. According to Peng and Hodges (1992) (33), presence of kanamycin in the culture medium may not allow regeneration of transformed protoplasts. Therefore transformed protoplasts have to be cultured in presence of other selectable markers like G418 in order to obtain regeneration.

Thus the three indica cultivars vary in their response to callus formation, protoplast isolation conditions and transient gene expression. The above investigation is substantial enough to provide a general outline for optimising protoplast isolation and transformation conditions in indica rice varieties.

4.5 REFERENCES

1) Zhang W. and Wu R. (1988)
Theor. Appl. Genet. **76** : 835 - 840.

2) Fujimura T., Sakurai M., Akagi R., Negishi T., and Hirose A. (1985)
Plant Tissue Cult. Lett. **18** : 659 - 668.

3) Kyozuka J., Hayashi Y. and Shimamoto K. (1987)
Mol. Gen. Genet. **206** : 408 - 413.

4) Shimamoto K., Terada R., Izawa T., and Fujimoto (1989)
Nature **338** : 274 - 276.

5) Wang D., Miller P.P. and Sondahl M.R. (1989)
Plant Cell Rep. **8** : 329 - 332.

6) Datta S.K., Peterhans A., Datta K. and Potrykus I. (1990)
Bio/Technology **8** : 736 - 740.

7) Peng J. and Hodges T.K. (1989)
In vitro Cell. and Dev. Biol. **25(1)** : 91 - 94.

8) Gupta H.S. and Pattanayak A. (1993)
Bio/Technology **11** : 90 - 94.

9) Christou P., Ford T.L. and Kofron M. (1991)
Bio/Technology **9** : 957 - 962.

10) Murashige T. and Skoog F. (1962)
Physiol. Plant. **15** : 473 - 497.

11) Chu C.C., Wang C.C., Sun C.S., Yin K.C., Chu C.Y., Bi F.Y. (1975)

Sci. Sin. **18** : 659 - 668.

12) Mueller A.J. and Grafe R. (1978)
Mol. Gen. Genet. **161** : 67 -76.

13) Krens F.A., Molendijk K.L., Wullems G.J., Schilperoort R.A. (1982)
Nature **296** : 72 - 74.

14) Jefferson R.A. (1987)
Plant Mol. Biol. Rep. **5** : 387 - 405.

15) Birboim and Doly (1979)
Nucl. Acid Res. **7** : 1513.

16) Hayashimoto A., Li Z.J., and Murai N. (1990)
Plant Physiol. **93** : 857 - 863.

17) Coulibaly M.Y. and Demarly Y. (1986)
Z. Pflanzenzucht **96** : 79 - 81.

18) Reiffers I. and Friere A.B. (1990)
Plant Cell, Tissue & Organ Culture **21** : 165 - 170.

19) Abe T. and Futsuhara Y. (1989)
Theor. Appl. Genet. **78** : 648 - 652.

20) Wang Y.L., Klein T.M., Fromm M., Cao J., Sanford J.C., and Wu R. (1988)
Plant. Mol. Biol. **11** : 433 - 439.

21) Peterson G. and Smith R. (1991)
Plant Cell Rep. **10** : 35 - 38.

22) Oard J.H., Paige D.F., Srmmonds J.A. and Gradziel T.M. (1990)
Plant Physiol. **92** : 334 - 339.

23) Baset A., Finch R.P. and Cocking E.C. (1991)
Plant Cell Rep. **10** : 200 - 203.

24) Thompson J.A. Abdullah R. and Cocking E.C. (1986)
Plant Sci. **47** : 123 - 133.

25) Cho M.S. and Zapata F.J. (1988)
Plant Sci. **58** : 239 - 244.

26) Lee L., Schroll R.E., Grimes H.D., and Hodges T.K. (1989)
Planta **178** : 325 - 333.

27) Jenes B. and Pauk J. (1989)
Plant Sci. **63** : 187 - 198.

28) Kemeya T. and Uchimiya H. (1972)
Planta **103** : 356 - 360.

29) Yamada Y., Zhi-Qi Y. and Ding-Tai T. (1986)
Plant Cell. Rep. **5** : 85 - 88.

30) Abdullah R., Cocking E.C. and Thompson J.A. (1986)
Bio/Technology. **4** : 1087 - 1090.

31) Negritiu I., Shillito R., Potrykus I., Biasini G., and Sala F. (1987)
Plant Mol. Biol. **8** : 363 -373.

32) Kosugi S., Ohashi Y., Nakajima K and Arai Y. (1990)
Plant Sci. **70** : 133 - 140.

33) Peng J., Kononowicz H., and Hodges T.K., (1992)
Theor. Appl. Genet. **83** : 855 - 863.

CHAPTER V

GENERAL CONCLUSIONS

Rice is one of the most widely cultivated cereals in the world and more than half of the world's population consumes rice as their staple food. Thus it is only natural that it should become a crop of great importance from a commercial as well as fundamental point of view.

Developmentally regulated genes are prime candidates for the study of basic plant mechanisms. Seed storage protein genes are expressed in a tissue-specific fashion with different levels of expression. Rice is one of the unique cereals which has glutelins as their major storage proteins and prolamins form only a minor fraction. Copy number estimation of these genes has revealed the occurrence of 8 - 10 copies of glutelin and 80 - 100 copies of prolamin genes per haploid genome of rice. Information regarding the mechanism of regulation of prolamin genes in rice is very limited. In fact, isolation of genomic clones of prolamin has been very sparse. Moreover, the disparity between high copy number and low expression levels of prolamin genes makes it an interesting system for investigation. In an attempt to attain this objective, molecular as well as tissue culture approaches were utilised. In the molecular approach, an attempt was also made to understand the regulation of storage

protein genes using a theoretical approach.

The first step towards isolation of prolamin genes from an elite indica cultivar of rice namely, Basmati-370 was the construction of a genomic library of rice in a replacement vector namely, Charon 4A. One of the critical steps in genomic library construction was preparation of a packaging extract. Though commercial preparations were available, packaging extracts lose their viability when stored for a long time. Hence I prepared the packaging extract from an E. coli SMR-10 culture. The library was screened for prolamin clones with the help of a heterologous probe namely, kafirin, a prolamin gene from S. bicolor. Kafirin was used as heterologous probe since homologous rice prolamin probes were not yet available at the time of initiating this work. This probe enabled me to identify three putative clones of prolamin. Another advantage of construction of this library is that it can also be used to isolate any other genes of interest from rice.

Parallel to the above work, I undertook a study of the 5' upstream regions of storage protein genes in order to identify regions that could be playing an important role in their expression. Free energy of these regions will indicate their ability to easily "open up" for transcription or translation. A lower

free energy implicates a lesser tendency to form secondary structures. Interestingly, free energy profiles of monocot and dicot storage protein genes were distinctly different. An additional peak exhibited in the free energy profile of monocot genes could probably indicate that monocot and dicot storage protein genes have different regulatory mechanisms. But these studies need to be explored further, both theoretically and experimentally, in order to reach any definitive conclusion. The consensus sequences of monocot and dicot genes also differ significantly. The number of conserved nucleotides in monocots is higher than in dicots which could probably be related to a more evolved nature of monocots as compared to dicots. There is a higher number of A and T nucleotides in the region preceding the translation start site which also proves that this region has a lesser tendency to form secondary structures as combinations of A and T dinucleotides have a lower free energy value. This is shown by the common unstable peak present -35 bp upstream of the translation start site in monocot and dicot storage protein genes. An indepth analysis of monocot and dicot upstream regions could also provide a clue towards their differential levels of expression. These studies could also help in identifying regions that may be influencing the regulation of storage

protein genes.

Once the seed storage protein genes were isolated, it was necessary to have a system in which its expression could be monitored. Since the prolamin genes were isolated from an indica variety of rice namely, Basmati-370, a gene expression system of an indica variety had to be optimised. The regeneration of japonica rice protoplasts is well established and is now possible using routine procedures. Such routine procedures are not yet available for indica varieties. Hence it was necessary to optimise conditions for callus initiation and protoplast isolation from indica cultivars. During optimisation, differences in isolation conditions for the three indica cultivars were noted. In order to assess their ability to express introduced DNA, protoplasts were transformed with control DNA. A plasmid containing the reporter genes gus and nptII was used for transformation of these protoplasts. The PEG method of transformation was preferred since it caused minimal injury to protoplasts. Though the protoplasts of all the three varieties did express GUS activity, they had different levels of background GUS-like activity.

Now further studies can be carried out in both molecular and tissue cultural aspects. Firstly, the

isolated putative prolamin clones can be further subcloned into smaller fragments in order to identify 5' upstream regions. These regions can then be fused to reporter genes like gus or nptII and their transient gene expression can be monitored in indica rice protoplasts. Secondly, conditions for regeneration of indica protoplasts can be optimised.